

STATISTICAL LEARNING IN A BILINGUAL ENVIRONMENT

by

Sin Mei Tsui

BEcon&Fin., University of Hong Kong, 2008

PCPsych, University of Hong Kong, 2011

THESIS SUBMITTED TO

THE UNIVERSITY OF OTTAWA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

in

EXPERIMENTAL PSYCHOLOGY

SCHOOL OF PSYCHOLOGY

Faculty of Social Science

University of Ottawa

© Sin Mei Tsui, Ottawa, Canada, 2018

Abstract

Statistical learning refers to the ability to track regular patterns in sensory input from ambient environments. This learning mechanism can exploit a wide range of statistical structures (e.g., frequency, distribution, and co-occurrence probability). Given its regularities and hierarchical structures, language is essentially a pattern-based system and therefore researchers have argued that statistical learning is fundamental to language acquisition (e.g., Saffran, 2003). Indeed, young infants and adults can find words in artificial languages by tracking syllable co-occurrence probabilities and extracting words on that basis (e.g., Saffran, Aslin & Newport, 1996a). However, prior studies have mainly focused on whether learners can statistically segment words from a single language; whether learners can segment words from two artificial languages remains largely unknown. Given that the majority of the global population is bilingual (Grosjean, 2010), it is necessary to study whether learners can make use of the statistical learning mechanism to segment words from two language inputs, which is the focus of this thesis. I examined adult and infant learners to answer three questions: (1) Can learners make use of French and English phonetic cues within a single individual's speech to segment words successfully from two languages?; 2) Do bilinguals outperform monolinguals?; and 3) Do specific factors, such as cognitive ability or bilingual experience, underlie any potential bilingual advantage in word segmentation across two languages?

In Study 1, adult learners generally could make use of French and English phonetic cues to segment words from two overlapping artificial languages. Importantly, simultaneous bilinguals who learned French and English since birth segmented more correct words in comparison to monolinguals, multilinguals, and sequential French-English bilinguals. Early bilingual experience may lead learners to be more flexible when processing information in new environments and/or

they are more sensitive to subtle cues that mark the changes of language inputs. Further, individuals' cognitive abilities were not related to learners' segmentation performance, suggesting that the observed simultaneous bilingual segmentation advantage is not related any bilingual cognitive advantages (Bialystok, Craik, & Luk, 2012).

In Study 2, I tested 9.5-month-olds, who are currently discovering words in their natural environment, in an infant version of the adult task. Surprisingly, monolingual, but not bilingual, infants successfully used French and English phonetic cues to segment words from two languages. The observed difference in segmentation may be related to how infant process native and non-native phonetic cues, as the French phonetic cues are non-native to monolingual infants but are native to bilingual infants. Finally, the observed difference in segmentation ability was again not driven by cognitive skills.

In sum, current thesis provides evidence that both adults and infants can make use of phonetic cues to statistically segment words from two languages. The implications of why early bilingualism plays a role in determining learners' segmentation ability are discussed.

Table of Contents

Abstract	ii
Table of Contents	iv
Acknowledgements	vi
Declaration of Academic Achievement	x
General Introduction	1
The Role of Statistical Learning in Language Segmentation Problem	2
Word Segmentation with Two Input Streams	6
Does a Bilingual Advantage in Dual-input Statistical Learning Tasks Exist?.....	13
Overview of Current Thesis and Hypotheses.....	30
References	37
Study 1: Statistical Word Segmentation from Accented Speech: A Bilingual Advantage...	58
Experiment 1	69
Method.....	69
Results	80
Discussion.....	84
Experiment 2A	86
Method.....	87
Results and Discussion	88
Experiment 2B.....	88
Method.....	89
Results and Discussion	90
General Discussion.....	93
References	101
Table 1. Participants' Language Backgrounds in Experiment 1	113
Table 2. Auditory Stimuli in the Language Segmentation Task.....	114
Table 3. Multiple Regression Models in Experiment 1	115
Figure 1. Transitional Probability Between Syllables in Experiment 1	116
Figure 2. Experimental Procedure of the Attention Network Task (ANT).....	117
Figure 3. Participants' Performance in Experiment 1	118
Study 2: Infant Statistical Segmentation in a Language-mixing Bilingual Environment..	119
Experiment 1	127

Method.....	128
Results and Discussion	132
Experiment 2	134
Method.....	135
Results and Discussion	137
Experiment 3	139
Method.....	139
Results and Discussion	140
Experiment 4	141
Method.....	142
Results	149
Discussion.....	150
General Discussion.....	151
Conclusion.....	160
References	161
Table 1. Auditory Stimuli in the Language Segmentation Task.	170
Table 2. Acoustic Measurement of the English- and French-Accented Speech	171
Table 3. Hierarchical Linear Regression Models.....	172
Figure 1. Transitional Probability Between Syllables in Experiment 2 and 3	173
Figure 2. Infants' Segmentation Performance in Experiment 1, 2 and 3	174
Figure 3. Procedure of the Anticipatory Eye Movement Paradigm in Experiment 4	175
Figure 4. Infants' Performance in Experiment 4.....	176
General discussion	177
Can Phonetic cues (i.e., Accents) help Learners Segment Words from Two Overlapping Languages?	178
Do Bilinguals Outperform Monolinguals when Segmenting Words from Two Overlapping Languages? If so, What Explains this Bilingual Advantage?	181
Implication of the Relationship Between Cognitive Ability and Statistical Segmentation	186
Limitations and Future Directions.....	189
Concluding Statement	192
References	193
Appendix A: Adult Language Background Questionnaire.....	198
Appendix B: Parental Questionnaire about Infant Language Exposure	206

Acknowledgements

This thesis marks the completion of my journey in the Ph.D. program and also the beginning of my career as a researcher. I am truly grateful to the many people who generously have helped and supported me over the past six years.

First and foremost, I would like to thank Dr. Christopher Fennell, for being the best doctoral supervisor in my career. Thank you for your invaluable guidance and enormous patience. I am hugely benefited from your excellent advice, constructive and inspiring discussions throughout the years. From you, I learned how to conduct research with high standards, how to write well in English, how to present research (both in text and speech) clearly at a level for readers with different backgrounds, and how to think critically when evaluating my own and other researchers' work. As a great role model, you have also demonstrated strong leadership and mentorship skills. You have deepened my research interest in infant bilingualism, strengthened my enthusiasm and facilitated my growth as a researcher. For this, I am forever thankful and I aim to repay this by expanding the knowledge that you have passed to me in future!

I am indebted to my thesis committee members, Drs. Cristina Atance, Alain Desrochers, and Laura Sabourin for dedicating their time and energy to my thesis, and for their insightful feedback and suggestions. In addition, I want to express gratitude for the following.

Cristina, thank you for the great opportunity to work in your lab. The saving marble game continues to be one of my favorite experiments and I hope I can further extend this research with you in future. I feel so fortunate to work with you as you are a genuine, kind and dedicated mentor.

Alain, thank you for being a great teacher of cognitive psychology, particularly in the course I took from you. I learned so much about the human memory processes. Importantly, I have

learned how to better present research orally from my direct involvement and indirect observation in your class.

Laura, thank you for your generosity of lending space in your lab for my data collection. In your lab, I have widened my horizon about adult bilingualism. I am also very grateful for your continuous support and encouragement, and the exciting opportunity to work with you and Santa.

During my Ph.D, I was very fortunate to meet with other mentors who have tremendous influence in my career. Dr. Veronika Huta has contributed much to my intelligence and knowledge in statistics. She is one of the most influential statistics teachers in my career (and my soulmate in statistics). I am very grateful for her genuine caring mentorship and friendship, which supports me emotionally and academically. Dr. Patricia Brosseau-Liard has been a wonderful mentor who genuinely supports me in further understanding statistics. I very much appreciate for the learning opportunity about simulations, great tips about oral presentations and her insightful comments on my work. In addition, I want to thank my collaborators. Dr. Krista Byers-Heinlein has been extremely supportive in my meta-analysis paper and research. She provides me with strong intellectual support and discussions, further pushing me to dig deeper in my research. Dr. Laurel Fais has been very generous with her time and her prompt feedback of my research project. She is also very friendly and nice. Dr. Lucy Erickson and Dr. Erik Thiessen are wonderful colleagues, giving me a number of great feedback to fine-tune the experiments in this thesis. Their expertise is invaluable and they have directed me toward relevant literature in the world of statistical learning.

I am very grateful to my good friends at Ottawa. Your friendship has been an important support over the years. Special thanks to Annick Tanguay (my best friend) who is always here, giving me excellent advice and supporting me emotionally, particularly when I was emotionally vulnerable. To AnGie Yong, who always listen to my rants and a great companion for food in

Ottawa. To Wendy Lee, a wonderful and amazing friend who I always enjoy talking to and who selflessly offers great tips and edits of my writing. To Myriam Beaudry, a wonderful friend and my Bayesian Statistics partner in Stats Club. To Robyn Carson, a great friend who I always have fun to talk to and a wonderful partner who I enjoyed working with in the ICP peer review session. To Nicole Atkinson, an amazing leader who started up Stats Club at uOttawa and a great friend who is selflessly providing me with a lot of guidance and materials for my job hunt. To Jeffry Quan, Kim Yurkowski and Andrey-Ann Deneault for their precious friendship. To Kylee Ack Baraly, Odilia Yim, and Gabriela Dinescu for being my pals in Ottawa.

I would like to express gratitude to a number of people in LDL (my volunteers and LDL RA) for recruitment, data entry and collection. Thanks to Sarah Fitzpatrick for her voice that makes great stimuli in this thesis. Thanks to my wonderful volunteers for their amazing job in coding the data in the thesis: Maxine Berthiaume, Anika Talukder, Emma-Lee Peller and Caroline Zamor. Thanks to my honour students for data collection and coding: Hector Alfonso Almendrades and Alia Chamberlain. Last but not the least, thanks to the lab coordinators and RA for recruitment of infant participants: Vanessa Hébert, Julia Zhu-Pawlowsky, Joana Mukunzi, Isabelle Cossette, Caitlin Atkinson, Thyra Driver.

In addition to people in LDL, I have been very fortunate to work with people in CCLL and ERP Ling lab (Special thanks to Cristina Atance and Laura Sabourin). For people in CCLL: Thanks to Julian Caza for helping me with technical difficulties and for being a fun friend. To Stephanie Strahm, Maiziel Serrao and Caitlin Atkinson for your friendship and great help in SAT data collection. To Bronwyn O'Brien, Lojain Hamwi, Leia Kopp and Joshua Rutt for being wonderful colleagues in the CCLL. For people in ERP Ling lab: Special thanks to Santa Vínerte for being a wonderful and caring friend, and a great person with strong and amazing organizing skills. To Joy

Li for being a great and fun friend. To Michele Burkholder who is a great friend to talk to and selflessly participated in my study with her kids.

I also want to express gratitude to my mentors and friends in my early research training period. Thanks to Drs. Anita Wong and Chia-huei Tseng for their guidance and supervision in Hong Kong. Special thanks to Dr. Anita Wong, a wonderful and friendly mentor, who led me to develop an interest in language development. I want to thank to Ms. Vicki Lam, my tutor at high school, who genuinely cares about me and has taught me so many things about Economics and Mathematics. Special thanks to Ms. Stella Yau and Ms. Vienna Yeung for your precious friendship since high school (almost 20 years now!) and Ms. Angela Mak for being a great friend in my undergraduate study.

Finally, I would like to thank all my participants (both adults and infants) for your time and interest in my studies. A very personal thanks to my mom and dad who always have faith in me, allowing me to do whatever I want and supporting me unconditionally. I can only succeed and accomplish all these with your love and support. My heart is full of gratitude for your selfless love. Words cannot express all my feelings here but I hope I have made you proud.

Declaration of Academic Achievement

Angeline Sin Mei Tsui is the primary author of Study 1 and Study 2. Under the supervision of Dr. Christopher Fennell, the author developed and designed the experiments, collected data, analyzed data and wrote the manuscripts. Dr. Lucy Erickson and Dr. Erik Thiessen were involved in the development phase of the statistical segmentation experiment in Study 1 and Study 2. They offered input and expertise in the design of the experiment and the experiment materials. Part of the data related to the statistical segmentation task in Study 1 was published in the *Proceedings of the 41st annual Boston University Conference on Language Development* (Full citation: Tsui, A. S. M., Erickson, L., Thiessen, E., & Fennell, C. T. (2017). Statistical learning from accented speech: A bilingual advantage. In M. LaMendola and J. Scott (Eds.), *Proceedings of the 41st annual Boston University Conference on Language Development*, 679-690. Somerville, MA: Cascadilla Press.)

Statistical Learning in a Bilingual Environment

Acquiring a new language is not an easy task. Most adult learners experience great difficulty when learning a new language (e.g., Johnson & Newport, 1989; 1991; Mayberry & Fischer, 1989; Mayberry & Lock, 2003; Newport, 1990; see Newport, Bavelier, & Neville, 2001 for a review). Yet, the speed and ease with which young infants acquire language during the first year is remarkable. By 6 months of age, young infants are already able to learn some nouns that are common in their everyday linguistic environment (Bergelson & Swingley, 2012, Tincoff & Jusczyk, 1999; Tincoff & Jusczyk, 2012). They also show comprehension of more abstract words (e.g., all gone, eat) at 10 months (Bergelson & Swingley, 2013). By the end of their first year, they demonstrate well-refined speech perception of their native language (Werker & Tee, 1984; Kuhl et al., 2006) and may even be able to produce their first word (Fenson et al., 1994). Thus, infants demonstrate impressive language learning abilities even in their first year.

Although the speed of acquiring language is incredible during infancy, the process of word learning remains complicated when considering all the processes involved. Even consider the simple example of when an infant hears a parent say, “Look at the car.” Minimally, the infant must first extract the word *car* from the rest of the sentence, which is all unfolding in a continuous stream of speech, hold the word *car* in mind, identify the object *car* among all the other objects in their immediate environment as the referent for the word, and finally associate the word *car* with its referent. Evidently, the entire process of word learning is complex and multifaceted. However, as mentioned above, infants are able to handle the challenges of word learning to some extent, as they are capable of learning words within the first year. So what fundamental abilities allow infants to succeed in word learning? In this thesis, I examine the first step of word learning process, that is, learners’ ability to segment words from continuous speech streams. In particular, I will focus

on the flexibility and durability of this base skill, as I will concentrate on how learners (both adults and infants) make use of statistical information to segment words from continuous speech in a bilingual environment.

The Role of Statistical Learning in Language Segmentation Problem

One of the biggest challenges that infant learners must overcome is to correctly identify words in their native language. Identifying words from continuous speech is difficult because word boundaries are not marked by consistent silent pauses (Cole & Jakimik, 1980) or invariant acoustic cues (Miller & Eimas, 1995). Some people argue that infants may not face the same challenge of finding words in fluent speech because adults tend to speak with infants in infant directed speech, a register that is characterized by a higher pitch, slower rate and longer pauses between words. This suggests that infants may learn words in isolation as they tend to hear speech that is slower and contains longer pauses between words. However, prior studies have suggested that infants do not usually hear words in isolation (Aslin, Woodward, LaMendola, & Bever, 1996; Brent & Siskind, 2001; van de Weijer, 1998). For example, Brent and Siskind (2001) found that only 9% of maternal utterances contained isolated words, suggesting that infants largely learn words from fluent speech. Thus, infant learners predominantly need to learn words by extracting them from continuous speech in naturalistic language environments. As an adult, we can rely on our a priori linguistic knowledge, such as our existing knowledge of how sounds normally combine together to represent words (i.e., word forms) in our native language when segmenting words from continuous speech (Norris & McQueen, 2008). However, young infants, who are just beginning to learn words, do not possess a large set of familiar word forms and therefore would need to rely on a learning mechanism that does not require a priori lexical knowledge to segment words from speech. One mechanism that infants can make use of is statistical learning.

Statistical learning is a powerful mechanism that allows for the detection of regularities in the structure of one's surrounding environment. This learning mechanism tracks stable patterns via sensory inputs and applies the learned patterns in subsequent learning tasks. Young infants were shown to be capable of detecting sound sequence regularities in fluent speech (Aslin, Saffran & Newport, 1998; Saffran, Aslin & Newport, 1996a), as well as tracking musical tones and shape sequences that have stable structures (Saffran, Johnson, Aslin, & Newport, 1999; Kirkham, Slemmer, & Johnson, 2002). This learning capacity is also widely observed in other populations, such as children and adults (Saffran, Newport, & Aslin, 1996b; Saffran, Newport, Aslin, Tunick & Barrueco, 1997) and even animals (Toro, Trobalon, & Sebastián-Gallés, 2005; Saffran et al., 2008). Statistical learning can exploit a wide range of different aspects of statistical structures in the environment, such as frequency, variability, probability distribution, co-occurrence probability, and correlations (see Lany & Saffran, 2013 for a review).

Ample research indicates that statistical learning mechanisms play a central role in early language acquisition (e.g. Saffran, 2003; Maye, Werker, & Gerken, 2002; Thiessen & Saffran, 2003). Let us consider one aspect of statistical learning: transitional probability. Transitional probability refers to the co-occurrence relationship between two events, say event X and event Y. It is calculated by dividing the co-occurrence frequency XY by the frequency X. For example, if X occurs a total of 100 times in a sample and is followed by Y for 80 of those events, then the transitional probability from X to Y is 0.8. Transitional probability is argued to be a more robust representation in comparison to the mere co-occurrence relationship between XY. This is because transitional probability not only takes into account the co-occurrence frequency of XY, but also controls for the raw overall frequency of event X (e.g. Aslin, Saffran, & Newport, 1998). In other words, even if the frequency of event X co-occurring with event Y is high, if the frequency of

event X alone is also very high, the transitional probability of event X with Y could actually be relatively low. As mentioned earlier, there are no consistent and invariant cues to word boundaries in continuous speech: known as the classic word segmentation problem. However, there are transitional probabilistic cues that denote word boundaries, which could help infants solve this word segmentation problem. In the following, I would illustrate how infants use transitional probability across syllables to segment words from fluent speech. Here, I used syllables, rather than phonemes, as the units for statistical word segmentation. This decision is based on studies reporting that 2- to 3-month-olds were able to represent syllables in detail (Jusczyk & Derrah, 1987) and kept syllable representations in memory for a short period of time (Jusczyk, Kennedy & Jusczyk, 1995). Together, these studies have suggested that syllables may be salient for infants' lexical representations during the process of segmenting words from their language.

Let us consider the English phrase “pretty baby” as an example of statistical word segmentation. The syllable “pre” strongly predicts the syllable “tty” because “pre” in English is followed by relatively few syllables (e.g., pre-pare, pre-tty, pre-dict). Therefore, there is a high transitional probability that “pre” will be followed by “tty”. However, the final syllable of “pretty” does not strongly predict the first syllable of “baby” since it can precede many other syllables (e.g., pretty dog, pretty cat, etc.). Further, these syllables do not constitute an English word (i.e., teebay). Thus, there is a low transitional probability that “tty” will precede “ba”. Indeed, Saffran (2003) reported that the transitional probability between “pre” and “tty” is approximately 80% in speech to English infants whereas the transitional probability between “tty” and “ba” is approximately 0.03%. Based on these statistical structures, young infants can infer that “pretty” is likely to be a word and the transition from “ty” to “ba” probably denotes a word boundary. The differences

between these statistical structures provide clues to young infants that “pretty” is likely to be a word and the transition from “ty” to “ba” probably denotes a word boundary.

One of the most influential findings in the statistical learning literature nicely illustrates the above segmentation strategy. Saffran, et al. (1996a) demonstrated that 8-month-old infants readily make use of the transitional probabilities between syllables to segment a novel artificial language. The researchers exposed infants to a continuous speech stream of synthesized artificial speech for 2 minutes. This synthesized artificial speech was specifically created such that there were no other acoustic cues to word identity except the transitional probability (TP) between syllables. The TP was controlled in a way that the TP for syllables within trisyllabic sequences (i.e. words) was always 1.0 whereas the TP for syllables between words (i.e. part words) was 0.33. At test, infants listened longer to trisyllabic sequences with low TP than trisyllabic sequences with high TP. That is, infants found part words (TP = 0.33) and non-words (syllables never co-occurred in the language; TP = 0) novel, but found the words (TP = 1.0) familiar. This finding suggested that infants were able to extract words with high TP from the continuous speech merely on the basis of the statistical structure in the language, demonstrating that statistical learning can support early infant word segmentation.

A number of studies have further explored how TP plays a role in language acquisition across typically developing and atypical populations (e.g., children with specific language impairment) (Evans, Saffran, & Robe-Torres, 2009; Saffran et al., 1997), how TP interacts with different linguistic cues such as lexical stress and phonotactic knowledge (Johnson & Jusczyk, 2001; Saffran & Thiessen, 2003; Thiessen & Saffran, 2007), and how the output of TP processes in word segmentation tasks aids subsequent learning of segmented words (Erickson, Thiessen, & Graf Estes, 2014; Graf Estes, Evans, Alibali & Saffran, 2007; Hay, Peluncchi, Graf Estes, &

Saffran, 2011). However, the majority of laboratory studies of speech segmentation have focused on statistical learning with a single input stream. What about learners who need to learn multiple input streams, such as infants learning two languages? Will statistical learning mechanisms be sufficient to tackle the word segmentation problem when learners receive multiple inputs?

Word Segmentation with Two Input Streams

The current evidence for statistical learning in infant language acquisition is mainly derived from experiments that present a single novel artificial language as input (i.e., a simple “language” consisting of very few words created for the experiment). There are very few studies examining how multiple novel language inputs might affect humans’ learning. Bilinguals, or multilinguals, need to extract separate statistical structures from different languages in their learning environment, or else they may segment languages incorrectly and fail to learn the target units (i.e., words). To clearly illustrate this concept, let us consider how a French-English bilingual infant faces the word segmentation problem in her language environment. When the infant hears two consecutive syllables “mo-to” in continuous speech, there are at least two different meanings: it could mean an actual word in French (i.e. *moto*: motorcycle in English) or word boundary between two words in an English sentence (e.g. *Elmo told Ernie about his day.*). To determine if this syllable connection is a word or not, the French-English infant needs to keep track of the statistical probability between these syllables separately across the two languages. That is, “mo-to” is a word in French: the transitional probability between these two syllables is high. Alternatively, “mo-to” is not a word in English and therefore denotes a word boundary in that language: the transitional probability is low. Hence, bilinguals or multilinguals need to be able to create separate structural representations for different languages in order to effectively learn multiple languages simultaneously.

To examine whether humans can extract and represent two different languages in continuous speech, Weiss and his colleagues (2009) designed a simulated bilingual environment by exposing monolingual adult participants to two interleaved artificial languages with different statistical cues. There were two different conditions in the study: congruent and incongruent. In the congruent condition, the probability cues across two languages were completely compatible with each other and the TP across word boundaries in the two languages remained the same even if learners combined the two languages. In other words, if participants could not separate two statistical structures of the two languages, it did not necessarily lead to a failure in segmenting the words because the two statistical structures of the languages could be combined with no adverse effects of word boundaries. This would be similar to a situation where one has to segment a set of words across English and French that are highly similar across the languages (e.g., “alarm” and “*alarme*”, “extra” and “*extra*”, and “sofa” and “*sofa*”). However, in the incongruent condition, the probability cues across the two languages were incompatible and in conflict with each other, such that the two languages were conveying two different statistical structures that would lead to unclear boundaries if combined. This would be similar to the example used earlier where the combination of syllables forms a word in one language (“*moto*” in French), but a word boundary in the other (e.g., “Elmo told Ernie.” in English) in the other. In order to segment words correctly, the learners had to separate the languages and keep track of two different statistical structures. Successful learners would have to perform independent calculations for the two different languages. As such, statistical learning in the incongruent condition was much more demanding and challenging than that in the congruent condition.

Across several experiments, the researchers found that participants had no difficulty in learning the languages in the congruent condition but failed in the incongruent condition when no

additional cues were given. Participants were only capable of learning two languages in the incongruent condition when the languages were presented in two voices differing in their gender (Weiss, Gerfen, & Mitchel, 2009). In another follow-up study, the same group of researchers examined whether visual contextual cues can facilitate dual-input statistical segmentation. Mitchel and Weiss (2010) presented participants with two incongruent languages that were identical to Weiss et al. (2009) study, but the two languages were not marked by differences in speakers' voices. Instead, participants watched accompanying visual displays while listening to the artificial languages. In a series of experiments, Mitchel and Weiss (2010) found that participants successfully segmented two incongruent languages only when each language was individually paired with a video of a distinct talking face. If the languages were paired with two different static pictures of a talker or two different coloured screens (e.g., purple screen – language 1, teal screen – language 2), participants failed to segment the languages. Therefore, visual contextual cues can facilitate representations of two conflicting statistical language inputs, however, their effectiveness appears to depend on whether the visual displays are natural informative cues to talkers' identities (e.g., learners expect to see dynamic moving faces, but not static pictures or unrelated coloured screens, when listening to speech sounds).

Are contextual cues necessary to determine two languages are present in the input? Gebhart, Aslin and Newport (2009) proposed that learners may also detect changes in two statistical structures by monitoring the consistency and variability of the surface statistics from sensory inputs over time (see Basseville & Nikiforov, 1993; Qian, Jaeger, & Aslin, 2012 for similar arguments). When learners hear one single language, they essentially receive a stable set of sensory inputs in which the surface statistics are constant over time. Learners will therefore conclude that there are no changes in the underlying structure of the input. On the contrary, when learners hear

two different languages, they are no longer receiving stable sensory input and they will notice the surface statistics of the input vary in systematic ways over time. In this case, there are two possibilities. Learners may treat the input variability as noise and conclude that they are hearing just one language, which would have the consequence of failing to learn the two languages. Alternatively, learners may treat the input variability as a signal of changes in the underlying structures of the input, which would help them detect the two different statistical sets (i.e., languages) in the environment.

In an initial attempt to address this latter possibility, Gebhart, Aslin, and Newport (2009) presented participants with two incongruent language structures in a sequential successive order, as opposed to the interleaved order in the studies outlined earlier. Participants listened to one language for 5 minutes and then listened to the other language for another 5 minutes. When there were no contextual cues (e.g., pauses, voices changes) that indicated structural changes across languages, participants showed a primacy learning effect: they only learned the first heard language but not the second. This primacy learning effect was again observed when the researchers added a subtle contextual cue: the two languages were marked by a change in overall pitch. However, if participants received salient contextual cues, such as when pauses were inserted between the presentation of the two languages and when participants were explicitly informed that they were listening to two languages, participants could then successfully segment both languages. Finally, Gebhart and colleagues (2009) found that, without contextual cues, participants could only learn the second-heard language when it was triple in length relative to the first. In sum, these experiments demonstrate that, while it is possible for learners to separate two different language structures without contextual cues, the resultant learning is weak in that only one language is acquired - unless a large over-exposure to the second language is provided.

Weiss and his colleagues further investigated why learners show a primacy learning effect when learning two different artificial languages in the absence of any contextual cues. Zinszer and Weiss (2013) found that participants successfully segmented two languages if the interleaved languages were presented in smaller blocks. This indicates that learners are more likely to treat statistical variability in sensory input as changes in the underlying structures if they receive more variable input (i.e., presenting languages in smaller blocks increased the number of switches between language 1 and 2). Moreover, subsequent studies have suggested that a shorter duration of language exposure can help learners avoid over-learning the first language. Under this account, learners are reluctant to change and learn the second language once they have established a solid representation of the first language. Bulgarelli and Weiss (2016) found evidence that longer exposure to the first presented language limits participants' learning of the second presented language. Bulgarelli and Weiss (2016) presented participants with the artificial languages in smaller language blocks. Crucially, the researchers included a short test after each language block to test participants' accuracy of the language. When participants showed mastery of Language 1 (e.g., they scored at least 7 out of 8 in Language 1), they immediately proceeded to Language 2. This design substantially reduced participants' exposure to Language 1 because participants would no longer hear Language 1 after they had learned it. Using this design, Bulgarelli and Weiss (2016) discovered that participants successfully learned two languages. This supports the argument that longer exposure to the first heard language may lead learners to establish a single stable linguistic structural representation and form a low expectation of future changes. Consequently, learners reduce their attention to input variability and treat the statistical variability of the second language input as noise.

Along the same line, Karuza et al (2016) used fMRI technique to examine the same question and found that participants who showed stronger learning of the first heard language (the primacy effect) have reduced activation in posterior parietal cortex (PPC), a brain area associated with monitoring and detecting changes in a dynamic shifting environment. Conversely, participants who showed stronger learning of the second heard language exhibited an opposite effect (i.e., increased activation in PPC). This study provides evidence that learners with a stronger representation of the first heard language are less likely to detect changes in the underlying statistical structures, and therefore fail to acquire the second heard language. In general, the above-mentioned studies have suggested the primacy effect observed in the dual-input segmentation studies may be related to overlearning of the first presented language. When learners are exposed to the first presented language for a longer period of time, they may form a stable representation of this language. This solid representation of the first presented language can lead learners to form an expectation that there is only one single language input, reducing their attention to new changes in their environments. This, in turn, may limit learners' ability to detect a second language. By contrast, if learners did not form a solid representation of the first presented language, they may be more flexible to accommodate the new changes in their environments and are more likely to detect a second language. This argument is supported by the studies showing shorter language exposure of the first presented language is related to better learning of the second presented language. Thus, it is possible that learners can segment two overlapping languages in the absence of contextual cues if they effectively monitor the consistency of incoming language input. However, their attention to the consistency of the language input can be affected by the length of first presented language. This is because learners may infer that there is no changes in the language input when they are exposed to the first presented language for an extensive period of time.

To summarize, there are two sources of information that facilitate segmentation of two conflicting artificial language inputs: (i) the availability of clear contextual cues and (ii) learners' monitoring skills of the language input over time. However, the current literature appears to suggest that adults learned best in the first case: when the two languages were marked by salient contextual cues. Nevertheless, these studies have only explored how speaker-specific contextual cues facilitate adult word segmentation from the two language inputs (e.g. two voices differing in gender or two different talker faces; see Weiss, Poepsel & Gerfen, 2015, for a review). This leads to a question of whether other contextual cues, especially those that might be present in naturally occurring bilingual environments, can also facilitate learners' segmentation of two languages within the same paradigm. In naturalistic language learning environments, two distinct languages are often not spoken by two different speakers. The very nature of being bilingual means that one person speaks two languages. A bilingual parent would speak one language in some contexts and the other language in other contexts. Further, bilinguals also often alternate between two languages even in a single conversation, a phenomenon that is known as code-switching (e.g., Poplack, 1980; see Heredia & Altarriba, 2001, for a review). It is thus important to explore whether contextual cues in a single individual's speech can facilitate adult segmentation of two artificial languages.

To fill in this research gap, I will investigate whether a single speaker's contextual cues can help learners to segment two different artificial language inputs. Here, I investigate whether language-specific phonetic cues within a single speaker's speech, which will be termed as accented speech in this thesis, is an effective contextual cue for successful statistical segmentation of two language inputs. Phonetic realizations of syllables predominantly vary across two languages. Consider our previous example about the word *moto*, this word has the same phonemes across French and English, but the phonetic properties differ across the languages. *Moto* would be

produced with tenser vowels and less duration in French than in English. Hence, these phonetic variations can be distinguishable indexical cues that signal a change of languages within a single individual's speech. As such, phonetic realizations can be ecologically valid contextual cues for learners to track two language inputs in a bilingual environment.

Another major issue in the statistical learning literature is that most dual-input statistical segmentation studies exclusively tested monolingual populations. Thus, how bilinguals perform with two artificial languages is not well known, despite the fact that they would have experience doing exactly that in their natural learning environment. Indeed, some recent studies have turned to bilingual populations and their results suggest that bilinguals may be advantaged in dual-input statistical learning tasks (e.g., Antovich & Graf Estes, 2018; Bartolotti, Marian, Schroeder, & Shook, 2011; Poepsel & Weiss, 2016). In the following, I will provide a review on the current literature that examine whether there is a bilingual advantage in a dual-input statistical learning task.

Does a Bilingual advantage in Dual-input Statistical Learning Tasks Exist?

There are several studies suggesting that bilinguals may outperform monolinguals in a dual-input segmentation task. Bartolotti, Marian, Schroeder, and Shook (2011) examined whether adult learners can segment two artificial Morse code languages under low and high interference conditions. Word boundaries were signaled by both statistical cues and pauses between words in the low interference condition. However, the segmentation task became harder under the high interference condition because the statistics indicating word boundaries did not align with pauses between words. Bartolotti et al. (2011) found that bilingual experience was correlated to better segmentation performance in the low interference condition while inhibitory control was correlated to better segmentation performance in the high interference condition.

More recently, Antovich and Graf Estes (2018) replicated the Weiss et al (2009) studies and investigated whether monolingual and bilingual infants at 14 months can statistically segment two interleaved languages in which the languages were marked by two gendered voices. Their results showed that bilingual infants successfully segmented two languages yet monolingual infants failed. Further, in an unpublished study, Haesung (2012) also examined whether bilingual and monolingual adults can learn two incongruent interleaved languages when the languages were presented with a foreign accent. She found that both monolingual and bilingual adults were able to track two distinct statistical sets when the speech was presented in a native American-English accent. However, only bilingual adults were capable of tracking two statistical sets when the languages were presented in a non-native Korean accent. However, there is one aspect of this unpublished study that may confound the findings about a bilingual advantage in statistical learning. Bilinguals in the study were only identified by asking whether participants can speak more than one language and whether they consider themselves as bilinguals. This criterion may not accurately differentiate bilinguals from monolinguals, since it is not based on self-rated and/or objective proficiency tests. Thus, the results about bilingual advantage in Haesung (2012) need to be interpreted with caution.

Apart from dual-input statistical segmentation tasks, there are emerging studies reporting that bilinguals generally outperform monolinguals in other statistical learning tasks. Wang and Saffran (2014) tested monolinguals and bilinguals on their ability to segment words by tracking statistics in a single artificial tonal language. The transitional probabilities across word boundaries were signaled by syllables and tones, with no conflicts between the two cues. The authors found that bilinguals, even those that had no prior exposure to a tonal language, succeeded in segmenting the words. Monolinguals, on the other hand, failed to do so, even if their native language was a

tonal language (but see Caldwell-Harris, Lancaster, Ladd, Dediu & Christiansen, 2015 for evidence that tonal language experience aids performance in a similar task). In another study, Poepel and Weiss (2016) compared monolinguals with bilinguals in a cross-situational statistical learning task. In this task, learners see multiple objects (e.g., two to four objects) displayed on computer screen and hear the corresponding labels across trials. There is no explicit information about which label goes with which object. However, each label always co-occurs with only one target object across trials and successful learners aggregate the co-occurrence data across trials to infer the most frequent and reliable label-object pairings (see Smith, Suanda & Yu, 2014 for a review). Poepel and Weiss found that bilinguals outperformed monolinguals, both in terms of overall proficiency and speed, in a more difficult task when they needed to acquire two-to-one mappings but not in an easier one-to-one mapping.

On the other hand, several studies have found no difference between bilinguals and monolinguals in statistical learning. For example, monolingual and bilingual children of ages 5 to 13 performed similarly in statistical learning tasks involving visual shapes and pure-audio tones (Yim & Rudoy, 2013). Bogulski (2013) also found that English-Spanish bilingual adults exhibited a primacy learning effect when they were tested with a dual-input segmentation task without any contextual cues, similar to the established findings involving monolingual-English adults. Lastly, Grey (2013) reported no group differences between Mandarin-English bilinguals and monolingual English adults in an implicit artificial grammar task (but see Nation & McLaughlin, 1986 for multilingual advantage in another artificial grammar task). As such, evidence regarding a bilingual advantage in statistical learning has been mixed.

The observed mixed findings can be due, at least partially, to variations of experimental design in statistical learning tasks and/or differing statistical power based on varying sample sizes.

However, these inconsistent findings could also imply that the core underlying statistical learning mechanism may not be affected by bilingual experience. Several researchers argued that statistical learning is a domain-general mechanism (e.g., Saffran, 2003; Saffran & Thiessen, 2007) as it can operate across multiple domains, such as language, visual geometric shapes, and musical tones. If bilingual experience improves the fundamental statistical learning mechanism, we should observe bilingual advantage in statistical learning tasks across different domains, such as audio tones, visual shapes and language segmentation. However, the current literature reveals that bilinguals seem to outperform monolinguals only in some linguistic statistical learning tasks, and not generally in statistical learning tasks across different domains. Importantly, for the linguistic learning tasks that bilinguals appear to have better performance, contextual cues that exist in bilingual environment are present in these tasks (e.g., two labels (one in each language) are paired with one object in a cross-situational statistical learning task). Perhaps bilinguals are advantaged in linguistic statistical learning tasks as they can process and readily utilize contextual cues in bilingual environments faster.

If bilinguals are indeed advantaged in some linguistic statistical learning tasks, what are the factors that drive the observed advantage? There are at least two possible factors. First, bilinguals may have certain enhanced cognitive abilities. This argument stems from the converging evidence that bilinguals possess better executive functioning (Bialystok, 1999; Bialystok & Martin, 2004; Kovács & Mehler, 2009a; see Bialystok, Craik, & Luk, 2012). Consistent with this argument, some studies have suggested a link between better cognitive/attentional control and performance in statistical word segmentation tasks (e.g. Toro, Sinnett, & Soto-Faraco, 2005; Weiss, Gerken, & Mitchel, 2010). Thus, it is possible that bilinguals have an enhanced executive function system that supports statistical learning.

An alternative, but not mutually exclusive, account is that bilingual experience may lead learners to be perceptually open to changes in the underlying structures when processing language input in novel environments. Bilinguals are often exposed to a much more variable linguistic environment, for example, learning two languages that are substantially different in their phonological inventory (e.g., English (i.e., non-tonal language) and Cantonese (i.e., tonal language)) or communicating with multiple speakers who speak their languages with non-native accents. Experience of handling variability in bilingual environments may have prepared bilinguals to form an expectation that language inputs may change, even under novel circumstances. Thus, they are more flexible than monolinguals when tracking language changes in statistical learning tasks. Here, I argue that bilinguals and monolinguals have the exact same fundamental statistical learning mechanism. Yet, bilinguals may be more flexible when interpreting variability of language input in new language environments, which subsequently helps them to segment words accurately from two overlapping languages in a dual-input statistical learning task. This argument suggests that bilingual experience provides learners more opportunities to deal with variability in language environments and strengthens their expectation that languages can change in new environments.

In the following sections, I will provide a review of how bilinguals' better executive functioning and/or more flexible perceptual processing skills can facilitate statistical segmentation in dual-input tasks.

Bilinguals' advantage: Better executive functioning. The root of bilinguals' cognitive advantage is hypothesized to stem from their joint activation of both languages during lexical processing (Gollan, Montoya, Fenneman-Notestine, & Morris, 2005; Thierry & Wu, 2007; Yan & Nicoladis, 2009). In order to successfully and efficiently communicate with other people, bilinguals need to selectively attend to and use the correct language while inhibiting the other competing language. This routine need for selectively attending to the relevant language and suppressing the irrelevant language can have a positive influence on bilinguals' executive function system.

The executive control system consists of three core components: inhibition of dominant responses, shifting between tasks or mental sets, and updating/monitoring working memory (Miyake et al., 2000)¹. Early studies on bilinguals' cognitive advantages mainly focused on inhibition, stimulated by Green's (1998) inhibitory control model. This model emphasizes the role of the supervisory attentional system in inhibiting the irrelevant competing language. To test bilinguals' inhibitory control advantages, researchers employed a number of non-verbal conflict tasks, including the flanker task (e.g. Costa, Hernández, Sebastián-Gallés, 2008; Costa, Hernández, Costa-Faidella, Sebastián-Gallés, 2009), the Simon task (e.g. Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Martin, & Viswanathan, 2005), and the Spatial-Stroop task (Bialystok, 2006; Bialystok, Craik, & Luk, 2008a). These studies found that bilingual children (both pre-school and school-aged), young adults, and older adults outperformed monolinguals in the conflict tasks. Bilinguals' enhanced performance in these conflict tasks was more pronounced

¹ In this thesis, I focus on the executive function model of Miyake et al (2000), as it is supported by large-scale latent variable analyses. There are some modifications of the model in Miyake and Friedman (2012), that is, inhibition was dropped and two components remained: shifting and updating. However, the current literature of bilingual executive function advantage mostly relies on Miyake's (2000) model. As such, I follow the traditional approach and define the executive functioning system in terms of three components.

in older adults (e.g. Bialystok et al., 2004; Bialystok, Craik, & Luk, 2008b), and when the cognitive demands in conflict tasks were high (e.g. Costa et al., 2009; Bialystok, 2006). Bilinguals also outperformed monolinguals in cognitive tasks which require learners to suppress interference from a competing cue in relation to the presented stimuli (Martin-Rhee & Bialystok, 2008; Luk, Anderson, Craik, Grady, & Bialystok, 2010).

Subsequent studies have extended research to other aspects of executive control components aside from inhibition. A major reason for this move was that researchers (e.g., Bialystok, 2006; Costa et al., 2008; 2009) not only found bilingual advantage in the incongruent trials of the above conflict tasks, which require inhibitory control, but also in congruent trials that do not require inhibition. Here, the difference between incongruent and congruent trials is the amount of attention effort required in the task. For example, in a flanker task, participants are presented with a row of five arrows and are asked to press a key to indicate the direction in which the central arrow is pointing. The attention effort is normally lower for “congruent trials” when the four irrelevant side arrows are pointing at the same direction as the central arrow. However, participants normally need to put forth more attentional effort when the four irrelevant side arrows are pointing to the opposite direction in comparison to the central arrow. These “incongruent trials” require participants to make additional attentional effort to ignore the side arrows in order to correctly respond to the direction of the central arrow.

If bilinguals have a cognitive advantage that is not limited to inhibition, perhaps they may have an overall domain-general executive functioning advantage (Bialystok, 2015; see Hilchey & Klein, 2011 and Coderre & van Heuven, 2014 for reviews). Costa et al (2008) proposed that bilinguals have a more efficient environmental monitoring system because they need to monitor and switch attention between two languages regularly. Evidence from bilinguals’ better

performance in shifting (Prior & MacWhinney, 2010; Prior & Gollan, 2011) and working memory (Morales, Calvo, & Bialystok, 2013; Luo, Craik, Moreno, & Bialystok, 2013) tasks support this perspective. Recently, infant researchers also found bilingual advantage in several cognitive processing tasks that do not only involve inhibition, such as learning and generalizing two structural regularities (Kovács & Mehler, 2009a; Kovács & Mehler, 2009b), memory generalization (Brito & Barr, 2012; Brito, Sebastián-Gallés, & Barr, 2015) and habituation paradigm (Singh et al., 2015). These studies together support that bilingual experience improves the executive functioning system.

However, there is a recent debate of whether bilingual executive functioning advantage is robust or not, since there are significant amount of findings that do not find a bilingual advantage in executive functioning (e.g., Antón et al., 2014; Duñabeitia et al., 2014; Gathercole et al., 2014; Kousaie & Philips, 2012a,b; Kousaie et al., 2014; Paap & Greenberg, 2013; Paap & Sawi, 2014; see Hilchey, Saint-Aubin, Klein, 2015 for a review). Some researchers have also found evidence of a publication bias in the literature on the bilingual cognitive advantage (de Bruin, Treccani & Della Sala, 2015). Adjudicating between the position of a robust bilingual cognitive advantage and that of a small (or non-existent) bilingual cognitive advantage is not the central focus of the current thesis. The main focus of the thesis is to investigate whether bilinguals' possibly advanced executive functioning system can explain why bilinguals may outperform monolinguals in a dual-input statistical segmentation task.

If bilinguals enjoy a better executive functioning system, one may ask how this cognitive advantage is linked to possible better performance in segmenting multiple statistical streams from continuous speech. To understand the link, we can turn to a memory-based theoretical account of statistical learning: the Extraction and Integration Framework (Thiessen, Kronstein, & Hufnagle,

2013, see Thiessen, 2017 for a review). This framework proposes two main types of regularities to which learners are sensitive in statistical learning: conditional statistics and distributional statistics. Conditional statistics reflect the likelihood of two or more elements co-occur in the sensory input and thus it describes a predictive relationship between event X and event Y. On the other hand, distributional statistics refer to a central tendency or prototypical characteristics of a set of elements. This type of statistics reflects the distribution of a set of exemplars in the sensory inputs, such as frequency (e.g., mean, median, and mode) and variability (e.g., standard deviation, variances). Sensitivity to these two distinct types of statistics is driven from two different processes: (i) Extraction and (ii) Integration.

Extraction refers to a process of identifying units (e.g. word forms) from the input (e.g. continuous speech) and storing these units in memory. The extraction process explains the underlying process of conditional statistics, such as the use of transitional probability in segmenting words from continuous speech. The Extraction and Integration framework focuses on a chunking model, PARSER (Perruchet & Vinter, 1998), to explain how the extraction process can give rise to the detection of conditional statistics. To illustrate, let us reconsider the classic word segmentation situation. The PARSER model proposes that when learners are presented with a continuous speech stream, syllables are randomly grouped into different chunks and stored in memory. The activation of these chunks decays over time, unless these chunks are re-experienced. Thus, chunks undergo different degrees of activation based on the frequencies and co-occurrence of syllables in the input. Chunks that are more frequent and statistically coherent (i.e. syllables that co-occur within word) will more likely be reactivated than those that are less frequent and less statistically coherent (i.e. syllables that co-occur across word boundaries). The high frequency of reactivated chunks makes its representation in the memory stronger and therefore new information

and noise in the data are less likely to interfere with these chunks. Finally, chunks that receive the most activation will be represented as words in memory and this output reflects the statistical structure of the inputs (i.e., syllabic transitional probability within word).

As mentioned previously, another common structure in statistical learning is forming a distribution across exemplars of the stimuli. To form a distribution, learners not only need to extract statistical structures from the input (i.e., the extraction process), but also integrate information across individual exemplars and abstract the central prototypical information. Integration is argued to be the process that gives rise to distributional statistics. The Extraction and Integration framework has relied on an exemplar memory model called “Integrative MINERVA”, or iMINERVA for short (Thiessen & Pavlik, 2013), to describe the integration process. In iMINERVA, a new exemplar is compared with the prior exemplars in the memory system and a weighted average will be generated after integrating information across current and prior exemplars. This weighted average is a function of similarity between current and prior exemplars. When current and prior exemplars are similar, similar exemplars will be highly activated and contribute strongly to the weighted average. A representation of an exemplar will be stored in the memory system but subject to decay over time if it is no longer similar to other exemplars. Finally, the model will integrate all information across exemplars and generate a weighted average that reflects the central tendency (e.g., mean, variance) of the distribution of a set of exemplars.

To illustrate this exemplar process, let us consider another classical example of how infants form phonemic categories. Maye, Werker and Gerken (2002) have demonstrated how infants categorize exemplars along a continuum of voice-onset time (VOT) between the two phonemic sounds /d/ and /t/ based on the distribution of sounds they heard. When infants were exposed to a unimodal distribution (i.e., sounds in the middle of the VOT continuum occurred most frequently),

infants treated all sounds they heard as one phonemic category. However, when infants were exposed to a bimodal distribution (i.e., sounds near the two endpoints of the continuum occurred most frequently), infants showed discrimination of sounds near the two endpoints (i.e. one near the phoneme /d/ and the other near the phoneme /t/), which indicated that infants identified two different phonemic categories. Applying iMINERVA model to this example, when infants are exposed to the unimodal distribution, the memory system encodes the exemplars in the middle of the VOT distribution and these exemplars are very acoustically similar to each other, thus the weighted average will produce a single distribution. However, in the case of bimodal distribution, the exemplars are most likely located near the two ends of the VOT distribution and the exemplars at different ends are very dissimilar, this signals two different modes and eventually produces two prototypes in the model.

In sum, the Extraction and Integration framework suggests two distinct processes, extraction and integration, to explain learning of conditional and distributional statistics. Both extraction and integration involve some basic memory processes: activation, decay and interference. However, they are distinct and separable in terms of their outputs: extraction process gives rise to conditional statistics (e.g., transitional probability) whereas integration process gives rise to distributional statistics (e.g., mean and variance). Thus, it is argued that extraction and integration processes rely on different memory systems. For example, extraction mostly relies on attention and working memory, because learners have to pay attention to grouping elements (e.g., syllables) into chunks and later hold the chunks (e.g., syllables chunks) in working memory for further processing such as activation and decay. In contrast, integration mostly requires long-term memory when learners need to compare new and prior exemplars to abstract prototypical information. In line with the above argument, some neurological findings have demonstrated that

extraction and integration processes produce different event-related potential (ERP) responses (e.g., Balaguer, Toro, Rodriguez-Fornells & Bachoud-Levi, 2007; Mueller, Girgsdies & Friederici, 2008).

In the current thesis, I will only focus on the extraction process because I am discussing the fundamental process of speech segmentation – the extraction of words from two interleaved continuous speech streams. First of all, I propose that attention and working memory play crucial roles in the process of segmenting words in continuous speech streams. This is based on the assumption that only syllables held in attention can be extracted and divided into chunks, and working memory is useful for storing the chunks for further processing (e.g., compare and reactivate previously stored chunks). This hypothesis is consistent with prior studies showing that attention is required for successful statistical learning (e.g., Baker, Olson, & Behrmann, 2004; Toro, Sinnett, & Soto-Faraco, 2005). Specifically in the task of learning two simultaneous language inputs, I argue that inhibitory control and set-switching are also needed for separating two inputs and shifting attention between the two languages. This is because learners need to suppress one statistical stream in order to attend to the other statistical stream, and simultaneously they also need to switch attention between two different languages (i.e., two languages were presented in an interleaved order). If bilinguals have enhanced executive functioning system that includes attention, working memory, inhibitory control and set-switching, they may perform better than monolinguals in a word segmentation task with two input streams. As such, I will examine whether the above-mentioned aspects of executive functioning will relate to participants' performance in the dual-input segmentation task.

Bilinguals' advantage: Flexible interpretation of input. Another hypothesis put forward to explain bilinguals' possible performance advantage in dual-input statistical tasks is that they are more flexible and open than monolinguals in how they interpret language inputs. This hypothesis is closely tied to bilinguals' regular exposure to a more variable language environment. In the following, I will first discuss why bilingual environment is more variable and then explain how variability in the language environment may lead to bilingual advantage in interpreting new linguistic inputs.

To understand the linguistic variability in a bilingual environment, let us first think about phonological² differences (i.e., sound contrasts that signal difference in word meaning) between two languages. Bilinguals have to learn two different phonological systems but the amount of time they hear each language is less than monolinguals (i.e., parents do not double the amount they speak to bilingual infants compared to monolinguals; Byers-Heinlein & Fennell, 2014; Costa & Sebastián-Gallés, 2014). A fundamental challenge of learning two phonological systems is that there could be substantial phonetic overlap between the two languages. For example, the phonetic contrasts /p/ and /b/ both exist in French and English, and speakers of both languages perceive the [b] phone as /b/ and the [p^h] phone as /p/. However, along the voice-onset-time (VOT) continuum between [b] and [p^h], French and English speakers place their phonological boundaries differently. VOT is the time difference between the release of a stop consonant and the beginning of the vibration of the vocal cords. This leads to different interpretations of phones lying in the middle of that particular VOT continuum across in English and French. For example, the [p] phone is

² There is clear distinction between phones (or phonetics) and phonology (or phonemic representations) in the literature. Phones refer to the basic acoustic units of the language sounds without carrying lexical meaning of the language. For example, pi:tər and p^hi:tər are phonetically different but these two words are not meaningfully different (i.e., both means the word “Peter” in English). On the other hand, phonology refers to the language sounds that indicates lexical meanings. For example, /b/ and /p/ are phonologically contrastive in English, therefore *bit* and *pit* are two lexically different words.

perceived as /b/ by English monolinguals and as /p/ by French monolinguals. French-English bilinguals thus need to remain sensitive to all three phones (i.e., [b], [p], and [p^h]), so that they can flexibly switch the interpretation of phones in different language contexts. For instance, a French-English bilingual needs to know that a change from the phone [b] to [p] reflects lexical differences in French but not English. Indeed, research has suggested that bilingual French-English infants as young as 10 months are able to discriminate all three phones whereas their monolingual English counterparts only show discrimination of their native categories: treating [p] as [b] (Burns, Yoshida, Hill & Werker, 2007, see also Sundara, Polka & Molnar, 2008 for similar findings of dental consonant /d/- /ð/). In a similar vein, using magnetoencephalography (MEG) techniques, Ferjan Ramirez, Ramirez, Clarke, Taulu, and Kuhl (2017) found that bilingual Spanish-English infants are sensitive to sound contrasts present in both of their native languages. Together, all these studies lend support to the hypothesis that, to a greater extent than monolinguals, bilinguals have to maintain flexibility for phoneme boundaries and acquire a broader phoneme inventory for successful phonological categorization.

Similar to the development of native phoneme perception, bilingual experience may affect the development of non-native phoneme discrimination. Generally speaking, during the first year of life, monolingual infants' phonetic perception undergoes a process of *perceptual narrowing*. This process attunes infants' perception from being sensitive to both native and non-native phonemes to being predominantly sensitive only to native language sounds (e.g., Narayan, Werker, & Beddor, 2010; Kuhl et al., 2006; Werker & Tees, 1984). Infants' loss of sensitivity to non-native sounds is thought to be driven by the regular exposure to native sounds in their language environment (Maye, Werker & Gerken, 2002). Recall that there is more phonetic variability in bilingual language environments. If bilingual infants are exposed to a more variant phonetic

environment and are less exposed to each of their native languages, it is possible that bilinguals may experience a protracted period of non-native sound sensitivity. There are fewer studies that have investigated bilingual infant non-native phonetic perception, but there is some evidence supporting the openness to non-native phoneme discrimination in bilingual population. Using functional Near Infrared Spectroscopy (fNIRS) technique, Petitto et al (2012) have shown that the left inferior frontal cortex (a brain region related to language processing) in both monolingual and bilingual infants decreases in activation for non-native Hindi sound contrasts as their age increases. However, bilingual infants' brain activation has a smaller decrease in activation when they grow older, suggesting that bilingual infants may have a prolonged period of sensitivity to non-native sound contrasts.

This extended period of non-native sensitivity has also been demonstrated in studies examining how humans differentiate between two different languages in a visual domain. Speech is multimodal in nature. Prior studies have shown that adults and infants can discriminate one language (e.g., Spanish) from another language (e.g., Catalan) by watching silent talking faces (e.g., Soto-Faraco et al., 2007; Weikum, et al., 2007). Sebastián-Gallés, Albareda-Castellot, Weikum and Werker (2012) tested whether 8-month-old infants learning Spanish and/or Catalan could visually differentiate between two non-native languages: English and French. Their results showed that bilingual Spanish-Catalan infants succeeded whereas monolingual Spanish or Catalan infants failed, further supporting the idea that bilingual experience may extend the period of non-native sound discrimination. If bilinguals have a protracted period of non-native phoneme sensitivity, perhaps a natural follow-up question is to ask whether sensitivity can be extended to a possible bilingual advantage in learning non-native words.

To date, there have been two studies (Graf Estes & Hay, 2015; Singh, 2017) suggesting that bilingual infants' protracted non-native phoneme sensitivity promotes flexibility of novel word learning. Graf Estes and Hay (2015) examined whether bilingual 14-, 19- and 22-month-old infants learning non-tonal languages will associate non-native lexical tone (i.e., tone variations that signal a change of word meanings) with objects in a Switch task. They found that bilinguals at 14 and 19 months, but not 22 months, are willing to associate lexical tone with objects. This finding is in conflict with the developmental trajectory of non-tonal monolinguals who stop associating lexical tones to objects at 17 months of age (Hay, Graf Estes, Wang & Saffran, 2015). A subsequent study also found that 19- to 20-month-old bilingual English-Mandarin infants, but not monolingual-English infants, are willing to map non-native click sounds to objects (Singh, 2017). Further, Singh found that bilinguals' flexibility in mapping sounds to objects are only restricted to linguistic sounds since bilinguals are not willing to accept hand-clap or finger-snap sounds as labels for objects. Thus, bilinguals' extended non-native phoneme sensitivity may relax children's word learning bias and facilitate novel non-native word learning.

Finally, bilingual children are also more flexible in other word learning processes. Bilingual experience can influence one aspect of word learning heuristics – disambiguation (also termed mutual exclusivity). Monolingual children often employ the disambiguation heuristic to infer the name of a new object in an underspecified context. For example, when children see a pair of objects, one familiar object (e.g., cup) and a new object (e.g., a stapler), children tend to associate a novel word (e.g., *bin*) with the new object rather than to the familiar object (Markman & Wachtel, 1988; Merriman & Bowman, 1989). Children rely on their prior lexical knowledge to quickly infer the association between a new word and object. Given that the disambiguation heuristics is based on children's previous lexical knowledge that one label maps on one object,

perhaps it is not surprising that bilingual experience will change children's use of this heuristics in word learning. In a bilingual language environment, children often encounter situations where at least two labels (e.g., the same cup is labelled as *cup* in English and *tasse* in French) are associated with one object, which is in apparent contrast with the monolingual environment where one label consistently maps on one object. Thus, the number of label-to-object in bilingual environment is a variable and this variability arguably leads bilingual children to adhere less to the use of disambiguation heuristics in word learning. Several studies have shown that 18-month-old bilingual infants do not adhere to the disambiguation heuristics when associating a novel word with a novel object in an experiment (Byers-Heinlein & Werker, 2009; Houseton-Price, Caloghiris & Raviglione, 2010; Kandhadai, Hall & Werker, 2017). Thus, these findings suggested that bilinguals are actually more flexible in interpreting novel information during the word learning processes.

Thus far, I have reviewed studies that demonstrate bilinguals are more open and flexible when interpreting new linguistic input. For example, bilinguals are found to be more flexible when associating non-native sounds to objects and accepting multiple labels for the same referent in several above-mentioned experiments. In the current thesis, I argue that bilinguals' flexibility guide them to form an expectation that structural changes in new language environments can occur, helping them to detect changes in statistical inputs faster and thus performing better when segmenting words from two language inputs. Learners' expectations of input changes are the key to success in a dual-language segmentation task. If learners are not expecting any changes in their environments, they will tend to treat variability of surface statistics as noise and consequently fail to detect changes in the underlying statistical structures of the language input. According to Qian, Jaeger, and Aslin (2012), learners' prior knowledge and experience will influence how they form

expectations of structural changes under novel circumstances. A bilingual environment is arguably dynamic and highly variable in terms of language input (see Sebastián-Gallés, 2010 for a review). Bilinguals may therefore have developed an expectation that new environments could be subject to changes and this leads them to be more open and flexible in interpreting the inconsistency of surface statistics in a dual-input statistical segmentation task. Moreover, bilinguals may be more familiar to certain contextual cues that are present in bilingual environments, such as phonetic cues associated with two different languages (Hudon, 2013). Bilinguals may be faster in identifying these contextual cues as markers of two language inputs in a dual-input statistical segmentation task. In closing, the bilingual environment is pivotal to learners' flexibility when they interpret linguistic inputs in a novel environment. I argue that early and continuing bilingual experience will be a leading factor of being open and flexible because of the accumulated time spent in bilingual environments. Thus, I expect that bilingual experience positively relates to the word segmentation performance in a dual input task.

Overview of Current Thesis and Hypotheses

Statistical learning is a mechanism by which one acquires and detects regular structures (e.g., frequency, variability, co-occurrence probability) from sensory inputs in the ambient environment. This ability is thought to be an important component of a fundamental step of language acquisition – segmenting words from continuous speech (e.g., Perruchet & Vinter, 1998; Swingley, 2005; Saffran, Aslin, & Newport, 1996; see Saffran, 2003 and Thiessen, Kronstein, & Hufnagle, 2013 for reviews). Over the past decades, the research into statistical learning has been growing substantially (Perruchet & Pacton, 2006; Thiessen, 2017), yet, the current literature has mostly focused on a monolingual language context, both in terms of participants and modelling inputs in the task. Given the prevalence of bilingualism in the world (Grosjean, 2010), it is essential

to extend research to a bilingual language context. Recently, there has been an increasing trend in research towards understanding of how learners statistically segment two different language inputs in a simulated bilingual environment. Pioneering work conducted by Weiss and his colleagues have shown that monolingual adults can statistically segment two incongruent language inputs when learners are provided with strong contextual cues. Subsequent studies have later extended this work to bilingual populations and investigated whether there are differences between monolinguals and bilinguals in their abilities in statistical learning tasks. However, empirical findings have been mixed with some studies suggesting a bilingual advantage in statistical learning whereas other studies finding no differences between the two language groups. In addition, even when studies have found a bilingual advantage in statistical learning, they rarely directly investigate why such an advantage is present (i.e., by incorporating additional tasks that measure possible explanatory factors).

The current thesis seeks to address the above research gaps. Rather than a simple replication of prior work investigating dual-input statistical segmentation, the current studies aim to explore how a single speaker's phonetic cues can facilitate learners' segmentation from two linguistic inputs. As mentioned above, most prior studies have only investigated how between-person contextual cues, such as two different visual faces were paired with two languages (Mitchel & Weiss, 2010), facilitate dual-input statistical segmentation. It is thus largely unknown of how within-person contextual cues in bilingual environments help learners to segment words. By nature, a single bilingual person can speak two different languages. In a context where a bilingual speaker needs to communicate with people with different language backgrounds (e.g., some are Anglophones and others are Francophones), it is necessary for this individual bilingual to switch between two languages in order to successfully communicate. In addition, bilinguals often mix

their two languages when they communicate with other bilingual speakers, a phenomenon that is called *code-switching* in the linguistic literature (Poplack, 1980). As such, an individual bilingual often alternates between two languages within a single conversation in reality. Exploring how within-person contextual cues facilitate dual-input statistical segmentation therefore provides ecological understanding of how naturalistic contextual cues in bilingual environments play a role in helping learners to segment words correctly in different languages. In the current thesis, I will explore whether language-specific phonetic cues within a single speaker's speech (i.e., accented speech) is effective for learners to statistically segment two language inputs.

Further, I will specifically investigate what accounts for the potential bilingual advantage in statistical learning. As I have proposed, there are two possibilities: (i) bilinguals may enjoy an advantage because they have a stronger executive functioning system and/or (ii) bilinguals are more open to the interpretation that there are potential changes in the incoming language inputs and therefore are better in detecting two different language systems in novel environments. In this thesis, I will measure participants' cognitive ability and bilingual experience, and examine which factor better explains the potential bilingual advantage in dual-input statistical segmentation. In addition, previous studies that demonstrated a bilingual advantage have mostly investigated heterogeneous bilingual populations with various language backgrounds (e.g., Antovich & Graf Estes, 2018; Haesung, 2012; Poepsel & Weiss, 2016). It remains a possibility that this variability in linguistic backgrounds may confound and influence outcomes as some studies found bilingual children from different language backgrounds differed in phonetic processing (Bialystok, Majumder, & Martin, 2003) and metalinguistic awareness (e.g., applying English morphological rules to new words, such as the use of past-tense and formation of plural nouns) (Barac & Bialystok, 2012). Therefore, I will investigate homogenous French-English bilinguals in Ottawa, Canada. In

the following, I will outline how two proposed sets of studies in the current thesis accomplish the above-mentioned goals.

The current thesis consists of two studies investigating two different populations: Study 1 tests adults and Study 2 involves infants. Study 1 investigates whether adult learners can segment two interleaved artificial languages, using English- and French-accented speech as contextual cues to the transition. If accented speech is an effective contextual cue, I expect learners to successfully segment words from both languages. The focus of this study is to examine whether and how bilingual experience over the lifespan leads to better performance in the two-language statistical segmentation task. This will be accomplished by two additional manipulations.

First, I will investigate whether there is a difference in segmentation performance among three language groups: English monolinguals, simultaneous French-English bilinguals (participants who learned French and English since birth) and sequential French-English bilinguals (participants learned either French/English as second language after age 1). Participants' language backgrounds are assessed by a detailed language background questionnaire developed by Sabourin and colleagues (2016) (See Appendix A for details). The language background manipulation allows me to test whether the age of acquisition (AoA) of the second language may influence learners' processing of information when segmenting words from two overlapping languages.

Here, I focus on AoA³ under age 1. Infants refine their sensitivity to native and non-native phonetic cues by age 1 (Werker & Tees, 1984). Thus, I hypothesize that bilingual French and

³ I operationalized AoA as the age of first exposure (AoE) to the second language. It is different from the AoA criteria used in Sabourin and colleagues' work (2016) where they were using the age of immersion. I am using AoE because a number of participants were confused with respect to when they were "immersed" in the second language. For instance, a simultaneous bilingual reported that his AoE was age 0 but his AoI was age 12. This is because he thought AoI meant the year that he officially was taught the English writing system in school, but the AoI should be 0 as he was immersed in both French and English since birth. Thus, I believe AoE is a more accurate measure of the participants' age of acquiring their second language.

English experience during the first year of life would help learners better utilize their native phonetic cues in order to segment words from two overlapping languages. In addition, I also hypothesize that early bilingual experience may help learners to be more flexible when interpreting new linguistic input based on studies showing that bilingual infants demonstrate such flexibility (e.g., Graf Estes & Hay, 2015; Singh, 2017). If early bilingual experience leads participants to be more flexible in interpreting changes in language inputs, I expect early bilingual experience (e.g., simultaneous French-English bilinguals who learned two languages from birth) will outperform all other participants in the dual-input statistical segmentation task.

Furthermore, early bilingual learners may have a stronger executive function system that supports their performance in the statistical learning task. Several recent studies have shown that early, but not late, bilingual learners have better inhibitory skills in cognitive tasks like the attention network task (Luk, de Sa & Bialystok, 2011; Marzavocá, Asanowicz, Krivá & Wodniecka, 2013, Sabourin & Vīnerte, 2015; 2018). This early bilingual advantage may be driven by the prolonged experience of monitoring two languages in daily life (Bialystok et al., 2012). For example, Sabourin and Vīnerte (2018) found that the amount of bilingual experience participants had correlated with their executive control skills. Simultaneous bilinguals demonstrated stronger executive control skills in comparison to late bilinguals (i.e., bilinguals who learned their second language after age 8) and monolinguals. As such, simultaneous bilinguals, who have acquired two languages for a longer period of time, may have a stronger executive function system that supports the segmentation of words from two languages in the statistical learning task. If simultaneous bilinguals do have enhanced executive functioning processes, I expect that these augmented skills will positively predict their performance in the dual-input statistical segmentation task. This simultaneous bilingual advantage would be mediated by the bilinguals' executive functioning

skills in a mediation analysis. In order to test this hypothesis, apart from testing participants with a dual-input statistical learning task, I will additionally test all participants in Study 1 with five executive functioning tasks that measure participants' attentional components, inhibitory control ability, general working memory, verbal working memory, spatial working memory and task switching ability.

Study 2 is an extension of Study 1 to an infant population. This study contributes to the current literature in two ways. First, the study of adults' learning of two statistical structures is mimicking what bilingual infants actually experience in their language environment. In the dual-input segmentation task, adult participants were expected to statistically segment words merely by listening to novel "languages" in the absence of any knowledge of the incoming input or how many languages they need to learn. This situation is exactly what bilingual infants are facing in their everyday life when they are complete novices in their native languages and need to segment two different sets of language statistical structures. It is thus essential to extend this study to young infants to understand whether learners are capable of simultaneously learning two linguistic systems based on statistical cues at the beginning of language acquisition. Second, it can shed light on the possible advantage that bilingual infants may enjoy in statistically segmenting two speech streams at a very young age. To do so, I will test monolingual and bilingual infants at 9.5 months on an infant-adapted dual-input statistical segmentation task. Also, similar to the design of Study 1, the two interleaved languages in this task will be marked by English- and French-accented speech. Moreover, the same group of infant participants will be invited to participate in a second task that measures their executive function ability (Kovács & Mehler, 2009). If bilingual advantage is driven by better cognitive ability, I expect bilingual infants to perform better in the executive functioning task and their executive functioning performance will predict infants' ability in the

statistical segmentation task. If bilingual advantage relates to bilinguals being flexible in processing new language inputs, I expect bilingual experience during infancy, as measured by the infants' percentage of exposure to their two native languages, will predict infants' ability in the statistical segmentation task. Infants' percentage of exposure to their languages will be measured by a parental language questionnaire (Bosch & Sebastián-Gallés, 1997) (See Appendix B for details).

In closing, the current thesis investigates how learners segment words from two languages on the basis of transitional probabilities in a simulated bilingual environment. The research comprises two large studies across adult and infant populations with specific focus on the following questions: 1) Can learners make use of phonetic cues within a single individual's speech to segment words successfully in a dual-input segmentation task?; 2) Do bilinguals outperform monolinguals in this dual-input segmentation task?; and 3) Do specific factors, such as cognitive ability or bilingual experience (i.e., age of acquisition), underlie any potential bilingual advantage in the dual-input segmentation task? These questions can inform researchers about how learners tackle the challenge of segmenting words from two languages, which is a fundamental step of word learning in bilingual environments.

References

- Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., Carreiras, M. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Frontiers in Psychology*, 5, 398. doi: 10.3389/fpsyg.2014.00398
- Antovich, D. M., & Graf Estes, K. (2018). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science*, 21(2), e12548. doi: 10.1111/desc.12548
- Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117-134). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321-324. doi:10.1111/1467-9280.00063a
- Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, 15(7), 460-466. doi : 10.1111/j.0956-7976.2004.00702.x
- Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, 2, 324. doi: 10.3389/fpsyg.2011.00324
- Basseville, M., & Nikiforov, I. (1993). *Detection of abrupt changes – Theory and application*. Englewood Cliffs, NJ: Prentice-Hall.

- Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, 127(3), 391-397. doi: 10.1016/j.cognition.2013.02.011
- Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), 3253-3258. doi: 10.1073/pnas.1113380109
- Barac, R., & Bialystok, E. (2012). Bilingual Effects on Cognitive and Linguistic Development: Role of Language, Cultural Background, and Education. *Child Development*, 83(2), 413-422. doi: 10.1111/j.1467-8624.2011.01707.x
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, 70, 636-644. doi:10.1111/1467-8624.00046
- Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology*, 60(1), 68-79. doi: 10.1037/cjep2006008
- Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, 9(2), 117-121. doi: 10.1111/cdep.12116
- Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, 19, 290-303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, 24(1), 27-44. doi: 10.1017/S014271640300002X

- Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, 9, 103-119. doi: 10.1177/13670069050090010701
- Bialystok, E., Craik, F. I. M., & Luk, G. (2008a). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 859-873. doi: 10.1037/0278-7393.34.4.859
- Bialystok, E., Craik, F. I. M., & Luk, G. (2008b). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, 21, 522-538. doi: 10.1016/j.jneuroling.2007.07.001
- Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, 16, 240-250. doi: 10.1016/j.tics.2012.03.001
- Bogulski, C. (2013). *Are bilinguals better learners? A neurocognitive investigation of the bilingual advantage* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database (UMI No. 1511655851).
- Bosch, L., & Sebastián-Gallés, N. (1997). Native language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, 65, 33-69. doi: 10.1016/S0010-0277(97)00040-1
- Brent, Michael R., & Siskind, Jeffrey Mark. (2001). The Role of Exposure to Isolated Words in Early Vocabulary Development. *Cognition*, 81(2), B33-B44. doi: 10.1016/S0010-0277(01)00122-6

- Brito, N., & Barr, R. (2012). Influence of bilingualism on memory generalization during infancy. *Developmental Science*, 15(6), 812-816. doi:10.1111/j.1467-7687.2012.1184.x
- Brito, N. H., Sebastian-Galles, N., & Barr, R. (2015). Differences in language exposure and its effects on memory flexibility in monolingual, bilingual, and trilingual infants. *Bilingualism: Language and Cognition*, 18(4), 670-682. doi: 10.1017/S1366728914000789
- Bulgarelli, F., & Weiss, D. J. (2016). Anchors aweigh: The impact of overlearning on entrenchment effects in statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(10), 1621-1631. doi: 10.1037/xlm0000263
- Burns, T., Yoshida, K., Hill, K., & Werker, J. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28(3), 455-474. doi:10.1017/S0142716407070257
- 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., & 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
- Caldwell-Harris, C., Lancaster, A., Ladd, D., Dediu, D., & Christiansen, M. (2015). Factors influencing sensitivity to lexical tone in an artificial language: Implications for Second Language Learning. *Studies in Second Language Acquisition*, 37(2), 335-357. doi:10.1017/S0272263114000849

- Coderre, E. L., & van Heuven, W. J. B. (2014). The effect of script similarity on executive control in bilinguals. *Frontiers in Psychology*, 5, 1070. doi: 10.3389/fpsyg.2014.01070
- Cole, R., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.
- Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, 106, 59–86. doi: 10.1016/j.cognition.2006.12.013
- Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, 113, 135–149. doi: 10.1016/j.cognition.2009.08.001
- Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, 15(5), 336–45. doi: 10.1038/nrn3709
- de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, 26(1), 99–107. doi: 10.1177/0956797614557866
- De Diego Balaguer, R., Toro J.M., Rodriguez-Fornells A., Bachoud-Lévi A-C. (2007) Different Neurophysiological Mechanisms Underlying Word and Rule Extraction from Speech. *PLoS ONE* 2(11): e1175. doi:10.1371/journal.pone.0001175

- Duñabeitia, J., Hernández, J., Antón, E., Macizo, P., Estévez, A., Fuentes, L., & Carreiras, M. (2014). The Inhibitory Advantage in Bilingual Children Revisited: Myth or reality? *Experimental Psychology*, 61(3), 234-251. doi: 10.1027/1618-3169/a000243
- Erickson, L.C., Thiessen, E. D., & Graf Estes, K. (2014). Statistically coherent labels facilitate categorization in 8-month-olds. *Journal of Memory and Language*, 72, 49–58. doi: 10.1016/j.jml.2014.01.002
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language and Hearing Research*, 52(2), 321-335. doi: 10.1044/1092-4388(2009/07-0189)
- Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech Discrimination in 11-Month-Old Bilingual and Monolingual Infants: A Magnetoencephalography Study. *Developmental Science*, 20(1), e12427. doi: 10.1111/desc.12427
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5), i-185. doi: 10.1111/j.1540-5834.1994.tb00169.x
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., ... Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, 5, 11. doi: 10.3389/fpsyg.2014.00011

- Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science*, 33(6), 1087-1116. doi: 10.1111/j.1551-6709.2009.01041.x
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, 33(7), 1220-1234. doi: 10.3758/BF03193224
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18(3), 254–260. doi: 10.1111/j.1467-9280.2007.01885.x
- Graf Estes, K., & Hay, J. (2015). Flexibility in Bilingual Infants' Word Learning. *Child Development*, 86(5), 1371-1385. doi: 10.1111/cdev.12392
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81. doi: 10.1017/S1366728998000133
- Grey, S. E. (2013). *A neurocognitive investigation of bilingual advantages at additional language learning* (Doctoral dissertation). Available from ProQuest Dissertations and Theses Global (UMI No. 3590106).
- Grosjean, F. (2010). *Bilingualism: Life and Reality*. Cambridge, MA: Harvard University Press.
- Haesung, K. (2012). *Effects of Multiple Inputs with an Accent Cue in Speech Segmentation Facilitated by Language Experiences*. Unpublished honours thesis. Carnegie Mellon University, Pittsburgh, U.S.

- Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child Development, 86*(1), 10-22. doi: 10.1111/cdev.12269
- Hay, J., Pelucchi, B., Graf Estes, K., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology, 63*, 93–106. doi: 10.1016/j.cogpsych.2011.06.002
- Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science, 10*(5), 164–168. doi: 10.1111/1467-8721.00140
- Hilchey, M., & Klein, D. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*(4), 625-658. doi: 10.3758/s13423-011-0116-7
- Hilchey, M., Saint-Aubin, J., & Klein, R. (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. Schwieter (Ed.), *Cambridge Handbook of Bilingual Processing*, pp. 586-613. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781107447257.026
- 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
- Hudon, T., E. (2013). *Bilingual Infants' Accommodation of Accented Speech*. (Doctoral dissertation). Retrieved from:

https://www.ruor.uottawa.ca/bitstream/10393/26163/1/Hudon_Tamara_2013_thesis.pdf

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, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-99. doi: 10.1016/0010-0285(89)90003-0

Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39(3), 215-258. doi: 10.1016/0010-0277(91)90054-8

Jusczyk, P. W., & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology*, 23(5), 648-654.

Jusczyk, P. W., Kennedy, L. J., & Jusczyk, A. M. (1995). Young infants' retention of information about syllables. *Infant Behavior and Development*, 18(1), 27-41. doi: 10.1016/0163-6383(95)90005-5

Kandhadai, P., Hall, D. G., & Werker, J. F. (2017). Second label learning in bilingual and monolingual infants. *Developmental Science*, 20(1), e12429. doi: 10.1111/desc.12429

Karuza, E. A., Balewski, Z. Z., Hamilton, R. H., Medaglia, J. D., Tardiff, N., & Thompson-Schill, S. L. (2016). Mapping the Parameter Space of tDCS and Cognitive Control

- via Manipulation of Current Polarity and Intensity. *Frontiers in Human Neuroscience*, 10, 665. doi: 10.3389/fnhum.2016.00665
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83(2), B35–B42. doi : 10.1016/S0010-0277(02)00004-5
- Kousaie, & Phillips. (2012a). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research*, 1446, 71-90. doi: 10.1016/j.brainres.2012.01.052
- Kousaie, S., & Phillips, N. A. (2012b). Ageing and bilingualism: Absence of a “bilingual advantage” in Stroop interference in a nonimmigrant sample. *Quarterly Journal of Experimental Psychology*, 65(2), 356-369. doi: 10.1080/17470218.2011.604788
- Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience*, 8, 250. doi: 10.3389/fnbeh.2014.00250
- Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, 106(16), 6556-6560. doi: 10.1073/pnas.0811323106
- Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science*, 325(5940), 611-612. doi: 10.1126/science.1173947
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12

months. *Developmental Science*, 9(2), F13-F21. doi: 10.1111/j.1467-7687.2006.00468.x

Lany, J., & Saffran, J. R. (2013). Statistical learning mechanisms in infancy. In Rubenstein, J. L. R. & Rakic, P. (eds.) *Comprehensive Developmental Neuroscience: Neural Circuit Development and Function in the Brain*, 3, pp. 231-248. Amsterdam: Elsevier. doi: 10.1016/B978-0-12-397267-5.00034-0

Luo, L., Craik, F.I.M., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and Aging*, 28, 28-34. doi: 10.1037/a0030875

Luk, G., Anderson, J.A.E., Craik, F.I.M., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition*, 74, 347-357. doi: 10.1016/j.bandc.2010.09.004

Luk, G., de Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, 14(4), 588-595. doi: 10.1017/S1366728911000010

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

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 11, 81-93. doi: 10.1017/S1366728907003227

- Marzecová, A., Asanowicz, D., Krivá, L., & Wodniecka, Z. (2013). The effects of bilingualism on efficiency and lateralization of attentional networks. *Bilingualism: Language and Cognition*, 16(3), 608-623. doi: 10.1017/S1366728912000569
- Mayberry, R., & Fischer, I. (1989). Looking through phonological shape to lexical meaning: The bottleneck of non-native sign language processing. *Memory & Cognition*, 17(6), 740-754. doi: 10.3758/BF03202635
- Mayberry, R. I., & Lock, E. (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain and Language*, 87(3), 369-384. doi: 10.1016/S0093-934X(03)00137-8
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–B111. doi: 10.1016/S0010-0277(01)00157-3
- Merriman, W., & Bowman, L. (1989). The Mutual Exclusivity Bias in Children's Word Learning. *Monographs of the Society for Research in Child Development*, 54(3/4), 1-129. doi:10.2307/1166130
- Miller, J. L., & Eimas, P. D. (1995). Speech perception: From signal to word. *Annual Review of Psychology*, 46(1), 467-492. doi: 10.1146/annurev.ps.46.020195.002343
- Mitchel, A. & Weiss, D.J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, 25(4), 456-482. doi: 10.1080/01690960903209888

- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100. doi: 10.1006/cogp.1999.0734
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21, 8–14. doi: 10.1177/0963721411429458
- Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 114, 187-202. doi: 10.1016/j.jecp.2012.09.002
- Mueller, J. L., Girgsdies, S., & Friederici, A. D. (2008). The impact of semantic-free second-language training on ERPs during case processing. *Neuroscience Letters*, 443(2), 77-81. doi: 10.1016/j.neulet.2008.07.054
- Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13(3), 407-420. doi: 10.1111/j.1467-7687.2009.00898.x
- Nation, R., & Mclaughlin, B. (1986). Novices and experts: An information processing approach to the “good language learner” problem. *Applied Psycholinguistics*, 7(1), 41-55. doi:10.1017/S0142716400007177
- Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, 14(1), 11-28. doi: 10.1016/0364-0213(90)90024-Q

- Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 481-502). Cambridge, MA, US: The MIT Press.
- Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, 115(2), 357–395. doi: 10.1037/0033-295X.115.2.357
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232-258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K., & Sawi, O. (2014). Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology*, 5, 962. doi: 10.3389/fpsyg.2014.00962
- Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge” hypothesis as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, 121(2), 130–143. doi: 10.1016/j.bandl.2011.05.003
- Perruchet, & Pacton. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10(5), 233-238. doi: 10.1016/j.tics.2006.03.006
- Perruchet, & Vinter. (1998). PARSER: A Model for Word Segmentation. *Journal of Memory and Language*, 39(2), 246-263. doi: 10.1006/jmla.1998.2576

- Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, 152, 9-19. doi: 10.1016/j.cognition.2016.03.001
- 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
- Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, 17(04), 682-691. doi: 10.1017/S1355617711000580
- Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13, 253–262. doi: 10.1017/S1366728909990526
- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to Represent a Multi-Context Environment: More than Detecting Changes. *Frontiers in Psychology*, 3, 228. doi: 10.3389/fpsyg.2012.00228
- Saffran, J. R. (2003). Statistical language learning mechanisms and constraints. *Current Directions in Psychological Science*, 12(4), 110–114. doi: 10.1111/1467-8721.01243
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928. doi : 10.1126/science.274.5294.1926
- Saffran, J. R., Hauser, M., Seibel, R. L., Kapfhamer, J., Tsao, F., Cushman, F. (2008). Grammatical pattern learning by infants and cotton-top tamarin monkeys. *Cognition*, 107, 479-500. doi: 10.1016/j.cognition.2007.10.010

- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27-52. doi: 10.1016/S0010-0277(98)00075-4
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621. doi: 10.1006/jmla.1996.0032
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2), 101-105. doi: 10.1111/j.1467-9280.1997.tb00690.x
- Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, 39, 484-494. doi: 10.1037/0012-1649.39.3.484
- Saffran, J.R., & Thiessen, E.D. (2007). Domain-general learning capacities. In E. Hoff & M. Shatz (Eds.), *Handbook of Language Development*. Cambridge: Blackwell (p. 68-86).
- Sabourin, L., Leclerc, J-C., Lapierre, M., Burkholder, M., and Brien, C. (2016). *The Language Background Questionnaire in L2 Research: Teasing Apart the Variables*. Proceedings of the Canadian Linguistic Association. Available online: http://cla-acl.ca/wp-content/uploads/actes-2016/Sabourin_etal_CLA2016_proceedings.pdf.
- Sabourin, L., & Vinerte, S. (2015). The bilingual advantage in the Stroop task: simultaneous vs. early bilinguals. *Bilingualism: Language and Cognition*, 18(2), 350-355. doi: 10.1017/S1366728914000704

- Sabourin, L., & Vinerte, S. (2018). Cognitive control among immersed bilinguals: Considering differences in linguistic and non-linguistic processing. *Bilingualism: Language and Cognition*, 1-16. doi: 10.1017/S1366728918000524
- Sebastián-Gallés, N. (2010). Bilingual Language Acquisition: Where Does the Difference Lie? *Human Development*, 53(5), 245-255. doi: 10.1159/000321282
- Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W., & Werker, J. (2012). A Bilingual Advantage in Visual Language Discrimination in Infancy. *Psychological Science*, 23(9), 994-999. doi: 10.1177/0956797612436817
- Singh, L. (2017). Bilingual infants demonstrate advantages in learning words in a third language. *Child Development*. doi: 10.1111/cdev.12852. [Epub ahead of print].
- Singh, L., Fu, C., Rahman, A., Hameed, W., Sanmugam, S., Agarwal, P., . . . Rifkin-Graboi, A. (2015). Back to Basics: A Bilingual Advantage in Infant Visual Habituation. *Child Development*, 86(1), 294-302. doi: 10.1111/cdev.12271
- Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences*, 18(5), 251–258. doi: 10.1016/j.tics.2014.02.007
- Soto-Faraco, S., Navarra, J., Weikum, W., Vouloumanos, M., Sebastián-Gallés, A., & Werker, N. (2007). Discriminating languages by speech-reading. *Perception & Psychophysics*, 69(2), 218-231. doi: 10.3758/BF03193744

- Sundara, M., Polka, L., & Molnar, M. (2008). Development of Coronal Stop Perception: Bilingual Infants Keep Pace with Their Monolingual Peers. *Cognition*, 108(1), 232-242. doi: 10.1016/j.cognition.2007.12.013
- Swingle, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology*, 50(1), 86-132. doi: 10.1016/j.cogpsych.2004.06.001
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 12530–12535. doi: 10.1073/pnas.0609927104
- Thiessen, E. D. (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Phil. Trans. R. Soc. B*, 372, 20160056. doi: 0.1098/rstb.2016.0056
- Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin*, 139(4), 792–814. doi: 10.1037/a0030801
- Thiessen, E. D., & Pavlik, P. I. (2013). iMinerva: a mathematical model of distributional statistical learning. *Cognitive Science*, 37(2), 310-343. doi: 10.1111/cogs.12011
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology*, 39(4), 706–716. doi: 10.1037/0012-1649.39.4.706

- Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development*, 3(1), 73–100. doi: 10.1080/15475440709337001
- Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, 10(2), 172-175. doi: 10.1111/1467-9280.00127
- Tincoff, R., & Jusczyk, P. W. (2012). Six-Month-Olds comprehend words that refer to parts of the body. *Infancy*, 17(4), 432-444. doi: 10.1111/j.1532-7078.2011.00084.x
- Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2005). Effects of backward speech and speaker variability in language discrimination by rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 31(1), 95-100. doi: 10.1037/0097-7403.31.1.95
- Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97(2), B25-B34. doi: 10.1016/j.cognition.2005.01.006
- van de Weijer, J. (1998). *Language input for word discovery*. Nijmegen: Max Plank Institute for Psycholinguistics.
- Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, 5, 953. doi: 10.3389/fpsyg.2014.00953

- Weikum, W., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. (2007). Visual language discrimination in infancy. *Science*, 316(5828), 1159. doi: 10.1126/science.1137686
- Weiss, D.J., Gerfen, C. & Mitchel, A.D. (2009) Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development* 5, 30-49. doi: 10.1080/15475440802340101
- Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2010). Colliding cues in word segmentation: the role of cue strength and general cognitive processes. *Language and Cognitive Processes*, 25(3), 402–422. doi: 10.1080/01690960903212254
- Weiss, D. J., Poepel, T., & Gerfen, C. (2015). Tracking multiple inputs: The challenge of bilingual statistical learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages*, pp. 167-190. Amsterdam, the Netherlands: John Benjamins Publishing Company.
- 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
- Yan, S., & Nicoladis, E. (2009). Finding le mot juste: Differences between bilingual and monolingual children's lexical access in comprehension and production. *Bilingualism: Language and Cognition*, 12, 323–335. doi: 10.1017/S1366728909990101

- Yim, D., & Rudoy, J. (2013). Implicit statistical learning and language skills in bilingual children. *Journal of Speech Language and Hearing Research*, 56(1), 310–322. doi: 10.1044/1092-4388(2012/11-0243)
- Zinszer, B. D., & Weiss, D. J. (2013). When to hold and when to fold: Detecting structural changes in statistical learning. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3858–3863). Austin, TX: Cognitive Science Society.

Statistical Word Segmentation from Accented Speech: A Bilingual Advantage

Sin Mei Tsui^{1*}

Lucy C. Erickson^{2,3}

Erik D. Thiessen²

Christopher T. Fennell¹

1. School of Psychology, University of Ottawa. 2. Carnegie Mellon University. 3. University of Maryland

*Corresponding author

Abstract

Previous studies have demonstrated that learners succeed in segmenting two overlapping artificial languages when supported by strong speaker-specific contextual cues (i.e., each language paired with a different person). However, by nature, a single bilingual individual has to alternate between two distinct languages, sometimes even in one conversation (i.e., code-switching). It is therefore important to explore whether phonetic contextual cues in a single individual's speech can facilitate adult segmentation of two artificial languages. In the current study, we explore whether phonetic variation cues within a single individual's speech, which we term accented speech, can facilitate adult segmentation of two artificial languages. We presented two artificial languages that differed in terms of the phonetic cues found in either Canadian English or French to four groups of participants: English monolinguals, simultaneous French-English bilinguals (learned both languages before age 1), sequential French-English bilinguals (learned their second language after age 1) and multilinguals (proficient in French, English and other language(s)). Overall, participants successfully segmented two languages only when the second artificial language was presented in an English accent, indicating difficulties in processing certain phonemes in French-accented speech. We found a simultaneous bilingual advantage: simultaneous bilinguals, but not sequential bilinguals or multilinguals, segmented more correct words when compared to monolinguals. We conclude that accented speech can facilitate the separation of two language inputs in adults. Our evidence for a simultaneous bilingual advantage suggests that early bilingual experience may help learners to detect multiple structures and promptly adapt to phonetic variability in a novel language environment.

Keywords: Statistical learning; Word segmentation; Bilingualism; Cross-linguistic phonetic cues; Accented speech

Statistical Word Segmentation from Accented Speech: A Bilingual Advantage

Finding words in the speech stream is not a simple task. Unlike the white spaces that mark boundaries between words in a written passage, pauses do not reliably and consistently mark word boundaries in continuous speech (Cole & Jakimik, 1980). Statistical learning, a mechanism that allows learners to track regularities in sensory input from the ambient environment, has gained traction over the past two decades as a solution to the above word segmentation problem (e.g. Saffran, 2003). Specifically, one type of statistical regularity that has been proposed to play a significant role in solving the word segmentation problem is transitional probability: the probability of event Y given event X. For example, consider a short phrase in English - *happy baby*. There is high transitional probability that the syllable *hap* precedes the syllable *py* because they constitute a word. But there are many words that could potentially occur after *happy* (e.g., happy boy, happy girl and etc). Thus, it is much less probable that the syllable *py* would be followed by the syllable *ba*. Many prior studies have shown that infants and adults are sensitive to the transitional probabilities between syllable sequences and can segment words from continuous speech based on transitional probability (e.g., Aslin, Saffran & Newport, 1998; Saffran, Aslin & Newport, 1996a; Saffran, Newport & Aslin, 1996b). However, most studies to date have examined how learners' sensitivity to transitional probability helps them to segment words from a single language input. However, outside of the lab, many learners encounter two or multiple language inputs, such as infants learning two languages at home. Thus, it is important to extend this research to discover how learners track two different language inputs in a statistical segmentation task.

To begin, it is, in all probability, more difficult for learners to statistically segment two language inputs in comparison to a single language input. There is an intrinsic challenge in a bilingual environment when using transitional probability to segment words from two input

languages. Syllables across two languages tend to overlap, and such phonological overlap can generate conflicting statistics across the two languages. For example, suppose a bilingual learner hears the syllable co-occurrence of *mo* and *to*. In French, *moto* is an actual word, meaning motorcycle (e.g., *Une belle **moto***. A beautiful motorcycle.). These two syllables will co-occur with greater probability since they form a word. The syllable *to* followed by another syllable would be less probable, thus marking the word boundary. However, in English, *moto* is not a word. Thus, the transition from *mo* to *to* would not be as probable as in French, indicating a word boundary at *mo* (e.g., *El**mo** told Ernie*). Hence, to acquire two different languages successfully, learners need to differentiate between two inputs and form separate representations of each language.

Weiss and his colleagues (2009) first began to address the above question by examining how adult learners statistically segment words from two language inputs. Participants were presented with two interleaved artificial languages across two conditions: congruent and incongruent. In the congruent condition, the transitional probabilities across languages were not in conflict with each other. Specifically, the languages did not share any syllables and therefore the statistics from one language did not interfere with those in the other, learners therefore could combine statistics across both languages without any impact on their word segmentation performance. On the other hand, in the incongruent condition, the two languages shared some syllables and these syllables were located at different positions across the two language streams. For example, *bo* was a within-word syllable in one language, but it was a word-boundary syllable in the other language, mimicking our previous example *moto* in French and English. Thus, the transitional probabilities across two languages were incompatible with each other because of syllable overlap. Participants succeeded in segmenting words in each language in the congruent condition without the presence of any contextual cues. However, adults only learned the two

conflicting statistical structures and segmented words in the incongruent condition when they were supported by strong contextual cues. For example, adults demonstrated learning when the two languages were associated with individuals of different gendered voices (Weiss, Gerfen & Mitchel, 2009) or two distinct faces (Mitchel & Weiss, 2010).

Other researchers have shown that it is possible for adult learners to segment words from two overlapping languages without contextual cues, although performance appears to suffer in their absence. These researchers (Basseville & Nikiforov, 1993; Gebhart, Aslin & Newport, 2009; Qian, Jaeger, & Aslin, 2012) argued that learners can learn two conflicting language structures by perceptually monitoring the consistency of inputs. When learners receive two conflicting language structures, they will notice the statistics cues vary over time. If the learners treat this variability as noise, they will fail to separate the two language structures. However, if the learners treat this variability as signals of changes in the underlying structures of the sensory input, they will eventually detect that there are two language structures and successfully track the transitional probability in each individual language. Experimental findings have suggested this is possible, but that adults generally only learned the first heard language during training. They failed to learn the second heard language when they did not receive any contextual cues, thus exhibiting a primacy learning effect (Bulgarelli & Weiss, 2016; Gebhart et al., 2009; Zinszer & Weiss, 2013). Taken together, the above findings suggest that adults are able to segment two different incongruent statistical language inputs in the absence of contextual cues, but their performance is limited. Thus, learners generally perform best when salient contextual cues that mark the transition of languages are provided.

However, it is important to note that previous studies have only explored how speaker-specific contextual cues facilitate adult word segmentation from two languages (e.g. two voices

differing in gender or two different talker faces; see Weiss, Poepsel & Gerfen, 2015, for a review). In naturalistic bilingual environments, two distinct languages may not always be spoken by two different speakers. Obviously, a single bilingual speaker would produce both of their languages across situations. Even in a single conversation, bilinguals often alternate between two languages when speaking with other bilinguals, a phenomenon that is known as code-switching (e.g., Poplack, 1980; see Heredia & Altarriba, 2001, for a review). As such, this leads to a question of whether other contextual cues, especially those that might be present in naturally occurring bilingual environments, can also facilitate learners' segmentation of two languages within the same paradigm. To fill in this research gap, we will explore whether contextual cues in a single individual's speech can facilitate adult segmentation of two artificial languages. In our paper, we focus on how phonetic cues, which we term accented speech, can facilitate learning in a dual-input segmentation task. Phonetic realizations of the same syllables often vary across two languages. Consider our previous example about the word *moto*, this word has the same phonemes across French and English, but the phonetic properties differ across the languages. *Moto* would be produced with tenser vowels and less duration in French than in English. Therefore, the phonetic variations across languages could be considered as distinguishable indexical cues that signal a change of languages within a single individual's speech. If these are effective contextual cues that mark the transitions between two language inputs, we expect participants will be able to segment two languages successfully in the dual-input statistical task.

Another issue present in the past studies is that they were predominantly testing monolingual populations. The experimental design in Weiss and colleagues' studies is necessarily simulating a bilingual environment where syllables overlap and conflicting linguistic statistics are present across two languages. If bilingual learners regularly need to manage and handle two

language systems in their daily environment, bilingual experience may help learners to statistically segment words from two conflicting languages better. Indeed, there is burgeoning literature investigating whether bilinguals enjoy advantages in statistical learning in comparison to monolinguals.

Regarding the dual-input segmentation task, there are at least two studies suggesting that bilinguals may outperform monolinguals. Bartolotti, Marian, Schroeder and Shook (2011) examined whether adult learners can segment two artificial Morse code languages in both low and high interference conditions. In the low interference condition, statistics-based word boundaries were paired with pauses that aligned with said boundaries; whereas, in the high interference condition, the pauses did not align with statistics-based word boundaries. Bartolotti et al (2011) found that bilingual experience was correlated with better segmentation performance in the low interference condition while inhibitory control was correlated to better segmentation performance in the high interference condition. This suggests that bilingual experience may improve learners' performance in a dual-input segmentation task but this improvement may not be sufficient to support segmentation under high cognitive demands. More recently, Antovich and Graf Estes (2018) have shown that bilingual infants of 14 months successfully segmented two language inputs when the languages were marked by two different gendered voices but their monolingual counterparts failed in the same task. Together, these two studies have suggested that bilingual experience may help learners perform better in a dual-input statistical segmentation task. However, we should note that, while some other studies have found that bilinguals outperformed monolinguals in statistical learning tasks (e.g., Poepsel & Weiss, 2016; Wang & Saffran, 2014), others found no evidence supporting a difference between these groups (Bogulski, 2013; Grey, 2013; Yim & Rudoy, 2013). The mixed findings in the literature, can partially, be explained by

the differences in methodology such as the use of different statistical learning tasks (some used statistical segmentation task, others used visual statistical learning task).

If there is a bilingual advantage in segmenting words from two languages, the underlying mechanism of why bilinguals outperform monolinguals in statistical segmentation tasks has yet to be researched. One possible mechanism is that bilinguals may be better in the cognitive processing skills known as executive functioning: attentional control, inhibitory control, working memory, and cognitive flexibility. This is based on the argument that bilinguals inevitably activate both of their languages when speaking. Thus, bilinguals regularly need to suppress the irrelevant language in order to speak the relevant language (Bialystok, 2001; see Bialystok, Craik, & Luk, 2012 for a review). This regular suppression of the language not currently being used may help bilinguals develop stronger executive function processing (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004; Costa, Hernández, Sebastián-Gallés, 2008; Chung-Fat-Yim, Sorge, & Bialystok, 2017; Morales, Yudes, Gomez-Ariza, & Bajo, 2015; Salvatierra & Rosselli, 2011). Prior research has suggested that better cognitive processing skills, such as attention and inhibitory control, aid learners in segmenting words in statistical learning tasks (e.g. Toro, Sinnett, & Soto-Faraco, 2005; Weiss, Gerken, & Mitchel, 2010). As such, bilinguals' enhanced executive function processing may help them to track and segment two statistical linguistic inputs and therefore bilinguals outperform monolinguals in statistical segmentation task.

An alternative, but not necessarily mutually exclusive, account is that specific bilingual experience may explain the advantage in statistical learning. The bilingual environment is arguably more variable. This variation can include: learning two similar languages with substantial overlapping sound contrasts (e.g., Spanish and Catalan); learning from speakers with foreign accents; and/or learning in a language-mixing environment (e.g., code-switching; see Byers-

Heinlein & Fennell, 2014, for a review). Importantly, variability in a bilingual environment may signal a change in language inputs. For example, when a proficient speaker alternates between two languages in a conversation, the phonetic cues vary across the two languages, signaling a change in language inputs within this conversation. Here, bilinguals need to be flexible when coping with the variability in this conversation. If they think that the phonetic variability is just noise, they would fail to separate the languages. This experience of handling the variability in bilingual environments may have prepared learners to be more flexible when interpreting variability in a new language environment. In a dual-input segmentation task, when learners initially combine two overlapping languages together, they necessarily receive unstable statistical cues to word boundaries (e.g., sometimes *moto* is a word (French) and sometimes *moto* is a word boundary (English)). This consistent variability in statistical cues is a signal that the underlying language structure may have changed. If bilingual learners are more flexible in treating the variability of statistics cues as signals of possible changes in language inputs instead of noise, they will be more likely to differentiate between two language inputs and start segmenting words from each language separately.

To sum up, our goal is to first investigate whether a natural intra-speaker cue to language change – the phonetic cues across languages (i.e., accented speech) – can facilitate statistical segmentation from two language inputs. We specifically used Canadian French- and Canadian English-accented speech to mark the differences in two artificial languages because they are local accents in the participants' community. We expected that our participants would be familiar with both Canadian French and English accents and could readily utilize these cues for segmenting words from the artificial languages.

In addition, we investigated how bilingualism influences performance in dual-input statistical task. To examine whether bilinguals enjoy a cognitive advantage in statistical learning, we tested participants with a battery of cognitive tasks, which include (i) color-shape switching task (Prior & MacWhinney, 2010), (ii) attention network task (Fan et al., 2002), (iii) operational span task, (iv) symmetry span task and (v) reading span task (Oswald, McAbee, Redick & Hambrick, 2015). This battery of tasks captured a wide range of executive functioning abilities, allowing us to investigate which cognitive aspect(s) contribute(s) to bilinguals' potential outperformance in segmentation task. Further, the above tasks have all been used in previous studies to investigate bilingual cognitive advantages, which allows for direct comparisons to past work (e.g., Prior & MacWhinney, 2010, Costa, Hernández & Sebastián-Gallés, 2008; Ratiu & Azuma, 2015).

Finally, we rigorously controlled for and assessed bilinguals' backgrounds. Bilingualism is a continuum that varies not only in length of exposure but also in different degrees of bilingual experience (Kroll & Bialystok, 2013). The linguistic environments of two bilinguals who have learned their second language at the same time can be substantially different in terms of their proficiency in the two languages, how they acquired their languages, and how frequently they use the languages. Thus, we controlled for the following aspects of bilingualism in our study. First, we only tested French-English bilinguals in the current study. This allowed us to control for possible confounds, such as variability of language-specific properties (e.g., phonemes, phonotactics, syllable frequency, etc.) across different linguistic backgrounds (e.g., Barac & Bialystok, 2012; Bialystok, Majumder, & Martin, 2003). Crucially, we explored an important factor along the bilingualism continuum: the age of acquisition of the second language (AoA), which we counted as their first exposure to the second language. A number of studies have demonstrated that the

AoA of the second language influences bilingual advantages in linguistic and cognitive tasks (Tao et al., 2011; Sabourin & Vinerte, 2015; Sabourin, Brien & Burkholder, 2014; Silverberg & Samuel, 2004). In our paper, AoA is not treated as a continuous variable but a categorical variable as this is how previous studies (e.g., Sabourin & Vinerte, 2015; Sabourin et al., 2014) have measured the influence of AoA in linguistic and cognitive tasks. We used a cut-off (age 1) to divide our bilingual participants into two groups based on their AoA: simultaneous bilinguals who learned English and French between ages 0 to 1 and sequential bilinguals who learned English and French after age 1. We set the AoA cut-off at age 1 because native phonetic development is mostly complete by this age (Kuhl et al., 2006; Werker & Tee, 1984) and our central task involves phonetic cues. Bilinguals, who learned two language before age 1, may have more well-refined phonetic perception of French and English since they were exposed to these two languages during the critical period of refinement. Thus, they may be better able to utilize English and French phonetic cues when segmenting words in our experiment. Furthermore, we carefully matched the two bilingual groups with their proficiency levels and frequent usage of the two languages (measured by average self-reported percentage of using the two languages in daily life). Proficiency levels and use of the two languages reflect how bilinguals learned and use their second language in a naturalistic or instructional context (Sabourin, Burkholder, Vinerte, Leclerc & Brien, 2016), which can affect how bilinguals process their second language, including its phonetic aspects (e.g., Flege, & Liu, 2001; Guion, Flege, Akahane-Yamada & Pruitt, 2000).

In short, we hypothesize that bilinguals will outperform monolinguals in the dual-input statistical segmentation task. Further, if the bilingual advantage is driven by enhanced executive function abilities, we expect that their performance on cognitive tasks will positively predict their segmentation accuracy in the dual-input segmentation task. Finally, if specific bilingual experience

(e.g., AoA) contributes to an advantage in segmentation, we expect that simultaneous bilinguals will outperform both sequential bilinguals and monolinguals in the segmentation task due to their experience with the phonetic cues of both languages during the key period for phonetic acquisition.

Experiment 1

Method

Participants. The final sample consisted of 115 adults (62 female). See Table 1 for participant characteristics. In this experiment and all other experiments (Experiment 2A and 2B), we recruited participants from the undergraduate research pool at the University of Ottawa (the Integrated System of Participation in Research System, or ISPR) and from the community by posting recruitment blurbs online. If participants were recruited from the, ISPR, they were provided with refreshments and also compensated with course credits. If participants were recruited from the community, they were provided with refreshments and a chance of winning a gift card of CAD 25. The Language Background Questionnaire (LBQ) (Sabourin et al., 2016) was used to measure participants' self-rated language proficiency level and how often they use their language(s) in daily life. Participants also completed an English cloze test (Brown, 1980) and a French cloze test (Tremblay, 2011), which were two objective measurements of their English and French proficiencies. Based on the LBQ questionnaires and participants' scores on these cloze tests, we identified four groups of participants: 33 functionally English monolinguals with little knowledge of French, 32 simultaneous French-English bilinguals who learned both languages before age 1 (mean English exposure during infancy = 44.7%; $SD = 16.7\%$), 34 sequential French-English bilingual adults who learned their second language (either French or English) between age 1 to age 12, and 16 multilingual adults who were proficient in English, French and other language.

Here, we categorized participants into different language groups based on their proficiency levels. All participants were proficient in English, the dominant language of the community. When participants indicated that they learned French, they would be categorized as English monolinguals if their scores in the French cloze test were below 30%. Participants would only be categorized as English-French bilinguals if their scores in the French cloze test were at least 30% or above. We admit that this was a post-hoc decision, as there was no clear reference as to where the cutoff should be in a self-declared monolingual population that always has exposure to the second language during schooling (as in Canada). After obtaining LBQ and cloze test scores, we discovered that no self-declared monolingual, according to their LBQ, had a cloze test score that exceeded 30%. We therefore set the French cloze test criterion to 30% to delineate English monolinguals from English-French bilinguals. Furthermore, a segment of our participant sample indicated that they learned a language in addition to English or French. We felt that these individuals could not be classified as simultaneous/sequential bilinguals in the study. Given that adults' self-rated and objective proficiency are highly correlated (e.g., Marian, Blumenfeld & Kaushanskaya, 2007; Sabourin & Vinerte, 2015), we categorized these participants as a separate group, called multilinguals, if their self-rating proficiency of this third language were at least 11 out of 24, as we did not have objective cloze tests for other languages. If a participant's self-rating proficiency of his/her third language were below 11, we would categorize this participant as a French-English bilingual speaker (the French cloze test is at least 30% for these participants as well). We used this criterion because scores under 11 in the long LBQ indicate low or intermediate self-proficiency in each language aspect (e.g., oral comprehension, oral production, etc.)

For the bilingual participants' proficiency levels, there were no significant differences in cloze task scores between simultaneous bilinguals, sequential bilinguals, and multilinguals ($F_s <$

2.6, $ps > 0.08$). In addition, the average percentage of English usage also did not differ significantly among simultaneous bilinguals, sequential bilinguals and multilinguals ($F_s < 1.8$, $ps > 0.18$). Thus, English and French proficiency, as well as the average percentage of usage of English and French, were similar among simultaneous bilinguals, sequential bilinguals and multilinguals. An additional 8 participants were tested but not included in the final sample due to the following reasons: diagnosed with phonological disorder during childhood ($n = 1$), had completed a previous experiment involving artificial language ($n = 1$), did not follow instructions from experimenter ($n = 2$), or withdrew from the study in the second session ($n = 4$).

Experimental Materials, Designs and Procedure. The experiment consisted of a battery of tasks, including one statistical segmentation task, five executive function measures (details described below), the LBQ, and French and English cloze tests. All instructions in the tasks (both written and verbal) were in English.

Participants completed all the tasks across two sessions, each conducted on a separate day. Each session lasted for approximately 70 minutes and the task order was fixed. In the first session, participants completed the statistical segmentation task first, followed by a color-shape switching task, and then the LBQ at the end of the session. Participants were offered a choice of completing the LBQ at home. If they did so, they were instructed to return the LBQ on the second session. In the second session, participants completed the attention network task first, followed by the operation span task, the symmetry span task, and the reading span task. At the end of the second session, participants completed the French and English cloze tests. The order of all tasks was fixed because this is the standard practice for studies examining individual differences using multiple tasks that are correlational in nature (Carlson & Moses, 2001). Experiments in the first and second session were completed in different testing rooms due to the specific locations where specific

experiments were programmed to run, but there were minimal differences between the set-ups (i.e., 1-inch difference in monitor size, difference in operating systems – Windows 7 versus Vista).

The following information contains details of each task in the study:

Statistical segmentation task. The task was programmed by using E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). All audio stimuli were adopted from scripts used in the incongruent condition of Experiment 2 in Weiss et al (2009). We only made one phoneme change from the original script, the vowel /æ/ became /a/ in our scripts, so that all phonemes co-exist in both Canadian English and Canadian French (Table 2). This adjustment ensured that our speaker did not add, substitute or delete phonemes in the scripts. The speech stream was composed of two artificial languages, termed Artificial Language 1 (AL1) and Artificial Language 2 (AL2), with each language consisting of four trisyllabic CV.CV.CV words. Each word in each language was presented an equal number of times and no words were repeated twice in a row. AL1 and AL2 were concatenated in an interleaved order. The marker transition cues across AL1 and AL2 were English and French accents. Therefore, AL1 and AL2 were always in different accents. For instance, English-accented AL1 was paired with French-accented AL2, whereas French-accented AL1 was paired with English-accented AL2. The transitional probability signaling the word boundaries in each language were 1.0 within word and 0.33 across words, presuming that participants can separate the two languages and calculate the statistical set of each. However, if participants were to combine the two languages, the transitional probability would not provide consistent statistical cues: the transitional probability within words would vary from 0.5 to 1.0 whereas the transitional probability across word boundaries would vary from 0.17 to 0.33 (see Figure 1). Apart from statistical cues, there were no pauses or other acoustic cues (e.g. stress) between syllables within each language to signal any word boundaries.

A balanced French-English bilingual produced the audio stimuli. She was instructed to produce the two languages once as if she were speaking Canadian English and once as if she were speaking Canadian French. Twelve local English-French native bilinguals were recruited to rate the stimuli on a scale of 1 to 5 in four aspects: (i) naturalness (1 indicating unnatural-sounding and 5 indicating natural-sounding); (ii) French-accent rating: whether the language sounds like French (1 indicating very French and 5 indicating very non-French); (iii) English-accent rating: whether the language sounds like English (1 indicating very English and 5 indicating very non-English); and (iv) overall sound quality of the language (1 indicating bad and 5 indicating excellent). The English-accented languages had an average of 3.02 for naturalness, 4.38 for French-accent rating (higher rating means the stimuli sounded very non-French), 1.75 for English-accent rating (lower rating means the stimuli sounded very English) and 3.38 for overall sound quality. The French-accented language had an average of 3.17 for naturalness, 1.90 for French-accent rating, 4.04 for English-accent rating, and 3.59 for overall sound quality.

In the experiment, participants were instructed to listen and attend to a set of syllables. They were informed that they would subsequently be tested on what they had heard but did not receive any further information regarding the experiment. There were two phases in the experiment: exposure and test. In the exposure phase, participants were presented with the two interleaved languages for a total of 24 minutes, divided into three blocks of 8 minutes. Within each block, the two languages alternated between each other. Each language was presented in 2-minute strings and was presented twice per block. Participants were given a one-minute break between blocks. The orders of accent (hearing an English-accented language first or second) and presentation of the language (AL1 presented first or second) were counterbalanced across participants. Thus, half of the participants heard AL1 in an English accent and AL2 in a French accent and the other half

heard the opposite. In the test phase, participants completed 32 two-alternative forced-choice trials (i.e., 16 items for each language). Each test trial consisted of two trisyllabic words: (i) correct words - syllables with a transitional probability of 1.0 in a language and (ii) part words - syllables with the low transitional probability of 0.33 in a language. Participants were required to press a key to indicate which of the two heard words sound more like a word from the languages they heard during exposure. Test items from each language were presented in the same accent as in the exposure phase. For example, if Language 1 were presented in an English accent, then test items from Language 1 would also be presented in English accent.

Color-Shape Switching task. This task was adapted from the switching task conducted by Prior and her colleagues (Prior & MacWhinney, 2010) and was programmed by using E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Participants were required to categorize visual stimuli based on its colour or shape, using corresponding keys on the keyboard (clear instruction and labels were provided). The visual stimuli were bivalent stimuli that were a circle (3 cm radius) and a triangle (3 cm base, 2.5 cm height) in either red or green. The participants received cues (5 cm X 1.2 cm) before the target presentation to determine whether it was a colour or shape trial. If they saw a rainbow gradient, they had to categorize the stimulus by its colour. If they saw a shape gradient, they had to categorize the stimulus by its shape. During the experiment, participants were instructed to respond by using one hand for each task. For example, they used their left hands to press target keys (“q” and “w”) when it was a colour task, whereas they used their right hands to press target keys (“o” and “p”) when it was a shape task. The labels for the tasks (i.e., colour and shape stickers) were provided on the keyboard. The hands used for each task were counterbalanced across participants.

Each trial began with a central fixation cross staying on the screen for 350 ms followed by a blank screen for 150 ms. Then a task cue (rainbow or shape gradient) appeared 3.5 cm above the fixation cross and remained on screen. After 250ms, the target stimuli appeared in the center of the screen. The target remained on the screen until the participants responded, or until 4 seconds passed. A 100 ms beep sound was presented if participants made an incorrect response. A blank screen serving as an inter-trial interval was presented before the next trial.

The experiment adapted a block-sandwich design. First, participants completed two single-task blocks that contained only color or shape cues (each block provided 8 practice trials and 36 test trials). Next, participants received 16 practice trials of a mixed-task block in which colour and shape cues were randomly mixed within the block. After completing the practice trials, participants proceeded to three mixed-task blocks that each included 48 test trials and 2 dummy trials. The dummy trials were presented at the beginning of each block, they were employed to ensure participants are accustomed to the procedure and therefore were not included in the analysis. The mixed-task blocks produced two types of trials for final analysis: repeated trials and switch trials. A repeated trial contained the same colour/shape cues as the preceding trial, whereas a switch trial contained different colour/shape cues from the preceding trial. Lastly, participants completed two additional single-task blocks after completing three mixed-task blocks. In total, there were 144 trials single-task trials (half colour-cued, half shape-cued), 72 repeated trials, and 72 switch trials. The color-shape task generated two variables of interest: switching cost and mixing cost. Switching cost was the difference in average response time (RT) between repeated and switch trials. This captures the additional RT participants needed to switch their response in accordance with the new cues in the current trial during the mixed task blocks, thus measuring how participants flexibly switch between tasks. On the other hand, mixing cost is a measure of monitoring and

maintaining two task sets in mind. The calculation of mixing cost involves the RT of the single-task trials and repeated trials. The reasons are as follows. Participants only needed to respond to one task cue during the single task blocks. Average RT during the single task blocks thus measures participants' baseline RT. However, participants needed to maintain two task cues to prepare for possible changes during the mixed task blocks. Average RT of the repeated trials during mixed task blocks is a measure of the additional RT participants needed to prepare for a possible switch between task cues. Importantly, RT of the repeated trials did not take into account the additional RT of actual switches in task cues (i.e., switch trials). Thus, mixing cost was the difference in average RT between repeated and single-task trials. Both the switching cost and the mixing cost are the additional RT participants needed to complete a task. Thus, smaller values in these respective measures mean that participants are better at switching and monitoring task sets.

Attention Network task. The Attention Network Task (ANT), based on the methodology in Fan et al., (2002), was programmed using Presentation® software (Version 14.9, Neurobehavioral Systems, Inc., Berkeley, CA). It measures three different attentional network processes: alerting (attain and maintain an alert state to prepare for incoming sensory inputs), orienting (select information from sensory inputs and turn attention towards to sensory signals), and executive control (monitor and resolve conflicts in the presence of competing information) (Posner & Petersen, 1990). The ANT task is a hybrid task that contains both a flanker task and a cue reaction task. For the flanker task components, participants were instructed to respond to the direction of the central arrow presented on the monitor. They were required to press a left response button of a Cedrus RB-830 button box when the central arrow pointed to the left and press a right response button when the central arrow pointed to the right. There were three different conditions: neutral, congruent and incongruent. In the neutral condition, a central arrow was pointing to either

the left or right, with two dashes on each side of the arrow. These trials provided an average baseline reaction time. In the congruent condition, the central arrow - either pointing to the right or left - was presented with four side arrow flankers pointing to the same direction. By contrast, in the incongruent condition, the central arrow was presented with four side arrow flankers pointing in the opposite direction. The average response time (RT) difference between incongruent condition and congruent condition yielded the measure of executive control in ANT task. A lower value in the measure of executive control means that participants could inhibit irrelevant responses in the incongruent condition better. All attentional cues were carried by the presence of asterisk cues before the target appears. The exact procedure would be described in detail below.

In each trial, a fixation cross appeared at the centre of the screen and remained in the same position throughout the whole trial. The time between the mere presence of fixation cross and the appearance of the asterisk cues varied from 250 to 1500 ms. The variation was used to prevent participants from forming expectations of when the asterisk cues appear in each trial, thus minimizing the learning effect merely on the basis of expectation and adaption to the task procedure. An asterisk cue (cue conditions listed below) was then presented for 100 ms and disappeared, with only the fixation cross remaining on the screen for an additional 250 ms. The target (i.e., one of the three flanker conditions) then appeared and stayed on the screen until either the participant responded or 2000 ms passed. See Figure 2 for a detailed description of the procedure of the ANT task.

Following the standard ANT paradigm (Fan et al., 2002), four additional different cue conditions were created. In the “No Cues” condition, participants saw only the fixation cross. In the “Double Cues” condition, participants saw two asterisks above and below the fixation cross. This condition alerted participants that there would be new information (target) without providing

any information about the target's location. Subtracting the average RT of "Double Cues" condition from that in the "No Cues" condition gave an alerting effect measure. In the "Spatial Cues" condition, the asterisk cue signaled the location of the incoming target by appearing at the same location as the target. In the "Central Cues" condition, the asterisk cue appeared at the location of the fixation cross. This condition served to control for the additional alerting effect when an asterisk appeared onscreen in the "Spatial Cue" condition. However, it was expected that the alerting effect in the "Double Cues" condition was larger than that in the "Central Cue" condition. The orienting effect in ANT task was the RT difference between the "Spatial Cue" and "Central Cue" conditions. Higher values of the alerting and orienting effects mean that participants are better at the corresponding attentional components.

The ANT task consisted of 288 trials: three testing blocks 96 trials. All trials were randomly ordered. Flanker conditions were randomly assigned across testing blocks, one third of trials were neutral, one-third congruent, and one-third incongruent. Each target was also randomly paired with one of the four cue conditions and an equal number of cue conditions were assigned to each block.

Working memory tasks. The working memory tasks, presented using E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA), were based on the shortened complex working memory tasks in Oswald, McAbee, Redick and Hambrick (2015; see <http://englelab.gatech.edu/tasks.html#English>). Prior studies (Foster et al., 2015; Oswald et al., 2015) have shown that these shortened span tasks have high test reliability and good predictive validity in comparison to the standard long span tasks, thus we decided to test participants with a shortened version of span tasks in order to reduce the length of the experiment. Participants' responses in the working memory span tasks were recorded via a computer mouse. In each span task, practice trials were provided in three blocks: a simple distractor task (i.e.,

math/symmetry/sentence problem), a simple main task (i.e., recall letters/position of squares), and a combination task (i.e., combining both distractor and main tasks in one trial). There was a time limit (i.e., average response time + 2.5 *SD*) for responding to the distractor task in all working memory span tasks. This time limit prevented participants from rehearsing information in the main task when solving distractor problems. If participants' response time was longer than the time limit, the program automatically moved on to the next trial and counted that distractor trial wrong. Participants' score was the total number of correctly recalled items in a correct presented order.

Operational Span task. In this task, participants were required to solve simple mathematical problems (distractor task) while remembering a set of unrelated letters (main task). Participants were presented with a series of "math-letter" pairs. On each trial, participants were first presented with a mathematical problem (e.g., Is $(6*1) + 2 = 8$?) and they later responded by clicking one of the boxes "True" or "False" on the screen. A letter would appear for 1000ms after participants answered the mathematical problem. When the letter disappeared, another "math-letter" pair would be presented. At the end of the trial, participants were shown a 4 X 3 letter matrix and required to select the letters in the order that they were previously presented. The number of "math-letter" pairs randomly varied from 4 to 6 on each trial and the total number of "math-letter" pairs was 30.

Symmetry Span task. In this task, participants were required to judge whether black and white square matrix images were symmetrical (distractor task) while remembering the positions of red squares in a 4 X 4 matrix (main task). Similar to the operation span task, distractor task was also presented first on each trial, followed by a main task. Participants responded to the distractor task by clicking one of the boxes "True" or "False" when asked whether the preceding image was symmetrical. A red square would appear in one of the cells in a 4 X 4 matrix for 650ms after

participants answered the symmetry question. At the end of each trial, participants were shown a 4 X 4 square matrix and required to indicate the positions of red square in the sequence order that they were previously presented. The number of “symmetry-red square” pairs randomly varied from 3 to 5 on each trial and the total number of pairs was 24.

Reading span task. In this task, participants were required to read English sentences and judge whether the sentences made sense (distractor task) while remembering a set of unrelated letters (main task). On each trial, participants were first presented with a sentence (e.g., Sandy likes swimming by the wall.) and they were asked whether the sentence made sense. Participants responded by clicking one of the boxes “True” or “False” on the screen. A letter would then appear for 1000ms. At the end of the trial, participants were shown a 4 X 3 letter matrix and required to select the letters in the order that they were previously presented. The number of “sentence-letter” pairs randomly varied from 4 to 6 on each trial and the total number of pairs was 30.

Results

Data screening. When calculating average response time, individual data from all executive functioning tasks were screened to exclude inaccurate trials. This ensured that participants were not sacrificing accuracy for better performance in the executive functioning tasks. We removed participants’ scores when their accuracy was below 80% (2 in the Color-Shape Switching task, 2 in the ANT task, 6 in the Operational span task, 10 in the Symmetry span task and 2 in the Reading span task). However, it is important to note that the above procedure did not remove these participants from all analyses. It simply means that participants’ score in that particular executive functioning task was not entered in the regression model as a predictor.

For the Color-Shape Switching task and ANT task, trials with response times shorter or longer than 2.5 standard deviations were classified as outliers and subsequently removed. On

average, only a small portion of RT were classified as outliers (approximately 2.85% for Color-Shape Switching task and 4.98% for ANT task).

Further, all independent variables were screened and winsorized for univariate outliers using the criteria of $Z = \pm 2.5$. Lastly, assumptions about multicollinearity, normality, linearity and homoscedasticity of residuals for multiple regressions were checked and met before running the regression model below.

Statistical segmentation task. The main dependent variable was the word segmentation accuracy, measured by the proportion of selecting words over part words. The average proportion for Artificial Language 1 (AL1) was 0.58 ($SD = 0.17$) and the average proportion for Artificial Language 2 (AL2) was 0.50 ($SD = 0.15$). Accuracy in AL1 was significantly above the chance level of 0.5 [$t(114) = 5.13, p < 0.001$, two-tailed] but accuracy in AL2 did not significantly differ from chance level [$t(114) = -0.16, p = 0.877$, two tailed].

To investigate which factors in the experiment (e.g., accent of the languages, presentation order of languages in the exposure phase) contribute to segmentation accuracy in AL1 and AL2, we entered the data into two separate mixed-factors ANOVA models. We also explored whether participants performed differently across different language groups, entering language background as a factor in the ANOVAs. Lastly, we entered age as covariate in the ANOVA analyses because there was a wide range of ages in our final sample ($M = 20.29, SD = 3.25, \text{range} = 17 - 32$).

To examine what factors predicted AL1 segmentation accuracy, we performed a 4 (Language background: monolinguals, simultaneous bilinguals, sequential bilinguals⁴,

⁴ We did not differentiate between early bilinguals and late bilinguals within the sequential French-English bilingual group. This is because the correlational analyses have suggested that there are no significant relationships between AoA and participants' performance in any statistical segmentation and executive functioning tasks [$rs < 0.3, ps > 0.09$].

multilinguals) X 2 (Gender: female, male) X 2 (Language accent: AL1 in English or French) X 2 (Language Presentation Order: AL1 presented first or second) mixed-factors ANOVA, after controlling for participants' age. The main effect of language accent was significant [$F(1, 107) = 5.96, p = 0.016$]. This result suggested that participants performed significantly better when AL1 was in a French accent ($M = 0.60$) than in an English accent ($M = 0.53$). Nevertheless, AL1 segmentation accuracy was significantly above chance level across both accent conditions [$ps < 0.03$, two-tailed]. Further, the main effect of language background was significant [$F(3, 107) = 4.58, p = 0.005$]. A finding that will be explored in the regression models below. No other main effects were significant [$ps > 0.8$]. We ran the model again to investigate whether the above-mentioned factors predicted AL2 segmentation accuracy. The only significant predictor was accent of the languages [$F(1, 107) = 25.48, p < 0.001$]. Our result indicated that participants segmented AL2 successfully when AL2 was in an English accent [$M = 0.56, p = 0.001$, two tailed] but they failed to do so when AL2 was in a French accent [$M = 0.43, p < 0.001$, two tailed]. Note that the average segmentation accuracy was below 0.5 (chance level) when AL2 was in a French accent. This result suggests that participants selected part words, but not words, as correct words at the test. However, this task explicitly asked adults to select words that sounded like words to them. Thus, participants may have differentiated between part words and words but they have fundamentally failed the task because they indicated that part-words (statistically incoherent item) were words. All other main effects were not significant [$ps > 0.35$].

Here, it is important to note that when AL1 was in English accent, AL2 would be in French accent. Thus, the above analyses suggested that participants were able to segment both AL1 and AL2 when AL1 was in French accent and AL2 was in English accent. However, they could only

segment AL1 but not AL2 when AL1 was in English accent and AL2 was in French accent. See Figure 3 for details of the average proportion of word accuracy in each language across conditions.

Contributions of linguistic and cognitive factors in statistical segmentation. We ran two separate multiple regression models to explore how linguistic and cognitive factors predicted AL1 and AL2 segmentation accuracy. In this paper, we did not enter the average language usage and language proficiency (both self-rated and objective) in the regression models. This is because these two factors were matched among the simultaneous bilingual, sequential bilingual and multilingual language groups (refer to Table 1 for details). As indicated in the previous ANOVA model, participants' language backgrounds predicted their segmentation accuracy in AL1. Thus, we investigated which group performed best in the regression models. To do so, the monolingual group served as the baseline and we created three dummy indicator variables to represent simultaneous bilinguals, sequential bilinguals and multilinguals in the models. Entering these three dummy variables also allowed us to see whether bilinguals' age of acquisition (AoA) would be a predictor of participants' segmentation accuracy. Regarding the cognitive factors, we entered the following factors in the models: alerting effect, orienting effect, executive control, switching cost, mixing cost, and composite working memory scores. Note that we used a composite working memory score because operation, symmetry and reading span tasks were all positively and significantly correlated ($r_s > 0.30$, $p_s < 0.004$) and high correlations across independent variables may lead to a multicollinearity problem. As such, we created a composite working memory score by averaging Z scores across operation, symmetry and reading span tasks. Lastly, we entered age and accent conditions in the regression model as covariates since the previous ANOVA models suggested that they were significantly related to the segmentation accuracy in AL1 and AL2. All independent variables and covariates were entered simultaneously in the regression models.

In the first regression model (model 1), AL1 segmentation accuracy was the dependent variable. All independent variables and covariates as a set significantly predicted the performance in AL1 [$F(11, 99) = 2.85, p = 0.003, R^2 = 0.49$]. Similar to the ANOVA analysis, accent condition was significant predictor for AL1 performance [$\beta = 0.22, p = 0.018$]. We found that simultaneous bilinguals significantly segmented more words accurately in AL1 when compared to monolinguals [$\beta = 0.32, p = 0.005$]. However, sequential bilinguals and multilinguals did not significantly outperform monolinguals [$\beta s < 0.17, ps > 0.13$]. We also found that mixing cost positively predicted the performance in the AL1 [$\beta = 0.21, p = 0.032$]. All other cognitive factors were not significant predictors of AL1 accuracy [$ps > 0.10$]. The second regression model (model 2) used AL2 segmentation accuracy as dependent variable. Again, all independent variables and covariates as a set significantly predicted performance in AL2 [$F(11, 99) = 2.64, p = 0.005, R^2 = 0.48$]. However, the only significant predictor for AL2 was the accent condition [$\beta = -0.44, p < 0.001$]. Language group and all other cognitive factors did not significantly predict AL2 accuracy [$ps > 0.30$]. See Table 3 for the results of the regression models.

Because we found simultaneous bilingual advantage in AL1 as well as a positive relationship between mixing cost and AL1 performance, we conducted a follow-up mediation analysis with the bootstrapping method (Preacher & Hayes, 2008) to examine whether mixing cost was a mediator of the relation between simultaneous bilinguals and AL1 performance. Our results showed that mixing cost was not a significant mediator in the model [$B = -0.01, SE = 0.009, 95\% CI (-0.0361, 0.0001)$].

Discussion

Overall, participants successfully segmented AL1 but not AL2. Importantly, we discovered that participants' language backgrounds predicted AL1 segmentation accuracy. Our multiple

regression analyses indicated that simultaneous bilinguals significantly outperformed monolinguals in segmenting words, whereas sequential bilinguals and multilinguals did not significantly differ from monolinguals with respect to their segmentation accuracy.

Although cognitive abilities generally did not predict participants' performance in the segmentation task, we found that mixing cost positively predicted segmentation accuracy. Higher mixing cost implied that learners needed more time to monitor task cues and maintain two competing tasks in mind. Thus, our results suggested that poorer ability of monitoring and maintaining two tasks was related to better segmentation accuracy in AL1. This result is surprising because we expected that skills related to monitoring two task sets should positively predict better segmentation performance. This finding may be due to an overlearning effect when participants focused on learning one particular language at the expenses of paying attention to two different task sets. We will return to this point in the general discussion section. Taken together, the above results reveal that participants' performance in the segmentation task cannot primarily be explained by their general cognitive abilities. But we found that simultaneous bilinguals, but not sequential bilinguals or multilinguals, segmented more words when compared to monolinguals. Thus, our findings support the argument that early bilingual experience (AoA), but not general cognitive abilities, facilitates word segmentation in a dual-input statistical task.

Finally, our results suggest that success in segmentation depended on the accent condition of the languages. If AL1 was in a French accent and AL2 was in an English accent, participants were able to segment both languages. However, if AL1 was in an English accent and AL2 was in a French accent, participants successfully segmented AL1 but not AL2. This result was not expected since we adopted the script (with only a change in one vowel) from Weiss et al (2009) and the participants in that study were successful in segmenting both AL1 and AL2. One possible

account for the current results is that learners may have preferences for the part word test items in the AF2 when it was presented in French accent. Our results have suggested that participants generally selected part-words as words candidates in the artificial language when they were tested in AL2 with a French accent. So it is possible that participants failed to segment words from AL2 because they had a preference to part-word items of this language when it was in French accent. To investigate whether this is the reason of participants' failure of segmenting AL2 in Experiment 1, we designed the second experiment.

Experiment 2A

The purpose of this experiment was to examine whether learners switch their preference to part words over words in AL2 when they heard the language presented in French accent. Learners may have a preference for part words if certain phonemes in the part words in AL2 sounded more natural or familiar to them in the French accent. For example, several part-words may have aligned better with learners' phonological knowledge of some real French words, thus participants may think that part-words produced in French accent are words in AL2. If learners have item-based preference for part words over words in this artificial language, they will demonstrate biases towards part words even if they are not previously exposed to AL2. To test this hypothesis, we presented another group of participants the test items from AL2 without first exposing them to the artificial language. In this experiment, participants heard two-word items (one part word and one word from AL2) on each trial and were asked to choose which item sounded more like a word to them. We expected that participants would, on average, choose part words over words in AL2 if they have initial item-based preference for the former.

Method

Participants. Thirty-six adults (18 female, Mean age = 23.19 (*S.D.* = 4.68)) participated in this study: half of listened to AL2 in the French accent and the other half listened to AL2 in the English accent. Participants again completed the language background questionnaire (LBQ) as well as French and English cloze tests. In this experiment, we did not control for the participants' language backgrounds. This is because the ANOVA and multiple regression models in Experiment 1 have suggested that participants' language background is not a factor that explains the variances in the segmentation accuracy of AL2. As such, we argue that the participants' language background would not explain variances related to participants' preferences for test items in AL2 in Experiment 2A

Based on LBQ and cloze tests, 7 were functional English monolingual adults, 3 were simultaneous French-English bilingual adults who learned both languages before age 1, 17 were sequential French-English bilingual adults who learned their second language (either French or English) between age 1 to age 17, and 9 were multilingual adults who indicated that they were proficient in at least three languages (may include French and English).

Experimental stimuli and procedure. The stimuli used in this experiment were identical to the AL2 test items in Experiment 1. Participants were randomly assigned to the French accent condition or English accent condition. They were informed that they would hear a total of 16 trials in the experiment and they would hear two-word items on each trial. They were instructed to indicate their preference for which item sounded like a word to them based on their intuition. Participants were also told that their answers only reflected their preference, and not accuracy or performance in the experiment. On each trial, participants heard one part word and one word. The

presentation of part words and words were counterbalanced across 16 test trials. At the end of the experiment, participants filled out LBQ and cloze tests.

Results and Discussion

The dependent variable was the proportion of selecting words over part words in AL2. Two one-sample *t* tests in each accent condition were computed to check if participants had a preference for words over part words. We did not find evidence of item-based preferences for the French accent [$t(17) = 0.70, p = 0.492$] or the English accent [$t(17) = -1.12, p = 0.278$]. Independent *t* tests also revealed that word/part-word preference was not different across the two accent conditions (i.e., French vs versus English) [$t(34) = 1.24, p = 0.222$].

This result indicates that participants did not have any initial biases towards part words or words in AL2, suggesting that the failure of segmenting AL2 in Experiment 1 was not driven by an item-preference for part words over words. Another alternative explanation for participants' failure of AL2 segmentation in Experiment 1 is that accent may alter AL2 sound properties (e.g., prosody) and this alternation influences segmentation performance. If this is the case, we should expect learners to succeed in learning AL2 only when it was presented in a specific accent. For example, participants may succeed in segmenting AL2 only in an English accent but not when AL2 is in a French accent.

Experiment 2B

In this experiment, we tested whether learning differences were present when participants were tested with the AL2 in different accents. Participants were therefore presented AL2 alone (i.e., a single input statistical segmentation task) in either an English or a French accent.

Method

Participants. Thirty-six adults (20 female, Mean age = 19.19 (S.D. = 1.19)) participated in this study, half listened to AL2 in the French accent and the other half listened to the English accent. Again, participants completed the LBQ and cloze tests in French and English. We did not control for the participants' language backgrounds in this experiment as the results in Experiment 1 indicated that learners' performance did not differ in terms of participants' language backgrounds. Thus, we expect that participants with different language backgrounds would perform similarly in a single-input word segmentation task using AL2 as stimuli. Based on LBQ and cloze tests, 11 were functional monolingual-English adults, 3 were simultaneous French-English bilingual adults who learned both languages before age 1, 8 were sequential French-English bilingual adults who learned their second language (either French or English) between age 1 to age 9, and 14 were multilingual adults who indicated that they were proficient in at least three languages (may include French and English).

Experimental stimuli and procedure. Only the AL2 stimuli from Experiment 1 were used. Participants were randomly assigned to the French accent condition or English accent condition. They were instructed to listen and pay attention to a recording of a set of syllables. They were informed that they would subsequently be tested on what they had heard but did not receive any further information regarding the experiment. Similar to Experiment 1, there were two phases in the experiment: an exposure phase and a test phase. In the exposure phase, audio stimuli were presented in three 4-minute blocks. Participants were given a 30 second break between blocks. Transitional probability within words was always 1.0 whereas the transitional probability across word boundaries was always 0.33. In the test phase, participants were asked to complete a two-alternative forced-choice test in which each trial consisted of a word (transitional probability of

1.0) and a part word (transitional probability of 0.33) from AL2. There were 16 test trials in total. At the end of experiment, participants filled out the LBQ and cloze tests.

Results and Discussion

The dependent variable was the proportion of selecting words over part words (i.e., word segmentation accuracy). The average proportion for AL2 in French accent condition was 0.44 (SD = 0.16) and this proportion was not significantly higher than chance [$t(17) = -1.58, p = 0.132$, two-tailed]. In contrast, the average proportion for AL2 in English accent condition was 0.55 (SD = 0.09), which was significantly above chance [$t(17) = 2.56, p = 0.020$, two-tailed]. An independent t test revealed that the segmentation accuracy of AL2 in the French condition significantly differed from that in the English condition [$t(34) = 2.62, p = 0.013$, two-tailed]. It is important to note the difference between Experiment 1 and Experiment 2B. The sample in Experiment 1 was bigger ($n = 60$) when we examined whether participants learned AL2 when it was presented in the French accent. In contrast, the sample size in Experiment 2B was smaller ($n = 18$) and we discovered that participants failed to differentiate between part-words and words. This finding thus suggests that the learning effect size of AL2 in French accent may be small such that this effect could only be detected with a bigger sample size.

The results from Experiment 2B demonstrate that participants can successfully segment AL2 when it is presented in an English accent. This finding partially replicated Weiss et al (2009) because they also tested learners on AL2 in an American-English accent – an accent that is similar to Canadian English. Together, both studies have suggested that AL2 is only learnable when it is presented in an English accent. This suggests that French-accented speech may have altered the sound properties in AL2, leading to perceptual challenges that affect participants' segmentation

performance. We see two potential accounts: a conflict between prosody and statistical cues, or phonetic realizations causing perceptual confusion.

It is possible that French-accented speech may change the prosody of the AL2 stimuli. Thus, the failure to segment is a result of the conflict between AL2 statistical cues (transitional probability) and prosodic cues. Prior studies have shown that prosody plays an important role in segmenting words from fluent speech. Shukla, Nespors and Mehler (2007) found that adult learners only successfully segmented words when the prosodic cues were consistent with the transitional probability across word boundaries. In their experiments, Shukla and colleagues discovered that adult learners recognized few words that did not align with regular prosodic cues (e.g., lengthened middle syllable instead of lengthened final syllable). More importantly, the researchers found that learners failed to segment words when the prosodic cues were pitted against the transitional probability. For instance, participants did not learn groups of syllables with high transitional probability as words when the prosodic cues across these syllables were poor (i.e., prosodic cue of syllable lengthening located in the middle of the word and not at the edge of the word). Thus, one potential account of why AL2 in French accent is more difficult to learn is that prosodic cues, such as syllable duration, are inconsistent with the transitional probability in the language.

We investigated this possibility by comparing syllable prosodic properties between words and part words of AL2 in the French accent condition. Here, we focused on the first and final syllables of these words because prior studies suggested that prosodic cues often exist at the edges of the words (Shukla, Nespors & Mehler, 2007; Endress & Mehler, 2009). Using the PRAAT computer program (Boersma & Weenink, 2012), we measured the syllable duration, average fundamental frequency (f_0), pitch peak, average intensity and maximum intensity. If prosodic cues are inconsistent with the transitional probability of the words, we expect to find stronger prosodic

cues (e.g., lengthened syllable, higher f_0 and intensity) in part-word syllables than in word syllables. We ran several independent t-tests to examine whether the above-mentioned prosodic properties differed between part words and words when considering the first syllable of the words, but none of the differences were significant [$t_s < 1.2$, $p_s > 0.3$]. We found similar results when we did the same comparisons for the final syllable of the words [$t_s < 1.5$, $p_s > 0.25$]. As such, prosodic changes in French-accented speech are not in conflict with the statistical cues in AL2, suggesting that the alternation of prosodic cues in French accent is not the main reason why participants failed to segment AL2.

Another possibility is that French-accented speech may alter the phonetic realizations of certain phonemes in the AL2 stimuli. This may, in turn, lead to perceptual confusion of several important phonemes in the language and eventually impede participants' segmentation. One potential confusion may be related to the voice onset time (VOT) differences between the phonemes /b/ and /p/ in the AL2 script. The /b/-/p/ distinction in French is more subtle than in English because the VOT difference is shorter in French and the /p/ has no aspiration, unlike English (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; see Burns, Yoshida, Hill & Werker, 2007, for a review). Consistent with the literature, we found that the average VOT difference between /b/ and /p/ in the French-accent condition is smaller (Mean = 25.32ms) than that in the English accent condition (Mean = 39.57ms). Given that our bilinguals and multilinguals were on average English-dominant (see self-rating and cloze task results) and the monolinguals were completely English-dominant by definition, our participants may find it difficult to perceive the smaller VOT difference between /b/ and /p/ in French-accented speech than is typically found in English. In our experiment, we note the phonemes /b/ and /p/ only exist in AL2 but not AL1. If participants have difficulty perceptually classifying certain phonemes (i.e., /b/ versus /p/), they

will incorrectly represent certain syllables and therefore conflate transitional probabilities across different words, and consequently fail to segment correctly.

General Discussion

The present study investigated whether a stable distinction within a single person's speech, phonetic realizations associated with particular languages, can be an effective contextual cue for segmentation of dual language inputs. Additionally, we investigated whether bilingualism facilitates segmentation of two statistical language inputs, with a focus on exploring factors that explain why bilinguals may outperform monolinguals in this task.

For the first question, we employed English- and French-accented speech to mark the differences across two different language inputs in a statistical segmentation task (Experiment 1). Overall, we found that participants successfully segmented AL1 but not AL2. Follow-up analyses revealed that, while participants successfully segmented AL1 across both accent conditions, they could only segment AL2 in an English accent and not in a French accent. These results demonstrate that accented speech can only facilitate successful segmentation of two overlapping language inputs under specific conditions. However, we would like to stress that participants' successful segmentation of AL1 across conditions demonstrates the power of using stable phonetic cues to indicate language streams. Despite not succeeding with AL2 in one condition, participants still correctly segmented AL1 in the face of interference from the interleaved AL2. This demonstrates that AL2's transitional probability cues were not conflated with those in AL1, which means that participants were successfully separating the languages based on accent cues.

In Experiment 2A, we sought to determine whether the failure of segmenting AL2 in Experiment 1 was driven by participants' item-based preference for part words over words. Across two experiments, participants showed no preference for part words in AL2, discounting this

possibility. In Experiment 2B, we simplified the task to a single language condition to confirm if that participants' success in segmenting AL2 strongly depended on the accent conditions. Again, participants were only able to successfully segment AL2 when it was presented in English accent and failed with the French-accented AL2.

So what makes AL2 difficult to segment? We can rule out multiple possibilities. First, participants' poor performance is unlikely to be driven by atypical or rare syllabic combinations in AL2 words. If syllabic combinations present in AL2 words are phonotactically uncommon in participants' native/dominant languages, we should see item preferences for part words when the AL2 was presented in an English accent, as most of our participants' dominant language was English. Yet, our participants did not show such item-preference. Further, they also successfully segmented words when AL2 was in English accent. An alternative account is that syllabic combinations present in AL2 words are uncommon in French. We also think that this account is improbable. If the syllabic combination in AL2 is rare in French, we should see bilinguals and multilinguals, who are more familiar with the French accent and phonetic inventory, perform poorer than their monolingual counterparts. However, the language backgrounds of our participants did not explain variances in the item-based preferences and segmentation performance of AL2, thus discounting this possibility.

It is our supposition that French-accented speech altered the phonetic properties of several important phonemes in AL2, which led to learners' perceptual confusion. Indeed, we found evidence via acoustic analyses of the stimuli that the VOT difference between /b/ and /p/ is much smaller in the French-accented speech than that in English-accented speech. If participants had difficulty differentiating /b/ and /p/ in the French-accented speech, they were likely to incorrectly combine transitional probabilities across different words in the AL2. Thus, we suggest that

participants' failure to segment AL2 may be driven by difficulties in perceiving the subtle phonetic distinction between /b/ and /p/ in the French accent condition.

Taken together, our findings suggest that, unsurprisingly, acoustically closer phonetic distinctions in one language may pose challenges to learners when classifying key phonemes in an artificial language, leading to difficulties in accurately tracking the transitional probabilities of words in said language. However, in general, accented speech within a single individual's speech can still be treated as effective cues for the statistical segmentation of words in two overlapping languages. Learners succeeded in learning both languages in one condition and one language in the other, which reveals that segmentation of one language did not influence their segmentation of the other language in either case. Thus, accented speech allows learners to separately represent two conflicting linguistic statistical structures.

Regarding the question about the influence of bilingualism on the dual-input statistical segmentation task, we first divided participants into four language groups: functional English monolinguals, simultaneous French-English bilinguals, sequential French-English bilinguals, and multilinguals. Our results suggested that only simultaneous bilinguals outperformed monolinguals when segmenting words in AL1. We also tested participants with a battery of cognitive tasks to examine which cognitive ability may account for a bilingual advantage in the segmentation task. Our findings provide insights of the underlying mechanism of why some bilinguals perform better in our statistical learning task. First, we found that simultaneous bilinguals' cognitive skills were not related to their segmentation performance, ruling out the possibility that bilingual advantage in statistical learning is due to more efficient cognitive processing such as attention, inhibitory control, set switching and working memory. This result is perhaps not very surprising, given that a large number of studies have not found bilingual cognitive advantages (e.g., Gathercole et al.,

2014; Kousaie & Philips, 2012; Kousaie et al., 2014; Paap & Greenberg, 2013; Paap & Sawi, 2014; see Hilchey, Saint-Aubin, Klein, 2015 for a review). Crucially, our results points to the importance of early bilingual experience (i.e., age of acquisition) in statistical segmentation, as only the simultaneous bilinguals outperformed monolinguals. Stronger language proficiency and more frequent use of two languages cannot explain the observed bilingual advantage since our sequential bilinguals and multilinguals were on par with simultaneous bilinguals in terms of the above two aspects. So, how does early bilingual experience, such as learning two languages since birth, supports better statistical segmentation in an environment with two language inputs?

Simultaneous bilinguals have learned two different languages since birth, and therefore must be capable of differentiating two parallel language inputs from the beginning of language development in order to acquire their native languages successfully. However, sequential bilinguals did not face the same challenge since they established a rudimentary lexicon and phonetic inventory of their first language before acquiring the other language(s). It is important to note that the multilingual group who performed similarly to monolinguals were predominantly sequential bilinguals, with only 4 of 16 reporting simultaneous acquisition. Here, we argue that early bilingual experience may hone simultaneous bilinguals' skills in identifying multiple parallel inputs based on subtle contextual cues (i.e., perceptual cues, such as accented speech, rhythm of the languages) in their environment. Further, several recent studies have suggested that simultaneous bilingual infants are more open to processing non-native sounds and words (e.g., Graf Estes & Hay, 2015; Petitto et al., 2012; Singh, 2017). Therefore, it could be argued that simultaneous bilinguals may be more open-minded and flexible in interpreting linguistic information under novel circumstances. An essential key to learning two parallel language inputs during infancy is to assume possible changes in the language environment. Accordingly, we argue

that simultaneous bilinguals may be more open to expecting possible variability in their language environment. This assumption plays an important role in helping learners to be less entrenched to one linguistic input, which in turn facilitates simultaneous bilinguals' detection of two different statistical structures in a dual-input statistical learning task (Basseville & Nikiforov, 1993; see Gebhart, Aslin & Newport, 2009; Qian, Jaeger, & Aslin, 2012 for reviews). Hence, early language experience may influence bilinguals' expectation of variability in their environments and help them to be flexible when interpreting new language inputs.

Another contribution of our paper is that mixing cost, a measure of learners' ability to maintain and monitor two different task sets, positively associates with segmentation accuracy in AL1. Mixing cost is the additional reaction time that participants need to monitor two task sets in the mixed block in comparison to monitoring only one task set in the pure block. Our result therefore shows that poorer ability of monitoring two task sets is linked to better segmentation accuracy in AL1. This result, on the surface, is surprising since we expect that dual-task monitoring skills should positively relate to better performance in segmentation. Yet, we suggest that this finding may actually reflect that overlearning of one particular language can reduce attention to two different task sets. Karuza et al (2016) employed fMRI techniques to discover that learners who were most efficient in learning the first presented language in a dual-input segmentation task had reduced activation in posterior parietal cortex, a brain region that is associated with monitoring changes in a dynamic environment. In line with this study, other behavioral studies have also shown that when learners were exposed to one particular language longer, they tended to acquire this language at the expense of learning the second language (Bulgarelli & Weiss, 2016; Zinszer & Weiss, 2013). This may be because when participants are exposed to one language longer, they tend to focus on the AL1 statistical structure only when they are establishing a solid representation

of the AL1, resulting in reduced attention to monitoring other tasks or other language statistical cues in their environment. Thus, our finding is consistent with the current literature about over-learning of one language can divert learner's attention from monitoring other task sets/cues in ambient language environment.

However, it is important to note that we generally did not find a significant relationship between the ability of segmenting languages and executive function processing. This finding is unexpected because prior studies have suggested that attention is positively related to some forms of statistical learning, such as speech segmentation (Toro, Sinnett, & Soto-Faraco, 2005) and visual statistical learning (Baker, Olson, & Behrmann, 2004). In addition, other studies also found that better inhibitory control predicts better speech segmentation accuracy (Weiss, Gerken, & Mitchel, 2010). Thus, this raises a question of why previous studies found a positive relationship between segmentation ability and some cognitive processes whereas we failed to find such relationship in our paper. To answer this question, we offer two potential reasons: a non-unitary statistical learning mechanism and the construct validity issues of cognitive tasks.

First, several recent studies reported that the correlations between different statistical learning tasks, including tasks that employed stimuli within the same modality (e.g., auditory/visual), are generally low and often non-significant (Erickson, Kaschak, Thiessen & Berry, 2016; Misyak & Christiansen, 2012; Siegelman & Frost, 2015). These findings have suggested that different statistical learning tasks most likely tap into different underlying components rather than a unitary learning mechanism (see Siegelman, Bogaerts, Christiansen & Frost, 2017 for a review). It is thus possible that the dual-input statistical segmentation task in our paper is not measuring the same learning mechanism that prior single-input statistical segmentation tasks were measuring (e.g., Toro, Sinnett, & Soto-Faraco, 2005; Weiss, Gerken, &

Mitchel, 2010). Perhaps executive function ability correlates to the cognitive mechanism that underlies the learning mechanism of a single-input statistical segmentation task, but not to the learning mechanism of a dual-input statistical segmentation task. Therefore, we may be measuring a specific component of statistical learning that simply does not correlate to executive function ability.

A second possibility is that measures of executive functioning can be subject to a lack of convergent validity (i.e., low cross-task correlations between measures of similar or the same effect(s)). Some recent studies (e.g., Keye, Wilhelm, Oberauer, & van Ravenzwaaij, 2009; Paap & Greenberg, 2013; Paap & Sawi, 2014) have reported very weak and often non-significant correlations among tasks that are often used to measure executive functioning in bilingual cognitive advantage literature, such as the ANT, Simon task, Eriksen flanker task and antisaccade task. In addition, other studies even reported weak cross-task correlations for tasks that supposedly measure the same concept. For instance, Salthouse (2010) found a remarkably low correlation (i.e., +0.03) between two slightly different flanker tasks (i.e., one used arrows and the other used letters for the stimuli). As such, our measures of executive functioning could be subject to low convergence validity and these tasks may not be measuring exactly what we intended to measure. Additional research is needed to work on the methodology of the tasks and ensure tasks that intend to measure executive functioning have high convergence validity when exploring the link between statistical word segmentation and executive functioning.

In conclusion, we have demonstrated that phonetic cues within a single person's speech can facilitate participants' separation of two language inputs with different statistical structures. We also found learning differences among participants with different language backgrounds, suggesting a simultaneous bilingual advantage in segmenting the languages. We proposed two

reasons why simultaneous bilinguals performed better in the dual-input segmentation task. First, early bilingual experiences may improve learners' skills to detect subtle cues (i.e., phonetic cues) in their language environments that differentiate two language inputs. Another possibility is that early bilingual experience may qualitatively influence learners' assumption of possible changes in ambient environment and therefore help learners to adapt to structural changes in new environment. In addition, we discovered a negative relationship between the ability of monitoring two task sets and segmentation accuracy of a particular language, supporting that over-learning one particular language may impede success of segmenting two different languages. Finally, although limited in scope, we found non-significant relationships between general executive functioning processes and the success of segmentation ability. This result may point to the fact that statistical learning is a non-unitary learning mechanism and/or that there is no strong relationship between executive functioning and statistical learning. Further research with additional statistical learning and executive function tasks, particularly those with strong convergent validity, are required to discover if such a relationship exists. Our findings add to the limited literature on dual-input statistical learning, particularly by demonstrating that natural intra-speaker cues can lead to the separation of speech streams for segmentation and that early bilingual experience can facilitate later adult statistical segmentation across two language inputs.

References

- Antovich, D. M., & Graf Estes, K. (2018). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science*, 21(2), e12548. doi: 10.1111/desc.12548
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321-324.
doi:10.1111/1467-9280.00063a
- Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, 15(7), 460-466. doi :
10.1111/j.0956-7976.2004.00702.x
- Basseville, M., & Nikiforov, I. (1993). *Detection of abrupt changes – Theory and application*. Englewood Cliffs, NJ: Prentice-Hall.
- Barac, R., & Bialystok, E. (2012). Bilingual Effects on Cognitive and Linguistic Development: Role of Language, Cultural Background, and Education. *Child Development*, 83(2), 413-422. doi: 10.1111/j.1467-8624.2011.01707.x
- Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, 2, 324. doi: 10.3389/fpsyg.2011.00324
- Basseville, M., & Nikiforov, I. (1993). *Detection of abrupt changes – Theory and application*. Englewood Cliffs, NJ: Prentice-Hall.

- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.
- Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*, 290-303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences, 16*, 240-250. doi: 10.1016/j.tics.2012.03.001
- Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics, 24*(1), 27-44. doi: 10.1017/S014271640300002X
- Boersma, P., & Weenik, D. (2012). Praat: Doing phonetics by computer [Computer program]. Version 5.1. Retrieved from <http://www.fon.hum.uva.nl/praat/>
- Bogulski, C. (2013). *Are bilinguals better learners? A neurocognitive investigation of the bilingual advantage* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database (UMI No. 1511655851).
- Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *Modern Language Journal, 64*, 311-317. doi: 10.1111/j.1540-4781.1980.tb05198.x
- Bulgarelli, F., & Weiss, D. J. (2016). Anchors aweigh: The impact of overlearning on entrenchment effects in statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(10), 1621-1631. doi: 10.1037/xlm0000263

- Burns, T., Yoshida, K., Hill, K., & Werker, J. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28(3), 455-474. doi:10.1017/S0142716407070257
- 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
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America*, 54(2), 421-428. doi: 10.1121/1.1913594
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child development*, 72(4), 1032-1053. doi: 10.1111/1467-8624.00333
- Chung-Fat-Yim, A., Sorge, G. B., & Bialystok, E. (2017). The relationship between bilingualism and selective attention in young adults: evidence from an ambiguous figures task. *The Quarterly Journal of Experimental Psychology*, 70(3), 366-372. doi: 10.1080/17470218.2016.1221435
- Cole, R., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and Production of Fluent Speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.
- Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, 106, 59–86. doi: 10.1016/j.cognition.2006.12.013

- Endress, A. D., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language*, 60(3), 351-367. doi: 10.1016/j.jml.2008.10.003
- Erickson, L., Kaschak, M., Thiessen, E., & Berry, C. (2016). Individual differences in statistical learning: Conceptual and measurement issues. *Collabra: Psychology*, 2(1), 14. doi: 10.1525/collabra.41
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14(3), 340-347. doi: 10.1162/089892902317361886
- Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23(4), 527-552.
- Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition*, 43(2), 226-236. doi: 10.3758/s13421-014-0461-7
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., ... Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, 5, 11. doi: 10.3389/fpsyg.2014.00011
- Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science*, 33(6), 1087-1116. doi: 10.1111/j.1551-6709.2009.01041.x

- Graf Estes, K., & Hay, J. (2015). Flexibility in Bilingual Infants' Word Learning. *Child Development, 86*(5), 1371-1385. doi: 10.1111/cdev.12392
- Grey, S. E. (2013). *A neurocognitive investigation of bilingual advantages at additional language learning* (Doctoral dissertation). Available from ProQuest Dissertations and Theses Global (UMI No. 3590106).
- Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America, 107*(5), 2711-2724. doi: 10.1121/1.428657
- Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science, 10*(5), 164–168. doi: 10.1111/1467-8721.00140
- Hilchey, M., Saint-Aubin, J., & Klein, R. (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. Schwieter (Ed.), *Cambridge Handbook of Bilingual Processing*, pp. 586-613. Cambridge: Cambridge University Press. doi:10.1017/CBO9781107447257.026
- Karuza, E. A., Balewski, Z. Z., Hamilton, R. H., Medaglia, J. D., Tardiff, N., & Thompson-Schill, S. L. (2016). Mapping the Parameter Space of tDCS and Cognitive Control via Manipulation of Current Polarity and Intensity. *Frontiers in Human Neuroscience, 10*, 665. doi: 10.3389/fnhum.2016.00665
- Keye, D., Wilhelm, O., Oberauer, K., & Van Ravenzwaaij, D. (2009). Individual differences in conflict-monitoring: testing means and covariance hypothesis about the Simon and

the Eriksen Flanker task. *Psychological Research PRPF*, 73(6), 762-776. doi: 10.1007/s00426-008-0188-9

Kousaie, & Phillips. (2012). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research*, 1446, 71-90. doi: 10.1016/j.brainres.2012.01.052

Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience*, 8, 250. doi: 10.3389/fnbeh.2014.00250

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, 25(5), 497-514. doi: 10.1080/20445911.2013.799170

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9(2), F13-F21. doi: 10.1111/j.1467-7687.2006.00468.x

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940–967. doi:10.1044/1092-4388(2007/067)

Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: an individual differences study. *Language Learning*, 62(1), 302–331. doi: 10.1111/j.1467-9922.2010.00626.x

- Mitchel, A. & Weiss, D.J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, 25(4), 456-482. doi: 10.1080/01690960903209888
- Morales, J., Yudes, C., Gómez-Ariza, C. J., & Bajo, M. T. (2015). Bilingualism modulates dual mechanisms of cognitive control: Evidence from ERPs. *Neuropsychologia*, 66, 157-169. doi: 10.1016/j.neuropsychologia.2014.11.014
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232-258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K., & Sawi, O. (2014). Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology*, 5, 962. doi: 10.3389/fpsyg.2014.00962
- Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge” hypothesis as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, 121(2), 130–143. doi: 10.1016/j.bandl.2011.05.003
- Poepssel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, 152, 9-19. doi: 10.1016/j.cognition.2016.03.001
- 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

- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25-42. doi: 10.1146/annurev.ne.13.030190.000325
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879-891. doi: 10.3758/BRM.40.3.879
- Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, 17(04), 682-691. doi: 10.1017/S1355617711000580
- Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13, 253–262. doi: 10.1017/S1366728909990526
- Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from <http://www.pstnet.com>.
- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to Represent a Multi-Context Environment: More than Detecting Changes. *Frontiers in Psychology*, 3, 228. doi: 10.3389/fpsyg.2012.00228
- Mitchel, A. & Weiss, D.J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, 25(4), 456-482. doi: 10.1080/01690960903209888
- Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, 47(4), 1343-1355. doi: 10.3758/s13428-014-0543-2

- Ratiu, I., & Azuma, T. (2015). Working memory capacity: Is there a bilingual advantage? *Journal of Cognitive Psychology*, 27(1), 1-11. doi: 10.1080/20445911.2014.976226
- Sabourin, L., Brien, C., & Burkholder, M. (2014). The effect of age of L2 acquisition on the organization of the bilingual lexicon: Evidence from masked priming. *Bilingualism: Language and Cognition*, 17(3), 542-555. doi: 10.1017/S1366728913000643
- Sabourin, L., Leclerc, J-C., Lapierre, M., Burkholder, M., and Brien, C. (2016). *The Language Background Questionnaire in L2 Research: Teasing Apart the Variables*. Proceedings of the Canadian Linguistic Association. Available online.
- Sabourin, L., & Vinerte, S. (2015). The bilingual advantage in the Stroop task: simultaneous vs. early bilinguals. *Bilingualism: Language and Cognition*, 18(2), 350-355. doi: 10.1017/S1366728914000704
- Saffran, J. R. (2003). Statistical language learning mechanisms and constraints. *Current Directions in Psychological Science*, 12(4), 110–114. doi: 10.1111/1467-8721.01243
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928. doi: 10.1126/science.274.5294.1926
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621. doi: 10.1006/jmla.1996.0032
- Salthouse, T. A. (2010). Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. *Brain and Cognition*, 73(1), 51-61. doi: 10.1016/j.bandc.2010.02.003

- Salvatierra, J. L., & Rosselli, M. (2011). The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism*, 15(1), 26-37. doi: 10.1177/1367006910371021
- Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Phil. Trans. R. Soc. B*, 372(1711), 20160059. doi: 10.1098/rstb.2016.0059
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, 81, 105-120. doi: 10.1016/j.jml.2015.02.001
- Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, 54(1), 1-32. doi: 10.1016/j.cogpsych.2006.04.002
- Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, 51(3), 381-398. doi: 10.1016/j.jml.2004.05.003
- Singh, L. (2017). Bilingual infants demonstrate advantages in learning words in a third language. *Child Development*. doi: 10.1111/cdev.12852. [Epub ahead of print].
- Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in Psychology*, 2, 123. doi: 10.3389/fpsyg.2011.00123

- Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97(2), B25-B34. doi: 10.1016/j.cognition.2005.01.006
- Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the gap. *Studies in Second Language Acquisition*, 33, 339-372. doi: 10.1017/S0272263111000015
- Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, 5, 953. doi: 10.3389/fpsyg.2014.00953
- Weiss, D.J., Gerfen, C. & Mitchel, A.D. (2009) Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development* 5, 30-49. doi: 10.1080/15475440802340101
- Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2010). Colliding cues in word segmentation: the role of cue strength and general cognitive processes. *Language and Cognitive Processes*, 25(3), 402–422. doi: 10.1080/01690960903212254
- Weiss, D. J., Poepsel, T., & Gerfen, C. (2015). Tracking multiple inputs: The challenge of bilingual statistical learning. In P. Rebuschat (Ed.), *Implicit and Explicit Learning of Languages*, pp. 167-190. Amsterdam, the Netherlands: John Benjamins Publishing Company.
- 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

- Yim, D., & Rudoy, J. (2013). Implicit statistical learning and language skills in bilingual children. *Journal of Speech Language and Hearing Research*, 56(1), 310–322. doi: 10.1044/1092-4388(2012/11-0243)
- Zinszer, B. D., & Weiss, D. J. (2013). When to hold and when to fold: Detecting structural changes in statistical learning. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3858–3863). Austin, TX: Cognitive Science Society.

Table 1. *Participants' Language Backgrounds in Experiment 1.*

	Average age	English self-rated proficiency (%)	English objective proficiency (%)	English usage in daily life (%)	French self-rated proficiency (%)	French objective proficiency (%)	French usage in daily life (%)	Other language(s) self-rated proficiency (%)	Other language(s) usage in daily life (%)
English Monolinguals (n = 33, 16 female)	20.21 (2.63)	97.35 (9.13)	89.27 (9.54)	97.73 (2.89)	16.92 (20.06)	4.12 (6.67)	1.69 (2.24)	5.99 (11.44)	0.48 (2.06)
Simultaneous French-English bilinguals (n = 32, 19 female)	20.13 (3.71)	94.01 (11.05)	88.90 (8.97)	63.22 (15.74)	95.96 (8.74)	68.39 (12.65)	36.68 (15.73)	7.55 (12.62)	0.00 (0.00)
Sequential French-English bilinguals (n = 34, 20 female)	21.03 (3.35)	88.97 (17.24)	88.61 (7.75)	66.24 (25.63)	73.59 (29.67)	60.30 (16.56)	33.64 (25.69)	9.07 (12.84)	0.09 (0.31)
Multilinguals (n = 16, 7 female)	19.19 (3.17)	91.80 (15.11)	84.38 (10.23)	67.97 (18.37)	82.03 (20.33)	66.53 (13.15)	27.31 (17.01)	60.93 (18.25)	7.33 (4.76)

Note. The table presents the average language proficiency and language usage in daily life across participants. Number within parentheses denotes the standard deviation of the corresponding values. Bilingual and multilingual participants scored lower in the French objective proficiency test than in the English, however this test has been established to be more difficult (i.e., French monolinguals' scores on the French cloze test are similar to those we obtained with our bilingual and multilingual groups; see Sabourin & Vinerte, 2018).

Table 2. *Auditory Stimuli in the Language Segmentation Task.*

Stimuli for Language 1	
Original in Weiss et al., (2009)	Modified stimuli in our study
Words	
bə-tɪ-gu	bə-tɪ-gu
sɪ-tʃə-vi	sɪ-tʃə-vi
vʊ-bo-sæ	vʊ-bo-sa
to-gʊ-tʃa	to-gʊ-tʃa
Part words	
sæ-to-gʊ	sa-to-gʊ
gu-vʊ-bo	gu-vʊ-bo
tʃa-sɪ-tʃə	tʃa-sɪ-tʃə
vi-bə-tɪ	vi-bə-tɪ
Stimuli for Language 2	
Words	
gu-pæ-tə	gu-pa-tə
dʒi-ga-pʊ	dʒi-ga-pʊ
sæ-dʒu-bo	sa-dʒu-bo
ta-bi-sɪ	ta-bi-sɪ
Part words	
bo-dʒi-ga	bo-dʒi-ga
sɪ-gu-pæ	sɪ-gu-pa
pʊ-ta-bi	pʊ-ta-bi
tə-sæ-dʒu	tə-sa-dʒu

Table 3. *Multiple Regression Models.*

Independent variables	<u>Model 1 (AL1 as DV)</u>			<u>Model 2 (AL2 as DV)</u>		
	β	S.E.	p	β	S.E.	p
Constant	0.35	0.14	0.012	0.32	0.12	0.011
Simultaneous bilingual group (SimB) (SimB coded as 1)	0.32	0.04	0.005	0.11	0.04	0.316
Sequential bilingual group (SeqB) (SeqB coded as 1)	0.17	0.04	0.139	0.04	0.04	0.696
Multilingual group (Multilinguals coded as 1)	-0.10	0.05	0.350	0.06	0.05	0.563
Alerting effect from ANT task	0.03	0.02	0.755	-0.08	0.02	0.360
Orienting effect from ANT task	0.05	0.00	0.615	0.04	0.00	0.661
Executive control from ANT task	-0.14	0.00	0.122	0.06	0.00	0.511
Working memory task composite	-0.03	0.00	0.727	0.03	0.00	0.770
Switching cost from color-shape switching task	-0.13	0.00	0.186	0.09	0.00	0.345
Mixing cost from color-shape switching task	0.21	0.00	0.032	-0.01	0.00	0.891
Age	0.22	0.03	0.018	0.44	0.03	0.000
Tested language (DV) accent (French-accented speech coded as 1)	0.15	0.01	0.094	0.03	0.00	0.734
R ²	0.49			0.48		
Omnibus test of the model	F (11, 99) = 2.85, p = 0.003			F (11, 99) = 2.64, p = 0.005		

Note. Model 1 used AL1 segmentation accuracy as dependent variable (DV). Model 2 used AL2 segmentation accuracy as DV.

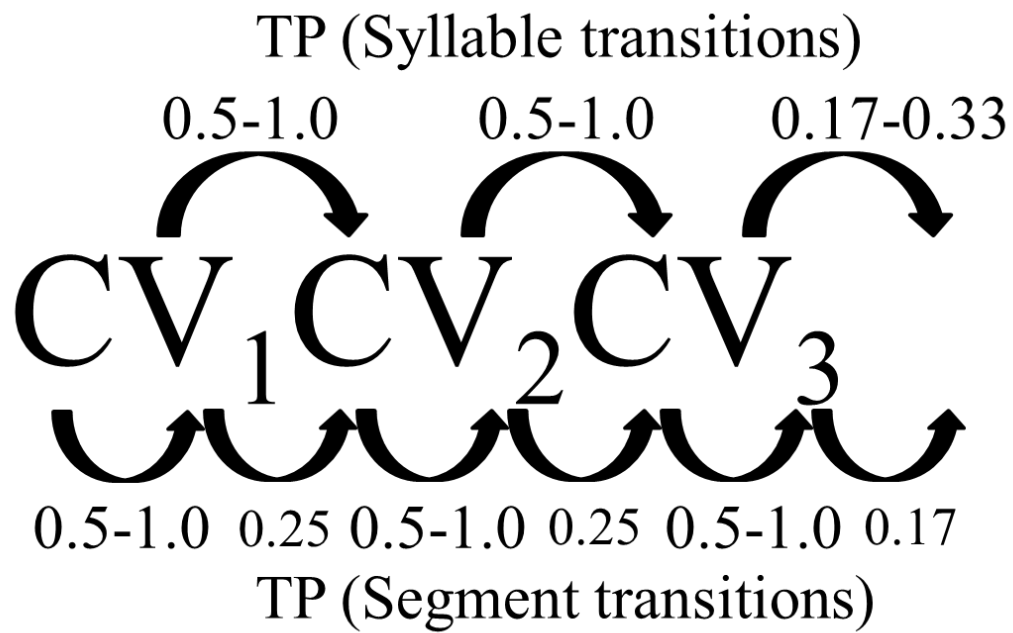


Figure 1. Transitional probabilities (transitional probability) of the underlying structures (i.e., inconsistent statistical cues) of the two languages when combined.

Figure 2a

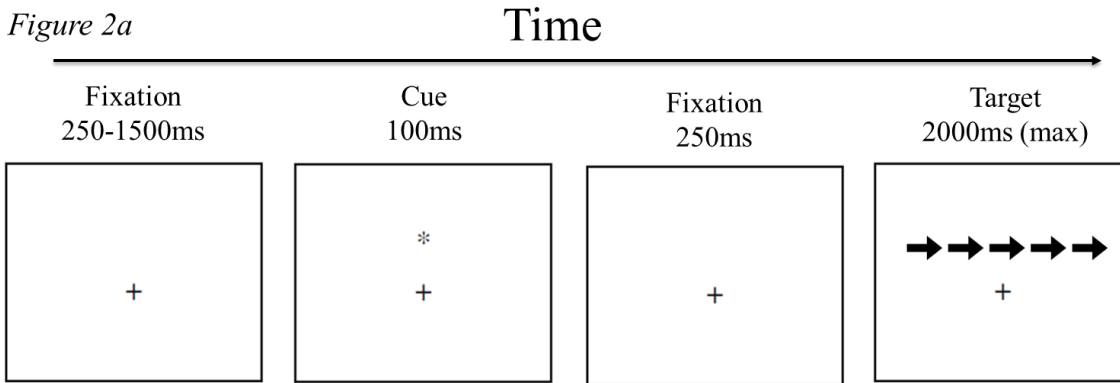


Figure 2b (i)

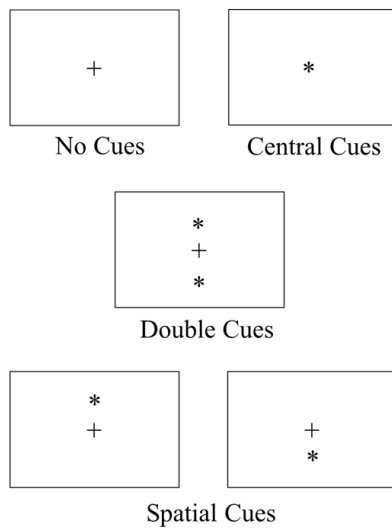


Figure 2b (ii)

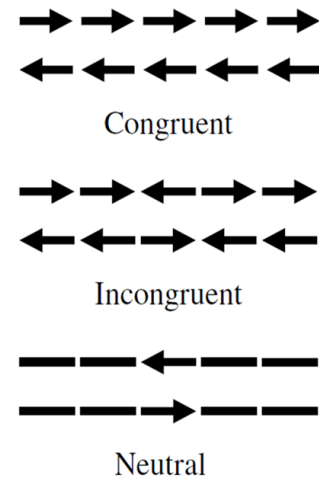


Figure 2. Experimental Procedure of the Attention Network Task (ANT). Figure 2a describes the procedure of the ANT task for each trial. Figure 2b describes the stimuli display for (i) the cue reaction components and (ii) the flanker components in the ANT task.

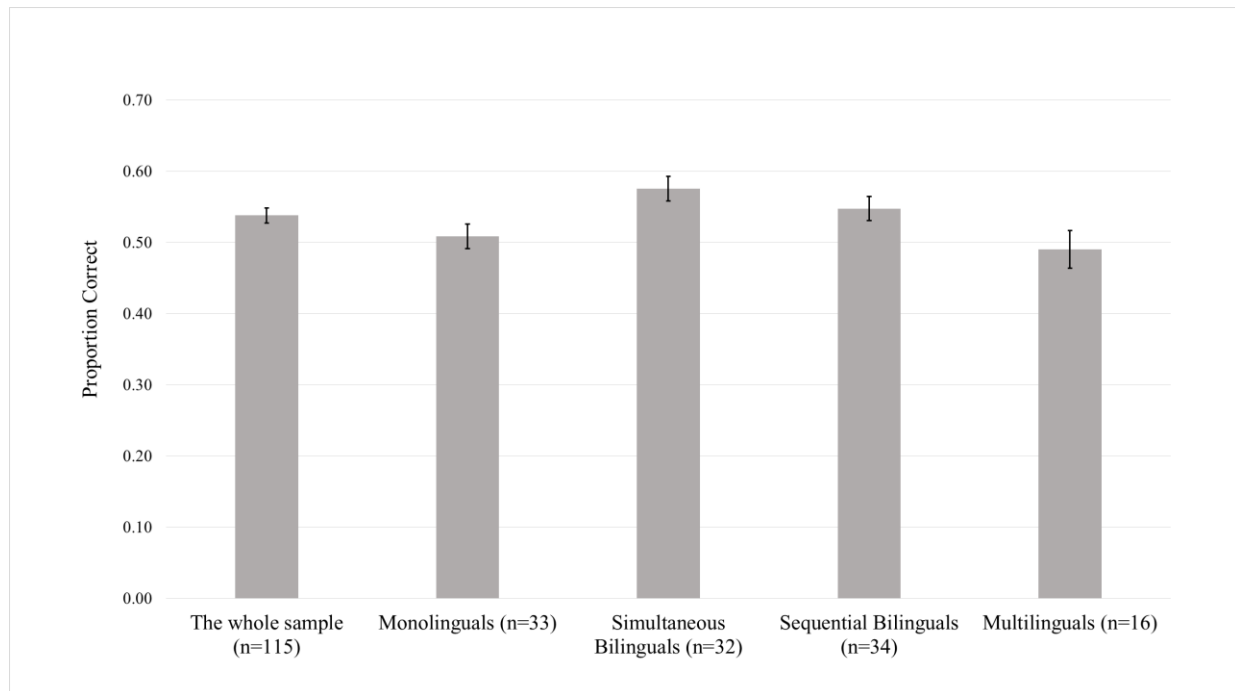


Figure 3. Average proportion of participants' scores of selecting correct words in each language.

Error bars indicate standard error of mean.

Infant Statistical Segmentation in a Language-mixing Bilingual Environment

Sin Mei Tsui^{1*}

Lucy C. Erickson^{2,3}

Erik D. Thiessen²

Christopher T. Fennell¹

1. School of Psychology, University of Ottawa. 2. Carnegie Mellon University. 3. University of Maryland

*Corresponding author

Abstract

A number of previous studies have shown that young infants can use syllable co-occurrence probabilities to segment words from fluent speech. However, most infant studies have investigated how infants statistically segment words from one language input. Whether infants can statistically segment two languages is largely unknown. An intrinsic challenge of statistically segmenting two languages is that languages tend to overlap at the syllable level. Successful segmentation thus relies on whether infants can represent two separate language inputs and track statistics of each individual language. In addition, parents often mix two languages in bilingual language environments (Byers-Heinlein, 2013), further making the task of segmenting two languages more difficult. Here, our paper investigated whether monolinguals and bilinguals of 9.5 months can make use of French and English phonetic cues (i.e., accented speech) to statistically segment words from two overlapping artificial languages within a single individuals' speech. In Experiment 1, our control study showed that monolingual English infants successfully segmented words from the individual languages. In Experiments 2 and 3, when infants were presented with two interleaved languages, we found that monolingual English infants, but not bilingual French-English infants, can make use of French and English phonetic cues to segment words from the two languages successfully. Experiment 4 demonstrated that infants' success of segmentation is not related to their cognitive ability. Our findings are discussed in terms of the differential processing of native and non-native phonetic cues in monolingual and bilingual infants.

Keywords: infant statistical learning, word segmentation, bilingualism, language mixing

Infant Statistical Segmentation in a Language-mixing Bilingual Environment

Learners often do not receive explicit information as to where to find words within fluent speech streams when hearing a new language. Words in the speech stream are not consistently marked by pauses and are rarely produced in isolation (Aslin, Woodward, LeMendola & Bever, 1996; Brent & Siskind, 2001; Cole & Jakimik, 1980). Given the lack of explicit cues, how do young infants, who are novice learners of their language(s), extract words from fluent speech stream? One possibility is that young infants can extract words based on the transitional probabilities across syllables in the speech stream. Syllable-based transitional probability refers to the probability of syllable X given the occurrence of syllable Y. Syllables that constitute words have higher transitional probability because syllables within words tend to reliably co-occur with each other than those across word boundaries. Take the phrase “pretty baby” as an example. The syllable “pre” strongly predicts the syllable “tty” because “pre” in English is followed by relatively few syllables (e.g., pre-pare, pre-tty, pre-dict). Therefore, there is a high transitional probability that “pre” will be followed by “tty” (approximately 80% in English speech to young infants, Saffran, 2003). However, the final syllable of “pretty” does not strongly predict the first syllable of “baby” since it can precede many other syllables (e.g., pretty dog, pretty cat, etc.). Further, these syllables do not comprise an English word (i.e., teebay is not an English word). Thus, there is a low transitional probability that “tty” will precede “ba” (approximately 0.03% in English, Saffran, 2003). Based on these statistical structures, young infants can infer that “pretty” is likely to be a word and the transition from “ty” to “ba” is probably denoting a word boundary. The mechanism that underlies the process of segmenting words using syllable transitional probabilities is termed “statistical learning”.

Statistical learning is a learning mechanism that allows learners to detect statistical regularities by tracking stable patterns via sensory inputs in the ambient environment. A number of prior studies have demonstrated that infants as young as 7 to 8 months are able to extract novel words merely based on syllable transitional probabilities after a short exposure to a fluent speech stream (e.g., Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996). In addition, several studies have shown that infants can readily use statistically segmented words for subsequent word learning tasks (Graf Estes, Evans, Alibali & Saffran, 2007; Hay, Pelucchi, Graf Estes, & Saffran, 2011; Erickson, Thiessen, & Graf Estes, 2014), supporting that statistical learning serves as a stepping stone for early lexical acquisition.

However, the majority of laboratory studies in infant statistical segmentation have focused on learning from a single input stream. While valid for monolingual infants, the results from this work cannot apply to those who necessarily need to segment two language input streams: bilingual infants. Statistical segmentation from two input streams is arguably more challenging when some syllables across the two languages overlap. For example, suppose a French-English bilingual learner hears the following syllables: *mo* then *to*. In French, *moto* is an actual word meaning motorcycle and thus these two syllables co-occur with greater probability. In contrast, *moto* is not a word in English. The co-occurrence between *mo* and *to* would therefore be less probable in English and likely to indicate a word boundary (e.g., *Elmo told Ernie*). In order to successfully segment words from two different language inputs, learners must represent two language systems and separately track transitional probabilities in each language.

To date, research on how learners segment words from two language inputs is relatively scarce. Pioneering work by Weiss and colleagues have suggested that monolingual adult learners are able to statistically segment words separately from two languages, even when the two

languages partially overlapped at the syllable level. The researchers found that learners can segment each language successfully only when provided with strong contextual cues that signaled the differences between the two languages. For example, when the two languages were associated with two different gendered voices (Weiss, Gerfen & Mitchel, 2009) or two different speakers' faces (Mitchel & Weiss, 2010).

The above studies have suggested that learners need contextual cues to segment words from two overlapping languages. Yet, some researchers have proposed that learners can segment two languages that overlap in syllables in the absence of contextual cues when learners monitor the variability of statistical cues in the language inputs (Basseville and Nikiforov, 1993; Gebhart, Aslin & Newport, 2009; Qian, Jaeger, & Aslin, 2012). This argument assumes that learners are able to track the variability of sensory inputs over time. When two languages partially overlap at the syllable level in a statistical segmentation task, the statistical cues to word boundaries are unstable if the learners combine two language inputs together (Weiss, Gerfen & Mitchel, 2009; see Weiss, Poepsel, & Gerfen, 2015 for a review). This means that when learners cannot separate two language inputs, the statistical structures of the incoming input should vary substantially. When the statistical cues are unstable, learners may think that this variability is noise. This leads learners to continue treating the input as one single language and thus they fail to segment words correctly from each individual language. However, if the learners treat this variability as an indicator of changes in underlying structures of the incoming input, they will begin to separate the two languages and successfully learn the languages (Qian, Jaeger, & Aslin, 2012).

Gebhart, Aslin and Newport (2009) discovered that monolingual adult learners were able to separate two language inputs that overlapped in syllables when there were no contextual cues that signaled the differences in the two languages. However, they only learned words from the first

presented language in the dual-input statistical segmentation task. The researchers found that adult learners could only learn two languages successfully when there were explicit and salient contextual cues that marked the differences of the two languages. For example, adult learners successfully learned two languages when the exposure length of second presented language was tripled in length. Learners also succeeded in segmenting two languages when they were directly informed that there were two inputs in the experiment and the languages were separated by 30 sec pauses. However, Gebhart and colleagues (2009) showed that not all contextual cues could help participants to learn two artificial languages in the task. When learners were cued by acoustic information (i.e., the two languages differed in terms of pitch), learners again could only learn the first presented language but not the second one. These studies have suggested that monolingual adult learners learned artificial languages best in a dual-input segmentation task when they were provided by strong and salient contextual cues that marked the differences in the two languages. A natural follow-up question is to ask whether infants, who are less cognitive mature, can statistically segment two overlapping languages when provided with contextual cues. This question is important because it allows us to investigate how infants overcome the challenges of differentiating two languages and discover words in a bilingual environment.

Two recent studies have shed light on whether infants can statistically segment words from two language inputs. Using the design found in Gebhart et al (2009), Bulgarelli, Benitez, Saffran, Byers-Heinlein and Weiss (2017) reported that 8-month-old monolingual infants failed to segment words in the dual-input segmentation task when there were no contextual cues signaling the differences in the two languages. The same group of authors also conducted a follow-up study to investigate whether pitch-shift cues can facilitate infant segmentation in the task but they found that infants still failed to learn the languages. On the contrary, Antovich and Graf Estes (2018)

have shown that bilingual infants of 14 months successfully segmented two language inputs using statistical cues when the languages were clearly marked by two different gendered voices but their monolingual counterparts failed in the same task. This result suggested that bilingual infants may be better in processing and segmenting two language inputs in a research setting, perhaps due to their daily experience in segmenting two languages or due to possible executive functioning advantages that arise out of attending to and inhibiting two languages (Bialystok, 2015). We will return to these possibilities later in the introduction. However, it is important to note that the two languages in Antovich and Graf Estes' study did not overlap at the syllable level. As such, it remains unknown whether infants can segment two languages that overlap in their syllabic inventory: a scenario that would occur in all natural bilingual learning situations (i.e., two languages cannot have mutually exclusive syllabic inventories due to the finite number of phonemes across the world's languages).

The present study aims to investigate whether infants can segment words from two languages when they share parts of their syllabic inventory. Further, we have extended the current literature in three ways: exploring whether phonetic cues within an individual's speech can facilitate statistical segmentation of two overlapping languages, whether bilingual infants outperform monolingual infants, and whether bilinguals' segmentation performance is related to their enhanced cognitive skills.

First, we explored whether contextual cues in a single individual's speech can facilitate infants' statistics-based segmentation of words. Previous studies have primarily explored how speaker-specific contextual cues (e.g. two voices differing in gender or two different talker faces; Mitchel & Weiss, 2010; Weiss, Gerfen & Mitchel, 2009) facilitate word segmentation from two languages. However, two distinct languages are not always spoken by two different speakers in

naturalistic infant bilingual environments. After all, a bilingual parent would need to switch between their languages depending on context (e.g., speaking English to one's partner, but French to one's parents). Furthermore, it is common that bilinguals often alternate between two languages in a single conversation when speaking with other bilinguals, a phenomenon that is known as code-switching (e.g., Poplack, 1980; see Heredia & Altarriba, 2001, for a review). Moreover, most bilingual parents have reported that they often code-switch between two languages when interacting with their infants (Byers-Heinlein, 2013). It is thus necessary to examine if contextual cues within a single person's speech are effective cues for statistical segmentation from two language inputs.

To address this concern, we investigated whether phonetic cues associated with French and English, which we call it as accented speech, can facilitate segmentation of two languages. Phonetic realizations of syllables often vary across two languages. In our previous example about the word *moto*, even though this word has the same phoneme sequence across French and English, the phonetic properties of this word differ across the languages. *Moto* would be produced with tenser vowels and less duration in French than in English. As such, we argue that the French and English phonetic variations (i.e., what we will call accented speech) could be considered a distinguishable contextual cue that can signal a change of languages within a single individual's speech, and therefore facilitate infants' segmentation of words from two languages. In our paper, we tested 9.5-month-old infants as monolingual infants at this age showed strong discrimination of accented speech (Schmale & Seidl, 2009; Schmale, Cristia, Seidl, & Johnson, 2010).

The second goal in the current paper is to test if early bilingualism may facilitate early statistical segmentation of two syllable-overlapping languages. Prior studies have shown that bilinguals outperformed monolinguals in several statistical learning tasks, such as segmenting

words from two languages (Antovich & Graf Estes, 2018; Bartolotti, Marian, Schroeder, & Shook, 2011), segmenting words from a tonal language (Wang & Saffran, 2014), and learning words across different scenarios (Poepsel & Weiss, 2016). Thus, it is possible that bilingual infants would outperform monolingual infants in our statistical segmentation task. We specifically tested French-English bilingual infants because we wanted to examine whether daily bilingual experience in hearing native French and English phonetic cues could bolster infants' processing of language inputs marked by those specific differences, thus facilitating infants' learning in the dual-input segmentation task.

The final goal of the current paper is to examine whether any possible bilingual advantage in the segmentation task could be explained by better cognitive skills. This motive was driven by recent research suggesting that bilingual infants have enhanced cognitive skills compared to their monolingual peers, including better inhibitory skills (Kovács & Mehler, 2009a; 2009b; Poulin-Dubois, Blaye, Coutya, & Bialystok, 2011), processing skills (Singh et al., 2015) and memory generalization (e.g., Brito & Barr, 2012). Given that past research has suggested a positive link between cognitive skills (e.g., inhibitory skills and attention) and statistical segmentation (Toro, Sinnott, & Soto-Faraco, 2005; Weiss, Gerken, & Mitchel, 2010), it is possible that if bilingual infants perform better in statistical learning task, it is related to enhanced cognitive skills.

Experiment 1

The purpose of this experiment was to test whether 9.5-month-old monolingual English infants can segment each artificial language individually. This experiment serves as a control study to demonstrate that each artificial language in either a Canadian-English or Canadian-French accent is learnable before testing infants with two interleaved artificial languages. In this experiment, we exposed 9.5-month-old monolingual English infants to an artificial language in

which the only cue to segmentation is transitional probability across syllables. If infants are able to use transitional probability to segment the language, they will attend differently to words (trisyllabic words with high transitional probability) and part-words (trisyllabic words with low transitional probability) at test. Since previous research has suggested that infants can segment words from an individual language stream on the basis of transitional probability when speech was unnaturally produced (i.e., monotonic, synthesized; e.g., Aslin, Saffran & Newport, 1998, Saffran et al., 1996), we expected that monolingual English infants would be able to track the transitional probability and segment words from our artificial languages, no matter the language was presented in English or the non-native French accent.

Method

Participants. We tested twenty-one typically developing 9.5-month-old infants (Mean age: 9.42, SD = 0.63, 10 females) from monolingual English homes (i.e., parents reported that infants were only exposed to English). Infants were randomly assigned to one of the four conditions: (i) heard Language 1 in a Canadian-English accent, (ii) heard Language 1 in a Canadian-French accent, (iii) heard Language 2 in a Canadian-English accent, or (iv) heard Language 2 in a Canadian-French accent. Four additional infants were tested but not included in the analysis because of crying (2), and fussiness (2).

Stimuli. We created the two artificial languages based on the scripts used in the incongruent condition of Experiment 2 in Weiss et al (2009), with the only change being that the vowel *a* replaced the vowel *æ* in our script. We made this change to ensure that all phonemes in the final script co-exist in both Canadian-English and Canadian-French (Table 1). This adjustment ensured that the speaker would not add, substitute or delete phonemes when producing the syllables in the scripts. Each language consisted of four trisyllabic CV.CV.CV words. Each word

in each language was presented an equal number of times and no word was repeated twice in a row. The transitional probabilities signaling the words and word boundaries in each language were 100% within a word and 33% at word boundary. For each language, test items consisted of two words (i.e., syllables co-occurred consistently with a transitional probability equal to 1.0) and two part-words (i.e., syllables co-occurred inconsistently with a transitional probability equal to 0.33). Apart from statistical cues, there were no pauses between syllables or other acoustic cues (e.g. stress) within each language to signal any word boundaries. It is important to note, while that each language consists of four words and four part-words, we only test infants in the current paper on two words and two part-words. Based on a pilot study, we found that a number of infants fussed out when they were tested on the exhaustive list of four words and four part-words. Importantly, the pilot study also revealed that infants may have preferences to some words and part-words in the two languages. Words and part-words that yielded longest and shortest looking times in the pilot study were not used in Experiment 1 because we want to minimize the influence of infants' preferences.

The stimuli in Table 1 were recorded by a balanced Canadian French-English bilingual female speaker. When recording the stimuli, the speaker produced all possible trisyllabic sequences in each language (Language 1 and Language 2). Only the middle syllable of each trisyllabic sequence was extracted when concatenating the syllables for the creation of the artificial languages. This would ensure that all syllables in the languages were fully co-articulated with their surrounding syllables. Thus, coarticulation was not a possible cue to word segmentation from the language speech streams. The speaker recorded the stimuli in two different sessions, producing all the stimuli in a Canadian-English accent on day 1 and all the stimuli in a Canadian-French accent on day 2. See Table 2 for the acoustic measurements of our stimuli.

Twelve local English-French simultaneous bilinguals then rated the concatenated artificial languages. The stimuli were rated on a scale of 1 to 5 in four aspects: (i) naturalness (1 indicating unnatural-sounding and 5 indicating natural-sounding); (ii) French-accent rating (1 indicating very French-sounding and 5 indicating very non-French-sounding); (iii) English-accent rating (1 indicating very English-sounding and 5 indicating very non-English-sounding); and (iv) overall sound quality of the language (1 indicating bad and 5 indicating excellent). The ratings for each language were averaged across all raters. The English-accented languages had an average of 3.02 for naturalness, 4.38 for French-accent rating (higher rating means the stimuli sounded very non-French), 1.75 for English-accent rating (lower rating means the stimuli sounded very English) and 3.38 for overall sound quality. The French-accented language had an average of 3.17 for naturalness, 1.90 for French-accent rating, 4.04 for English-accent rating, and 3.59 for overall sound quality.

Apparatus. Testing took place in a 2.38 m X 1.82 m sound-attenuated room dimly-lit by a 60 W lamp that was placed 80cm to the left of the infant and parent. Visual stimuli were projected onto a SmartBoard screen from a NEC Duocom LT280 projector. The auditory stimuli were played at 65 dB (+/- 5 dB) by two speakers placed below the SmartBoard screen. A digital camera, used to record infants' performance, was hidden under a table draped with black cloth below the screen. The camera lens poked out of a hole in the black cloth 10 cm below the screen.

Procedure. The experiment used a single-screen procedure. This procedure was a modification of previous adaption of the Headturn Preference Procedure for word segmentation studies (Altwater-Mackensen & Mani, 2013, Thiessen & Erickson, 2013). Infants seated on their parent's lap approximately 150 cm away from the screen during the whole experiment. To eliminate parental interference, parents were instructed not to talk to the infants and listened to

female vocals accompanied by music via headphones to mask the female voice of the stimuli. The experimenter observed the infants outside the testing room via a closed-circuit video system and used the Habit 2 software (Oakes, Sperka & Cantrell, 2015) to present stimuli and record looking time. The experimenter was blind to the nature of the stimuli being presented.

There were two phases in the experiment: a familiarization phase and a test phase. An animated cartoon baby was used to draw infants' attention to the display before the familiarization phase and before all subsequent test trials. When the infants oriented their attention to the screen, the familiarization phase began. During the familiarization phase, infants watched a colourful Winnie the Pooh video⁵ while listening to the artificial language. Infants heard a single language presented twice during the familiarization phase. Therefore, the duration of the familiarization phase was approximately 132 sec for French-accented speech and 140 sec for English-accented speech.

Infants immediately proceeded to the test phase after completing familiarization. In the test phase, infants received three blocks of four test items, resulting in 12 test trials. Within each block, infants heard 2 words and 2 part-words from the familiarized language. The order of test items were randomized and counterbalanced across infants. Accented speech was consistent across both familiarization and test phases (e.g. if infants heard Language 1 in French-accented speech, then they would be tested on Language 1 in French-accented speech). When listening to the audio stimuli during test trials, infants watched an animated green circle (changing in size over time) on

⁵ Previous studies (e.g., Thiessen & Erickson, 2013) presented infants with a checkerboard during familiarization. However, we found that checkerboard may be too boring for infants at 9.5 months as they tend to fuss out easily. Thus, we presented infants with a Winnie the Pooh video during familiarization. We argue that the Winnie the Pooh was attractive enough for infants to pay attention to the language but the objects displayed in the video were presented randomly such that infants could not form an association between objects in the video with the words they heard during familiarization.

the screen. Thus, infants' attention to the test items was measured by their looking time to the screen. Test trials ended when infants looked away for more than 1 sec or reached the maximum length of the test trial (20 sec).

Coding. Videos of all test trials were coded offline using a frame-by-frame analysis (1 frame = 30ms). Two experienced researcher assistants, who were blind to condition assignment, coded all data in the following analysis with high reliability ($r = 0.97$, $p < 0.0001$).

Results and Discussion

We identified one infant whose looking behaviour was more than 2.50 standard deviations away from the mean looking behaviour of the group. This infant was therefore excluded from the following analysis.

If infants learned the artificial language from familiarization, their average looking time to words and part-words should be significantly different, indicating successful discrimination of the target words for non-words. First, we ran a 2 (gender: male or female) X 2 (tested language: Language 1 or Language 2) X 2 (language accent: English or French) mixed ANOVA to test whether infants showed differential looking to word versus part-word test items. This ANOVA analysis revealed that gender, language accent and the tested language had no effect on infants' average looking time to the test items [$F_s < 2.0$, $p_s > 0.18$]. Therefore, we collapsed the data across these variables in the following paired t-test analysis.

To investigate whether there is a difference in looking time between words and part-words, we ran a paired t-test and found that infants looked longer to words (M: 10.76 sec) than part-words (M: 8.01 sec) [$t(19) = 5.67$, $p < 0.0001$]. This result suggested that infants segmented words from

both artificial languages, regardless of whether the language was presented in a French or English accent.

Nevertheless, the direction of infant preference to words and part-words has an implication of infant learning processes in this experiment. The classic work on novelty and familiarity preferences by Hunter and Ames (1988) demonstrated that infants tend to show a preference for novel stimuli only when familiarized stimuli are fully processed. If the infants' have learned aspects of the familiarization stimuli, but have not fully processed the information, they show a familiarity preference, as they remain attentive to the stimuli to continue to process the targets. In line with this perspective, many prior studies have demonstrated that infants generally paid more attention to part-words in statistical word segmentation experiments, as the part-words are novel (e.g., Aslin et al., 1998; Saffran et al., 1996). However, if the target language is more difficult than the typical segmentation experiment, infants may only partially encode the language stimuli and therefore would show a familiarity preference, looking longer to words at test (i.e., statistically coherent words are familiar).

Indeed, infants have paid more attention to words than part-words when the task is more difficult, such as when tested with a non-native language speech stream (Graf Estes, Gluck & Bastos, 2015) and when the test items were presented in different voices in comparison to those in familiarization (Graf Estes & Lew-William, 2015). Thus, our findings indicated that 9.5-month-old monolingual infants were able to segment words from our artificial speech stream but they generally may have found the segmentation task challenging. This perhaps is related to the fact that at least half of the languages were presented in a non-native (French) accented speech. However, we did not find a main effect of the language accent condition, suggesting that infants performed equally well when they were tested with French-accented speech or English-accented

speech. Thus, we are currently unsure why infants exhibited a familiarity preference in the experiment. Although infants showed a familiarity preference, the results in Experiment 1 suggest that monolingual English infants can successfully segment words based on the syllabic transitional probability in each artificial language. As we demonstrated that each artificial language is learnable, we could now move forward to test whether infants can segment words from two language inputs presented within the same learning session.

Experiment 2

This experiment investigated whether monolingual English 9.5-month-old infants can learn two interleaved artificial languages when there was overlap in the syllable inventory across the two languages. This overlap between the two languages increases the noise in the transitional probability cues of each language. This would lead to a failure of segmentation in the case where learners cannot separate the two languages, as they would be collapsing the statistics across the languages. As mentioned above, prior studies have suggested that adult learners performed best in this dual-input segmentation task if they were provided with strong contextual cues that highlighted the differences in the two languages (e.g., two different gendered voices, two different faces) (Mitchel & Weiss, 2010; Weiss et al., 2009). Thus, in the current study, we also provided infants with contextual cues that marked the differences in the two languages. First, the two artificial languages differ in terms of phonetic cues. For example, if Language 1 was presented in French accent, Language 2 would be presented in English accent. Accent within a single individual's speech is a more subtle cue because there are no changes in voices and speakers, but it is arguably more ecologically valid because bilingual parents often mix two languages together when talking to their infants (Byers-Heinlein, 2013). For example, it would be very rare that a bilingual English-French mother would only speak in English but never in French when talking in

the presence of her infant. As such, phonetic cues are arguably ecologically valid cues in a naturalistic language-mixing bilingual environment.

Further, following the design in Antovich & Graf Estes (2018), we added a 5 sec silent pause between the languages. Pilot data suggested that infants could not segment two artificial languages without this pause. However, it is important to note that a pause between languages alone did not support infants' segmentation of words from two languages (Antovich & Graf Estes, 2018; Bulgarelli et al., 2017). Thus, additional contextual cues (i.e., phonetic cues in our study) would still play the key role in successful word segmentation of the two overlapping languages.

If infants separate the two languages, they should be able to track the transitional probability of each language. Successful learners would demonstrate different preferences to items that conform the syllable transitional probability at test.

Method

Participants. Thirty-four monolingual English infants (Mean age = 9.6, SD = 0.64, 19 females) were tested. All infants were only exposed to English, as reported by parents. Four additional infants were tested but not included in the final analysis because of crying (1) and fussiness (3).

Stimuli. The artificial language stimuli were identical to those used in Experiment 1. The only difference is that we presented infants with both artificial languages in this experiment.

Apparatus and procedure. The apparatus and general procedure were the same as in Experiment 1, with the only difference being the addition of another language during familiarization. Language 1 and Language 2 were presented in an interleaved order and each language was presented twice to the infants. Therefore, the total length of exposure to the artificial

languages was longer than in Experiment 1 since we were presenting two artificial languages to the infants (approximately 409 sec total). In addition to adding a 5 sec silent pause between the two languages at each transition point, the key transition cues across Language 1 and Language 2 were English and French accents (i.e., Language 1 and Language 2 were always in different accents). For instance, English-accented Language 1 was paired with French-accented Language 2, whereas French-accented Language 1 was paired with English-accented Language 2.

As a reminder, the transitional probability signaling the word boundaries in each language were 1.0 within word and 0.33 across word. However, this presumes that participants can separate the two languages and calculate the statistical set of each. If participants combined the two languages together, the transitional probability would not provide consistent statistical cues. The transitional probability within words would vary from 0.5 to 1.0 whereas the transitional probability across word boundaries would vary from 0.17 to 0.33 (see Figure 1). Apart from statistical cues, there were no pauses or other acoustic cues (e.g. stress) between syllables within each language to signal any word boundaries. The orders of accent (hearing an English-accented language first or second) and presentation of the language (Language 1 presented first or second) were counterbalanced across infants. Thus, half of the infants heard Language 1 in an English accent and Language 2 in a French accent and the other half heard the opposite.

At test, infants were only tested on one of the artificial languages, allowing for a direct comparison to the test phase of Experiment 1. Half of the infants were tested on words and part-words from Language 1 and the other half were tested on items from Language 2. Accented speech was consistent across familiarization and test phases (e.g. if infants heard Language 1 in French-accented speech and Language 2 in English-accented speech, then they would be tested on either Language 1 in French-accented speech or Language 2 in English-accented speech).

Coding. As in Experiment 1, infants' looking time during test phase were coded by two trained research assistants. The inter-coder reliability was high ($r = 0.98$, $p < 0.0001$).

Results and Discussion

If infants were able to separate the two artificial languages and succeed in segmenting each language, they should look differentially to word test trials than to part-word test trials. In the following analysis, we added one more factor: presentation order of the tested language. It is possible that infants may allocate more attention to the first presented language in the familiarization phase and thus they may learn that language better. This hypothesized possibility stems from prior studies (e.g., Gebhart et al., 2009) that have suggested that adults only acquired the first presented language when learning two overlapping languages in a difficult context where no overt cues marked the differences between the two languages. Although we provided infants with phonetic cues to help them segment words from the artificial languages, it is still possible that infants found the task difficult due to their relative cognitive immaturity and the subtlety of our cues. This may lead to a similar primacy bias as that seen in adults.

Preliminary analysis indicated that there was no infant whose looking behaviour was more than 2.50 standard deviations away from the mean looking behaviour of the group. Thus, all data were entered in the following analysis. We conducted a 2 (test trial: words or part-words) X 2 (gender: male or female) X 2 (language accent: English or French) X 2 (tested language: Language 1 or Language 2) X 2 (presentation order: tested language presented first or second in familiarization) mixed ANOVA to see if infants show different looking preference to word and part-word test trials. We found a significant main effect of trial type [$F(1, 23) = 5.83$, $p = 0.024$], suggesting that infants looked significantly longer to the part-word test trials ($M = 10.07s$, $SD =$

2.63) than word test trials ($M = 8.66s$, $SD = 2.59$) (See Figure 2). There were no other significant main effects, nor any interactions between trial type and other factors [$F_s < 1.6$, $ps > 0.2$].

The results of this experiment suggest that monolingual English infants can separate the two artificial languages and successfully segment words from the interleaved dual-input speech stream. Specifically, this preference was not related to the accent of the language, or the tested language itself (Language 1 or Language 2), or the presentation order of the language during familiarization. Importantly, infants in this study demonstrated a novelty preference as they looked longer to the part-words than to the words. This result is in contrast to our findings in Experiment 1, since monolingual infants showed familiarity preference when they were tested with one language only. Infants' preference direction reveals the depth of infants' learning processes during familiarization phase (Hunter & Ames, 1988). Generally speaking, when infants completely encode the information in the familiarization phase, they would be ready to encode novel information and therefore shift from a familiarity preference (attending more to familiar stimuli) to a novelty preference (attending to novel stimuli). Thus, the findings in Experiment 2 suggest that monolingual infants could process the statistical cues to segmentation more thoroughly with two language inputs than with a single language input. Although the length of exposure to each individual language in Experiment 2 was the same as in Experiment 1, infants in Experiment 2 heard artificial language stimuli longer overall because they needed to learn two languages. The increased exposure to the task may be benefited the infants. This is because longer exposure to artificial languages overall may actually provide infants with more information about the nature of the segmentation task (i.e., tracking transitional probability across syllables). A better understanding of the nature of the segmentation task could help infants to fully process the

information and segment words better from the two languages, consequently showing a novelty preference to part-words at test.

In Experiment 2, we found that accented speech was an effective cue for monolingual infants to separate and learn two artificial languages. In the following, we extended this experiment to French-English bilingual infants to test whether the bilingual infants also succeed in the dual-input statistical segmentation task. Bilingual infants necessarily need to learn two different linguistic statistical structures in their language environments. It is thus ecologically valid to extend this study to a bilingual population and see if they can handle this task as well as monolinguals.

Experiment 3

This experiment investigated whether French-English bilingual infants can segment two artificial languages in a dual-input statistical segmentation task. Because monolingual infants in Experiment 2 were capable of using phonetic cues to statistically segment words from two overlapping languages, we expected that bilingual infants would do the same as past work (e.g., Antovich & Graf Estes, 2018) has suggested a bilingual advantage in the dual-input statistical segmentation task. Further, the French-English bilingual infants were regularly exposed to French and English phonetic cues in their daily language environments, thus the bilingual infants may readily use these cues to differentiate between two artificial languages.

Method

Participants. Thirty two French-English infants (Mean age: 9.51, SD: 0.70, 16 females) were tested. Infants were exposed to English and French since birth and had a maximum of 80% exposure to one language and a minimum of 20% exposure to the other language (Mean English exposure = 56.08%, SD = 16.23%; Mean French exposure = 43.65%, SD = 16.32%), as determined

by the Language Exposure Questionnaire (Bosch & Sebastián-Gallés, 1997). Nine additional infants were tested but not included in the final analysis because of crying or fussiness (7), exposed to English was up to 85% (1), being a Mandarin-Cantonese infant (1).

Stimuli. Stimuli were the same as in Experiment 2.

Apparatus, procedure and coding. Apparatus and procedure were the same as those described in Experiment 2. For coding, two research assistants coded all data included in the following analysis. The inter-coder reliability was high ($r = 0.93$, $p < 0.0001$).

Results and Discussion

An infant was identified as an outlier as this infant's looking behaviour was more than 2.50 standard deviations away from the mean looking behaviour of the group. This infant was therefore excluded from the following analysis.

We conducted a 2 (test trial: words or part-words) X 2 (gender: male or female) X 2 (language accent: English or French) X 2 (tested language: Language 1 or Language 2) X 2 (presentation order: tested language heard first or second in familiarization) mixed ANOVA to see if infants show different looking preferences to word and part-word test trials. None of the main effects or interaction effects were statistically significant [$F_s < 3.90$, $p_s > 0.10$]. Thus, our findings indicated that French-English bilingual infants failed to segment words from the interleaved dual-input speech stream.

It is important to note that bilingual infants differed in terms of their exposure to English and French. Some infants were more dominant in English or French (e.g., 70% exposure to English and 30% exposure to French) while others were more balanced in both English and French (e.g., 50% exposure to both English and French). Perhaps the degree of language exposure to French or

BILINGUAL STATISTICAL LEARNING

English may influence how infants perceive and utilize French and English phonetic cues to separate two language inputs in our study. If a bilingual infant has unbalanced exposure to the two languages, they may over-rely on the phonetic cues from the dominant language and/or not effectively use the phonetic cues of the non-dominant language. Thus, we also ran a correlational analysis to examine whether there was a relationship between the degree of bilingualism and infants' performance in the segmentation task. Degree of bilingualism was determined by the difference of percentage exposure between dominant language and non-dominant language. For example, an infant with 60% exposure to English and 40% exposure to French would get a score of 20, while an infant with 50% exposure to English and French would get a score of 0. Here, a smaller value in the degree of bilingualism reflected that the infant received a more balanced language exposure to French and English. On the other hand, performance in the segmentation task was measured by subtracting the looking time to words from that of part-word items at test. Positive and higher scores on this measure indicated that infants exhibited novelty preference and performed better at test. However, the Pearson correlation analysis revealed a non-significant relationship between degree of bilingualism and performance in segmentation task [$r = 0.22$, $n = 31$, $p = 0.230$], suggesting that more balanced exposure to English and French did not help infants to segment words from the two languages in Experiment 3.

Experiment 4

This experiment aimed to measure infants' cognitive skills in order to explore whether there is a link between infants' cognitive skills and their performance in the dual-input segmentation task. Infants who had participated in Experiment 2 and 3 were invited to participate in this experiment on a second visit to the lab two weeks later. Although the hypothesis that cognitive skills – particularly inhibition - would be related to the dual-input segmentation task was

inspired by the cognitive advantages seen in past work with bilinguals (e.g., Kovács & Mehler, 2009a), it turned out that monolingual infants actually performed better than bilingual infants in the segmentation task (Experiment 2 versus 3). Nevertheless, we still report this experiment in order to examine whether an individual's cognitive skills, regardless of language background, would facilitate segmentation in the dual-input task. This motive was driven by previous studies showing that there is a positive link between learners' statistical segmentation performance and their inhibitory control (Weiss, Gerken, & Mitchel, 2010) and attention (Toro, Sinnett, & Soto-Faraco, 2005).

This experiment involved a shortened anticipatory eye movement paradigm, modified from Experiment 3 in Kovács and Mehler (2009a) (Figure 3). Infants were trained to predict the locations of visual rewards based on the structures of tone-shape sequences. The key manipulation was that the structure of the tone-shape sequence would change in the middle of the experiment. Successful learners must inhibit the previously learned tone-shape sequence structure and then learn the new one. We argue that this task measures infants' general cognitive skills, including working memory, attention and inhibitory control. First, infants need to use their working memory to process the information of the tone-shape sequences. Next, infants need to pay attention to the common structures across tone-shape sequences and the association between those structures and the locations of visual rewards. Lastly, as mentioned above, infants must rapidly inhibit the previously learned tone-shape sequence structure in order to learn new tone-shape sequence structure during the post-switch phase.

Method

Participants. Twenty-four monolingual English infants from Experiment 2 and 23 French-English bilingual infants from Experiment 3 returned to the lab for this experiment two weeks after

they performed the dual input segmentation task. Additional four infants were tested but not included in the final analysis because of crying or fussiness.

Stimuli. All stimuli were organized into 3 tone-shape sequences of two structures: AAB or ABB. For AAB sequences, the first two tones and geometric shapes in the sequence were identical and the final pairing differed (e.g., participants were presented with a sequence “circle-tone C, circle-tone C, star-tone F”). By contrast, for ABB sequence, the last two shape-tone pairings were identical and the first pairing differed (e.g., participants were presented with a sequence “star-tone F, circle-tone C, circle-tone C”).

The visual stimuli comprised 9 AAB shape sequences and 9 ABB sequences. Following the methodology in Kovács and Mehler (2009a) study, specific geometric shapes were used to generate the sequences. Three geometric shapes (arrow, circle, pentagon) were used for objects in position A of the AAB and ABB sequences, whereas three other geometric shapes (moon, 5-pointed star, triangle) were used for position B objects. Each shape was presented in different colours. For symmetrical geometric shapes (i.e., circle, pentagon, 5-pointed star and triangle), the size of each shape was 32cm X 32cm. For asymmetric geometric shapes (i.e., arrow and moon), the size of each shape was 34cm X 32cm. The shape sequences appeared on screen in the following manner. The first shape of the sequences was presented on the left side of the screen alone. Next, the second shape was added in the middle of the screen while the first remained onscreen. Finally, the third shape was added on the right side of the screen so that all three shapes appeared simultaneously on the screen for 3 sec. All shape sequences were displayed against a black background.

The audio stimuli were sequences of three musical tones. Two tone structures (i.e., AAB and ABB) were constructed to pair with the corresponding visual shape sequences. Three musical

tones (A, D, E) could be paired with objects in position A, whereas three other musical tones (C, F, G) could be paired with objects in position B.

In this experiment, the tone-shape sequence would be followed by a visual presentation of two white squares and the visual rewards (see procedure for more details). The two white squares were presented side by side on the screen (each was 52.5cm to 53.5cm in size) for 1.5 sec. For the visual rewards, one of two puppets (i.e., a giraffe or hippopotamus toy) appeared inside one of the two white squares on the left or right side of the screen. The puppet loomed from 20cm X 30cm to 28cm X 49cm in size for 2 sec. The presentation of puppets was accompanied by a chime sound.

Apparatus and procedure. The apparatus remained identical to Experiment 1, 2 and 3. Infants were seated on a parent's lap during the whole experiment. The parent was wearing headphones to mask the sounds. The parent was also instructed to look at the infant or look at the centre of the screen in order to minimize their influence of infants' attention to a particular side of the screen. At the beginning of each trial, we presented infants with attention-getting stimuli (e.g., an image of a baby and audio of a baby giggle). Once infants oriented their attention to the screen, the experimenter pressed a key to present the orientation stimuli to the infants. For the first trial, infants saw a video where a rotating ball changed its position from the left side to the right side of the screen. The ball first appeared on the left side and remained onscreen for 3.3 sec. The ball then reappeared on the right side of the screen for another 3.3 sec. The trial served to accustom infants with the experiment procedure where they would be trained to look at the left and right side of the screen to predict different visual rewards based on the structures of tone-shape sequences.

After this orientation trial, there were two phases in this experiment: pre-switch and post-switch (see Figure 3). Each phase consisted of 6 trials. Although Kovács and Mehler (2009) presented infants with 9 trials for each phase, we decided to shorten the number of trials because

a pilot study suggested that infants tended to fuss out after the sixth trial in the pre-switch phase. In the pre-switch phase, infants were trained that one tone-shape sequence structure (either AAB or ABB) that predicted a visual reward in a particular location (i.e., on the right or left of the screen). On each trial, infants were first presented with a tone-shape sequence (e.g., AAB) for 3 sec. The tone-shape sequence would then be replaced by two white squares on the screen for 1.5 sec. During this window of time, infants could make an anticipatory eye movement by directing their eye gaze to the square where the object would appear (anticipatory window period). Finally, a looming puppet (e.g., giraffe) would appear on one side of the screen (e.g., right side) for 2 sec.

After presenting infants with 6 pre-switch trials, infants then entered a post-switch phase where a new structure of tone-shape sequences (e.g., ABB structure) predicted a different reward (e.g., hippopotamus) on a different location (e.g., left side). The procedure and length were identical to the pre-switch phase, aside from the differences above (i.e., sequence structure, visual reward type and location).

Coding. Following Kovács and Mehler (2009a), we coded infants' eye gaze during the anticipatory window period (1.5 sec) for the dependent variable (DV). We coded infants' eye gazes to the right and left positions. We only counted eye gazes to the appropriate reward location as correct responses. Eye gazes to the opposite side of the reward location were all counted as incorrect. Videos were coded frame by frame (30 frames per second). Two trained undergraduate coders coded all trials independently and the reliability between their coding was high ($r = 0.90$, $p < 0.0001$). To obtain the proportion of correct anticipatory looks, we divided the number of frames looking at the correct location by the sum of total frames that infants looked at the correct and incorrect positions during anticipatory window period. For example, on a test trial, an infant looked

at the correct position for a total of 20 frames and looked at the incorrect position for a total of 10 frames. The proportion of correct anticipatory looks is 0.67 for this infant at this particular trial.

Data analysis. As mentioned, the DV was infants' proportion of correct anticipatory looks in each trial. In addition, we expected that infants would increase the proportion of correct anticipatory looks over time, thus we tested whether the relationship between the DV and the number of trials (i.e., DV-trial slope) was positive or not. Here, we investigated whether infants' language background would influence the DV-trial slope because the infant bilingual cognitive advantage has suggested that bilingual infants may have stronger cognitive skills. For example, if bilingual infants have stronger cognitive skills that may support their learning in the dual-input segmentation task in Experiment 2, they may have faster rate of learning the association between tone-shape sequence and the corresponding visual rewards. Thus, the DV-trial slopes in the bilingual infant group would be steeper. To model the variations of the DV and the DV-trial slopes in the experiment, we employed hierarchical linear modeling in our analyses. Because we have different hypotheses about infants' performance in the pre-switch and the post-switch phases. We ran two separate hierarchical linear models, one examined infants' performance in the pre-switch phase (model 1) and the other examined infants' performance in the post-switch phase (model 2). We predicted that infants' inhibitory control ability would only be reflected in their performance during the post-switch phase. As such, we expected that monolingual and bilingual infants would have similar proportion of correct anticipatory looks and DV-trial slopes during the pre-switch phase. By contrast, monolingual and bilingual infants would differ in terms of their proportion of correct anticipatory looks and DV-trial slopes during the post-switch phase and these differences would be related to the infants' performance in the dual-input segmentation (measured by the looking time difference between words and part-words in Experiment 2 and 3).

We used the lme4 package in R (Bates, Mächler, Bolker & Walker, 2015) to perform the hierarchical linear regression models. The regression models were fit by the restricted maximum likelihood approach (REML) and the p values in the models were estimated by Satterthwaite approximations in the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016).

The DV were nested within infant participants. First, a random intercept was specified to allow differences in DV vary randomly across infant participants in both model 1 and 2. To estimate the intraclass correlations that measure the between-participant variability in model 1 and 2, we first fit the data into two null intercept-only models. The intraclass correlation for model 1 was 0.51 and the intraclass correlation for model 2 was 0.28, suggesting that the differences in DV vary substantially across infants in both models and therefore we were statistically justified to run hierarchical linear regression models.

In addition to allowing intercepts vary as random effects in the models, we also allowed the DV-trial slopes vary randomly across individuals in both model 1 and 2. We conducted further analyses to compare the null-models with random intercept only with the models including both random intercepts and slopes. We found that model 1 and model 2 were not significantly improved when it specified both intercept and DV-trial slope as random effects [χ^2 s < 3.30 , ps > 0.08], thus adding random slopes in the models were not statistically justified. However, we were still able to add random slope in model 2 for theoretical reasons (Huta, 2014). In the current paper, we expected that infants' performance would improve over time in both pre-switch and post-switch phases. This learning effect would be reflected by a positive DV-trial slope in the models. Thus we added random slopes in models 1 and 2. In all models, the number of trials was coded with trial 1 as the reference level in the pre-switch and post-switch phases.

Furthermore, we included infants' language background, infants' performance in the dual-input segmentation task (as measured by the difference in looking time between words and part-words in Experiment 2 and 3) as fixed effects in the model to explore whether these two independent variables (IV) explain variability in DV across infants. Monolingual English infants were served as the reference level (coded as 0) in the infants' language background IV. For the IV related to infants' performance in the dual-input segmentation task, we did not center this IV because zero in this IV has a meaningful interpretation which indicates that infants did not discriminate between word and part-words (i.e., no differences in looking time between words and part-words). Finally, an interaction term between infants' language background and random slope was added in the models to specifically explore whether infants' language background influences their learning rate (i.e., DV-trial slope) over time.

Data screening. For data screening, we first screened each variable for univariate outliers at the lowest level using the criteria $Z = 2.50$ and found no outliers across participants in DV and the two IV in model 1 and 2. We next examined whether the residuals of the random-intercepts and random-slopes in the model 1 and model 2 were normally distributed and whether there was residual outlier. For model 1, we found two outliers in the residuals using the criteria $Z = 2.50$. The residuals of the random-intercepts and random-slopes were normally distributed after removing the outliers. For model 2, we found one outlier in the residuals of random-slopes using the criteria $Z = 2.50$. Again, the residuals of random-intercepts and random-slopes were normally distributed after removing this outlier. Further, we examined the coefficient reliabilities of random-intercepts and random-slopes in model 1 and 2. The coefficient reliabilities in both models were higher than 0.05 and thus we did not need to simplify the model. We also found that the inter-correlations between random-intercepts and random-slopes in model 1 and 2 were moderate in size

(i.e., less than 0.9), suggesting that we did not have redundant coefficients in our models and therefore the outputs in our models would be stable. Lastly, we checked the assumptions of linearity, homogeneity of variances of random-intercepts and slopes, and multicollinearity of model 1 and 2. We did not find any violations of the above assumptions in all models.

Results

The coefficients of all models were all displayed in Table 3. In model 1, the effects of infants' language background [$\beta = 0.020$, $S.E. = 0.069$, $p = 0.770$] and segmentation ability [$\beta = -0.001$, $S.E. = 0.012$, $p = 0.928$] were not significant predictors for the variability of the DV in the pre-switch phase. We also found that the interaction between infants' language background and DV-trial slopes was not significant [$\beta = 0.020$, $S.E. = 0.019$, $p = 0.280$], suggesting that the learning rate between monolingual and bilingual infants are similar in the pre-switch phase. The average DV-trial slope was significantly higher than zero across all participants [$\beta = 0.020$, $S.E. = 0.008$, $p = 0.012$], suggesting that infants generally improve their proportion of correct anticipatory looks over time during the pre-switch phase. Finally, a significant intercept in the model 1 indicated that infants on average had a positive proportion of correct anticipatory looks in the pre-switch phase [$\beta = 0.49$, $S.E. = 0.051$, $p < 0.001$]. A one-tailed t test revealed that the average proportion of correct anticipatory looks in the pre-switch phase was significantly greater than the chance level [$M = 0.59$, $t(276) = 4.909$, $p < 0.001$].

In model 2, we also did not find significant effects of infants' language background [$\beta = 0.003$, $S.E. = 0.066$, $p = 0.961$] and segmentation ability [$\beta = 0.005$, $S.E. = 0.012$, $p = 0.660$] for the variability of the DV in the post-switch phase. The interaction between infants' language background and DV-trial slopes was not significant [$\beta = -0.016$, $S.E. = 0.018$, $p = 0.375$], suggesting that the learning rate between monolingual and bilingual infants were similar in the

post-switch phase. The average DV-trial slope was significantly higher than zero across all participants [$\beta = 0.032$, $S.E. = 0.012$, $p = 0.009$], suggesting that infants generally improved their proportion of correct anticipatory looks over time during the post-switch phase. Finally, a significant intercept in the model 2 indicated that infants on average has a positive proportion of correct anticipatory looks in the post-switch phase [$\beta = 0.34$, $S.E. = 0.063$, $p < 0.001$]. A follow-up one-tailed t test revealed that the average proportion of correct anticipatory looks in the post-switch phase was not significantly greater than the chance level [$M = 0.45$, $t(268) = -2.63$, $p > 0.99$]. See Figure 4 for infants' performance across trials during the pre-switch and post-switch phases.

Discussion

The results of model 1 were consistent with our hypothesis that monolingual and bilingual infants would have similar proportion of correct anticipatory looks and DV-trial slopes during the pre-switch phase. However, the results of model 2 has suggested that both monolingual and bilingual infants had difficulty overcoming the learned tone-shape structure in the pre-switch phase and failed to learn another tone-shape structure in the post-switch phase. Both monolingual and bilingual infants performed similarly in terms of their proportion of correct anticipatory looks and learning rate. Thus, our findings did not find evidence supporting that infant bilingualism improves infants' inhibitory control or their learning rate in this task. Importantly, we did not find that infants' segmentation performance in the dual-input task was related to infants' inhibitory control. As such, our findings suggested that infants' cognitive skills, at least the ones tested here, did not relate to their segmentation performance in the dual-input task. However, it is important to note that infants' average proportion of correct anticipatory looks during post-switch phase was not significantly above chance level. Thus, infants generally did not show any learning of the associations between

the tone-shape structure and the visual rewards during the post-switch phases. The current experiment may be too cognitively challenging for 9.5-month-old infants such that both monolingual and bilingual infants failed to learn during the post-switch phase in the task. It is important to note that infants' inhibitory control ability in the current task is reflected by their ability to switch response during the post-switch phase. Thus, infants' failure of learning in the post-switch phase implies that the current experiment was not sensitive to measure infants' inhibitory control ability. In short, our failure of finding a relationship between infants' segmentation performance and their inhibitory control ability may relate to the fact that we failed to measure infants' inhibitory control ability in the current experiment.

General Discussion

The current paper investigated whether monolingual and bilingual infants can segment words from two interleaved artificial languages based on transitional probability cues. Importantly, these two languages overlapped partially at the syllable level, thus successful word segmentation required infants to form two separate language representations in the dual-input segmentation task in order to track transitional probabilities from each individual language accurately. In our studies, we used phonetic cues (i.e., accents) from two different languages as the marker for infants to separate the two overlapping languages. Our experiment provides a better model of how infants segment words in a naturalistic bilingual environment because bilingual infants would usually hear one speaker producing two different languages, each with their associated phonetic cues (Byers-Heinlein, 2013). This reflects ecological validity and extends our understanding of how infants find words from two language inputs in the real world.

Results from Experiment 1 demonstrated that monolingual English infants were able to statistically segment words from an individual language speech stream when the speech was

presented in either native English or non-native French accents. This confirms that each language stream in Experiment 2 and 3 was learnable individually. In Experiment 2 and 3, we investigated whether monolingual English and bilingual French-English infants can successfully segment words from two interleaved languages that differed in terms of French and English accents. Results demonstrated that monolingual English infants, but not bilingual English-French infants, were able to segment words correctly from two overlapping languages. Finally, in Experiment 4, we attempted to examine whether individuals' inhibitory control skills relate to their segmentation performance. We did not find evidence that monolingual and bilingual infants differed in their inhibitory control skills. We also did not find a relationship between infants' segmentation performance and their inhibitory control skills, or that this relationship was influenced by infants' language background. In the aggregate, our findings suggest that phonetic cues across two languages (accents) can be effective for 9.5-month-olds to differentiate between two interleaved languages and successfully segment words from each when the cues involve salient non-native linguistic cues (i.e., English infants listening to French phonetic cues). Importantly, learners' success of segmenting words from two overlapping languages was not related to their inhibitory cognitive control. This suggests that early language experience, but not general cognitive ability, plays a key role in influencing how infants interpret language inputs and make use of linguistic cues to separate two overlapping languages.

Our paper is the first study that demonstrates infants at 9.5 months of age can separate two different language systems and accurately track transitional probabilities of the individual languages when they overlapped at syllable level. By showing that phonetic cues can facilitate dual-input statistical segmentation, we suggest that within-person acoustic cues are salient cues for infants to represent two languages separately in the process of word segmentation. This result

has an implication for infant word segmentation in a language mixing context. Language mixing is a common phenomenon in bilingual environments when speakers include two languages in one sentence, such as code-switching between two languages (Poplack, 1980) or borrow words from another language (Myers-Scotton, 1992). A common concern is that language mixing may pose a special challenge to infant word learning because young infants must determine which words are from which language in one sentence and they may have more difficulties to learn words in each language. For example, Byers-Heinlein (2013) found that higher rate of parental code-switching was associated with smaller vocabulary for children at age 1.5 years. Thus, language mixing may be particularly challenging for infants to segment words separately and accurately from each language. Yet, our study has suggested that infants can make use of the phonetic cues across languages to correctly segment words from two interleaved languages, providing support that infants can overcome the challenges in a language mixing context when learning words.

However, our findings are in stark contrast to findings in prior studies (e.g., Antovich & Graf Estes, 2018; Bartolotti, Marian, Schroeder, & Shook, 2011) where bilingual experience facilitated statistical word segmentation from two languages. We found that monolingual English infants, but not bilingual French-English infants, succeeded in utilizing phonetic cues to differentiate between two languages in a dual-input segmentation task. So, how can we reconcile the apparent difference between our findings and other prior studies?

One possibility is that bilingual infants may have impoverished phonetic representations of close contrasts in their two native languages as compared to their monolingual peers, leading to potential difficulties in detecting the differences in phonetic cues between English and French. Several studies have found that Spanish-Catalan bilingual infants have a temporary difficulty in discriminating certain native vowel contrasts. In a series of studies, Bosch & Sebastián-Gallés

(2003a) demonstrated that monolingual infants followed the usual patterns of perceptual narrowing of non-native phonetic contrasts and maintenance of native phonetic contrasts when discriminating a Catalan vowel contrast /e/-/ɛ/ that does not exist in Spanish. Monolingual Spanish infants showed a decline of discrimination of this non-native contrast whereas monolingual Catalan infants showed a maintenance of discrimination of this native contrast. Because this vowel contrast is a native phonetic contrast for bilingual infants, they should maintain their discrimination of this contrast over time. However, the bilingual infants demonstrated an unusual U shaped pattern such that they were sensitive to this contrast at 4 months and 12 months, but they failed to discriminate it at 8 months. This finding, along with other studies showing that Spanish-Catalan bilingual 8-month-old infants failed to discriminate other close contrasts, such as /o/-/u/ (Sebastián-Gallés & Bosch, 2009) and /s/-/z/ (Bosch & Sebastián-Gallés, 2003b), have suggested that bilingual infants between 8 months to 11 months may suffer a temporary decline in discrimination ability of certain contrasts in their native languages. Thus, our bilingual participants, who were 9.5-months-old, may not be able to discriminate certain close phonetic contrasts in French and/or English. If they cannot discriminate some phonetic cues between French and English, they will not make use of all possible phonetic cues to represent two statistical language structures separately and consequently fail to track transitional probability of each language structure.

However, there is counter-evidence to indicate that French-English bilingual infants may not have impoverished phonetic representations. First, prior studies have suggested that bilingual infants generally can discriminate sound contrasts in English and French quite well. For example, Burns and colleagues (2007) found that French-English bilinguals of 10-12 months of age could differentiate between English-specific and French-specific /p/-/b/ contrasts whereas English monolinguals could only differentiate the English-specific contrasts. Sundara, Polka and Molnar

(2008) also showed that French-English bilingual infants remain sensitive to both English and French /d/ sounds at the same age. These studies have demonstrated that French-English bilingual infants maintain sensitivity to sounds in both of their native languages and can even distinguish phonetic realizations of the same phoneme (i.e., English /d/ from French /d/). Further, the French- and English-accented speech in our study also possessed phonetic differences associated with the different rhythms of French and English (see Table 2). For example, on average, the French-accented speech had shorter duration and smaller intensity than the English-accented speech. Prior studies have shown that bilingual infants can discriminate two languages merely based on differences in rhythmic cues (Byers-Heinlein, Burns & Werker, 2010; Sundara & Scutellaro, 2011). Thus, the differences in rhythmic cues between French- and English-accented speech could provide infants with additional cues to differentiate two overlapping phonetic cues. Together, these studies have suggested that French-English bilingual infants are likely to be capable of differentiating the phonetic cues employed in the current experiment when separating the two language structures.

Another explanation is that bilingual experience may alter learners' attention and interpretation of the linguistic input. It is crucial to note that we tested French-English bilingual infants only with phonetic cues from their native languages, whereas we tested monolingual English infants with both native and non-native language phonetic cues. Our bilingual participants receive French and English phonetic variations on a regular basis in comparison to monolingual English infants, and are therefore highly familiar with French and English phonetic cues. In our experiment, the artificial language speech stream alternated between English- and French-accented speech. Due to their high familiarity to French and English accents, our bilingual participants may

not be surprised to hear two alternating French and English accents in one speech stream and therefore these phonetic cues may not greatly drive attention to the stimuli.

Additionally, the bilingual infants may have treated the cues as ambiguous to the separation of two languages due to experience with accented speech. For example, a bilingual infant could hear native-accented French from her father (his native language) and English-accented French from her mother (her second language). In this scenario, the base language would have the same statistical cues to word segmentation, in the face of differing phonetic cues. Thus, this may result in bilinguals treating the changes in phonetic cues as potential noise in the novel input and would not make use of the cues to detect changes in the underlying language structures. On the contrary, monolingual English infants are unfamiliar with French phonetic cues. They may be surprised to hear speech changing from native English-accented speech to a non-native French-accented speech, thus have increased attention to the latter. This increased attention would allow them to treat the differences in the phonetic cues as a signal of a change in language inputs, and successfully represented two languages separately.

Support for the argument that bilingual experience changes how infants treat accented speech comes from previous research that has demonstrated that bilingual experience improved infants' word recognition across different accented speech. A number of previous studies (e.g., Schmale & Seidl, 2009; Schmale, Cristia, Seidl & Johnson, 2010) have showed that 9-month-old monolingual English infants have difficulty recognizing familiar word forms in non-native accented speech (e.g., monolingual English infants at 9 months are not able to recognize the word "candle" when the word was produced by a non-native speaker with a Spanish accent). These findings have suggested that young monolingual infants are very sensitive to phonetic variations of their native language and they reject non-native accented speech as potential variations of their

own language environments (e.g., the word “candle” produced in Spanish accent is not the word “candle” in their native-accented language). However, bilingual infants were shown to be more open-minded and can accept subtle phonetic differences as an instance of the same language. Hudon (2013) found that 9-month-old bilingual French-English infants were capable of recognizing familiar word forms in both English- and French-accented speech whereas English monolinguals at the same age group failed to do so. If bilinguals are more open-minded when interpreting phonetic variations in language inputs, perhaps bilingual experience may lead infants to think that the two artificial languages in our study are instances of the same underlying language. For example, as the bilingual infants in our study were highly familiar with the English and French accents, they may think that the French- and English-accented speech in our experiment represent the same language, just like how English-accented French and native-accented French represent the same language (French) in reality. Consequently, bilingual infants may reduce their attention to these phonetic cues in our experiment, failing to separate the two artificial languages. This argument is speculative and further research with different accented cues is needed to elucidate these possibilities. For example, bilingual French-English infants may succeed with non-native accented speech (e.g., Mandarin accented speech). It is possible that bilingual infants could have heightened attention to non-native accented speech and may be better in using these non-native phonetic cues to segment two overlapping languages.

A third possibility is that bilingual may try to apply different strategies to segment words from different languages when they hear different phonetic cues. Previous studies have suggested that infants are highly sensitive to the prosodic properties of their native languages and will apply language-specific strategies to segment words from natural fluent speech. For example, English infants tend to segment words based on stress by treating stressed syllable as the onset of the words

(Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000; Jusczyk, Houston, & Newsome, 1999) whereas French infants tend to segment words based on syllable units by extracting each syllable, not extracting multi-syllabic words (Hohle, Bijeljac-Babic, Herold, Weissenborn & Nazzi, 2009; Nazzi, Iakimova, Bertoncini, Fredonie, Alcantara, 2006). Past research has shown that bilingual infants switch word segmentation strategies in accordance with language context (Polka, Orena, Sundara & Worrall, 2017). In our experiment, when bilingual infants heard the alternating French and English phonetic cues in the artificial language speech stream, they may think that they need to correspondingly switch their segmentation strategy. For example, they could be paying attention to the lexical stress when hearing the English speech stream and then paying attention to each syllable when hearing the French speech stream. Switching between segmentation strategies in our experiment would hinder infants' ability because they would be distracted from tracking the only informative cues in the speech stream – the transitional probability (i.e., there were no rhythmic cues to segmentation). The monolingual infants would not be switching strategies and could focus on the informative aspect of the stimuli, the statistical cues, rather than any uninformative strategy like stress.

In addition to the findings of word segmentation from two overlapping languages, our paper has also attempted to investigate whether infants' cognitive ability may be related to their segmentation performance. We did not find a correlation between individual's cognitive ability and segmentation performance, ruling out the possibility that a more effective cognitive system is the driving mechanism for infants to process information and segment two languages separately. However, this interpretation needs to be taken with a caveat. Infants in our study generally did not show learning of the associations between the tone-shape sequences and visual rewards in the post-switch phase, suggesting that our cognitive task was challenging for infants at 9.5 months. This

finding is unexpected because 7-month-old bilingual infants, who are less cognitively mature, succeeded in the same task in a previous study (Kovács & Mehler, 2009a). A possible account is related to the differences in methodology between our study and Kovács and Mehler (2009a). We reduced the number of trials in the experiment, which could mean that infants had insufficient training. However, the fact that infants successfully learned the associations in 6 trials in the pre-switch phase indicates that the number of trials were sufficient. In addition, Kovács and Mehler (2009a) used eye-tracker but we used behavioral oculomotor paradigm. This can be a less sensitive measurement of infants' eye-movements in the task. All these differences highlight the challenges of replicating an infant study that can be extremely age and paradigm-specific (Frank et al., 2017).

A possible solution would be simplifying the infant cognitive task and see if this will help reveal the underlying correlation between infants' cognitive task and their segmentation ability. For example, Reuter, Emberson, Romberg and Lew-Williams (2018) simplified the infant cognitive task and found a positive correlation between infants' cognitive skills and their vocabulary sizes. Importantly, Reuter and colleagues (2018) discovered that overall infants between 12 to 24 months successfully switched their responses during the post-switch phase by testing infants with a simple cognitive task. In their cognitive task, the authors simply showed infants a visual reward on the left side across the first 8 trials (i.e., the pre-switch phase) and then switched this visual reward to the other side for another 8 trials (i.e., the post-switch phase). This switching task is a much simplified task in comparison to our current task. It only required infants to switch their responses during the post-switch phase. In contrast, in addition to switching responses during the post-switch phase, our current task required infants to extract the patterns of the tone-shape sequences in both pre-switch and post-switch phases and associate different sequence patterns with different visual rewards. Our current task may be too cognitive demanding

to 9.5-month-old infants and thus infants failed to switch their responses during the post-switch phase. Future work perhaps needs to explore a range of age-appropriate and simplified cognitive tasks to fully address the question of a correlation between cognitive ability and segmentation performance in infancy.

Conclusion

In conclusion, the current paper has demonstrated that French and English phonetic cues can be effective indicators for successful infant statistical segmentation of two overlapping languages. Our findings have revealed that monolingual infants, but not bilingual infants, can successfully make use of the phonetic cues to separately represent the two languages when segmenting words from the languages. We conducted a follow-up study to examine whether individual cognitive abilities explain infants' segmentation performance but found that the success of the word segmentation is not related to the abilities tested. Instead, language experience, specifically bilingual experience, appears to impact how infants process information in our task. Bilingual infants may think that familiar French and English phonetic cues are possible cues within one language representation (i.e., the same language), thus failing to separate two language inputs when segmenting words. However, monolingual infants may treat any unfamiliar, non-native phonetic cue set as clearly indicating a different language input, and therefore they successfully represent each individual language and segment words from two language inputs. As such, the current study demonstrates that monolingual and bilingual experience influences how infants process phonetic cues differently when learning words.

References

- Altwater-Mackensen, N., & Mani, N. (2013). Word-form familiarity bootstraps infant speech segmentation. *Developmental Science*, 16(6), 980-990. doi: 10.1111/desc.12071
- Antovich, D. M., & Graf Estes, K. (2018). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science*, 21(2), e12548. doi: 10.1111/desc.12548
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321-324.
doi:10.1111/1467-9280.00063a
- Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117-134). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, 2, 324. doi: 10.3389/fpsyg.2011.00324
- Basseville, M., & Nikiforov, I. (1993). *Detection of abrupt changes – Theory and application*. Englewood Cliffs, NJ: Prentice-Hall.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1. doi: 10.18637/jss.v067.i01

- Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, 9(2), 117-121. doi: 10.1111/cdep.12116
- Bosch, L., & Sebastián-Gallés, N. (2003a). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46(2-3), 217-243. doi: 10.1177/00238309030460020801
- Bosch, L., & Sebastián-Gallés, N. (2003b). Language experience and the perception of a voicing contrast in fricatives: Infant and adult data. In *Proceedings of the 15th International Conference of Phonetic Sciences* (pp. 1987-1990). Barcelona, Spain: UAB/Casual Prods.
- Brent, M. R., & Siskind, J. Mark. (2001). The Role of Exposure to Isolated Words in Early Vocabulary Development. *Cognition*, 81(2), B33-B44. doi: 10.1016/S0010-0277(01)00122-6
- Brito, N., & Barr, R. (2012). Influence of bilingualism on memory generalization during infancy. *Developmental Science*, 15(6), 812-816. doi:10.1111/j.1467-7687.2012.1184.x
- Bulgarelli, F., Benitez, V., Saffran, J., Byers-Heinlein, K., & Weiss, D. J. (2017). Statistical learning of multiple structures by 8-month-old infants. In M. LaMendola & J. Scott (Eds.), *Proceedings of the 41st Annual Boston University Conference on Language Development* (pp. 128–139). Somerville, MA: Cascadilla Press.
- Burns, T., Yoshida, K., Hill, K., & Werker, J. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28(3), 455-474. doi:10.1017/S0142716407070257

- Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition*, 16(1), 32-48. doi: 10.1017/S1366728912000120
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343-348. doi: 10.1177/0956797609360758
- Cole, R., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–163). Hillsdale, NJ: Erlbaum.
- Erickson, L.C., Thiessen, E. D., & Graf Estes, K. (2014). Statistically coherent labels facilitate categorization in 8-month-olds. *Journal of Memory and Language*, 72, 49–58. doi: 10.1016/j.jml.2014.01.002
- Frank, M. C., Bergelson, E., Bergmann, C., Cristia, A., Floccia, C., Gervain, J., ... & Lew-Williams, C. (2017). A collaborative approach to infant research: Promoting reproducibility, best practices, and theory-building. *Infancy*, 22(4), 421-435. doi: 10.1111/infa.12182
- Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive science*, 33(6), 1087-1116. doi: 10.1111/j.1551-6709.2009.01041.x
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18(3), 254–260. doi: 10.1111/j.1467-9280.2007.01885.x

- Graf Estes, K., Gluck, S. C. W., & Bastos, C. (2015). Flexibility in Statistical Word Segmentation: Finding Words in Foreign Speech. *Language Learning and Development, 11*(3), 252-269. doi: 10.1080/15475441.2014.926730
- Graf Estes, K., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology, 51*(11), 1517. doi : 10.1037%2Fa0039725
- Hay, J., Pelucchi, B., Graf Estes, K., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology, 63*, 93–106. doi: 10.1016/j.cogpsych.2011.06.002
- Heredia, R. R., & Altarriba, J. (2001). Bilingual language mixing: Why do bilinguals code-switch? *Current Directions in Psychological Science, 10*(5), 164–168. doi: 10.1111/1467-8721.00140
- Hohle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: evidence from German and French infants. *Infant Behavior and Development, 32* (3), 262–274. doi:10.1016/j.infbeh.2009.03.004
- Houston, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review, 7* (3), 504–509. doi:10.3758/BF03214363
- Hudon, T., E. (2013). *Bilingual infants' accommodation of accented speech*. (Doctoral dissertation). Retrieved from:

https://www.ruor.uottawa.ca/bitstream/10393/26163/1/Hudon_Tamara_2013_thesis.pdf

Hunter, M.A., & Ames, E.W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in Infancy Research* (pp. 69–95). Stamford, CT: Ablex.

Huta, V. (2014). When to use hierarchical linear modeling. *The Quantitative Methods for Psychology*, 10(1), 13-28. doi: 10.20982/tqmp.10.1.p013

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

Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, 106(16), 6556-6560. doi: 10.1073/pnas.0811323106

Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science*, 325(5940), 611-612. doi: 10.1126/science.1173947

Kuznetsova, A., Brockhoff, P.B., Christensen, R. H. B. (2016). *lmerTest: Tests in Linear Mixed Effects Models*. R package version 2.0-29. Retrieved from: <http://CRAN.R-project.org/package=lmerTest>.

Mitchel, A. & Weiss, D.J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, 25(4), 456-482. doi: 10.1080/01690960903209888

Myers-Scotton, C. (1992). Comparing codeswitching and borrowing. *Journal of Multilingual & Multicultural Development*, 13(1-2), 19-39. doi: 10.1080/01434632.1992.9994481

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

Oakes, L.M., Sperka, D. J., & Cantrell, L. (2015). Habit 2. *Unpublished software*. Center for Mind and Brain, University of California, Davis.

Poepel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, 152, 9-19. doi: 10.1016/j.cognition.2016.03.001

Polka, L., Orena, A. J., Sundara, M., & Worrall, J. (2017). Segmenting words from fluent speech during infancy—challenges and opportunities in a bilingual context. *Developmental Science*, 20(1), e12419. doi: 10.1111/desc.12419

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

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

- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to Represent a Multi-Context Environment: More than Detecting Changes. *Frontiers in Psychology*, 3, 228. doi: 10.3389/fpsyg.2012.00228
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.
- Reuter, T., Emberson, L., Romberg, A., & Lew-Williams, C. (2018). Individual differences in nonverbal prediction and vocabulary size in infancy. *Cognition*, 176, 215-219. doi: 10.1016/j.cognition.2018.03.006
- Saffran, J. R. (2003). Statistical language learning mechanisms and constraints. *Current Directions in Psychological Science*, 12(4), 110–114. doi: 10.1111/1467-8721.01243
- 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
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it?. *Developmental Science*, 12(6), 874-887. doi: 10.1111/j.1467-7687.2009.00829.x
- Schmale, R., Cristia, A., Seidl, A., & Johnson, E. K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy*, 15(6), 650-662. doi: 10.1111/j.1532-7078.2010.00032.x

- Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, 12(4), 583-601. doi: 10.1111/j.1467-7687.2009.00809.x
- Singh, L., Fu, C., Rahman, A., Hameed, W., Sanmugam, S., Agarwal, P., . . . Rifkin-Graboi, A. (2015). Back to Basics: A Bilingual Advantage in Infant Visual Habituation. *Child Development*, 86(1), 294-302. doi: 10.1111/cdev.12271
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108(1), 232-242. doi: 10.1016/j.cognition.2007.12.013
- Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, 39(4), 505-513. doi: 10.1016/j.wocn.2010.08.006
- Thiessen, E. D., & Erickson, L. C. (2013). Discovering words in fluent speech: the contribution of two kinds of statistical information. *Frontiers in Psychology*, 3, 590. doi: 10.3389/fpsyg.2012.00590
- Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97(2), B25-B34. doi: 10.1016/j.cognition.2005.01.006
- Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, 5, 953. doi: 10.3389/fpsyg.2014.00953

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

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2010). Colliding cues in word segmentation: the role of cue strength and general cognitive processes. *Language and Cognitive Processes*, 25(3), 402–422. doi: 10.1080/01690960903212254

Weiss, D. J., Poepsel, T., & Gerfen, C. (2015). Tracking multiple inputs: The challenge of bilingual statistical learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages*, pp. 167-190. Amsterdam, the Netherlands: John Benjamins Publishing Company.

Table 1. *Auditory Stimuli in the Language Segmentation Task.*

Stimuli for Language 1	
Original in Weiss et al., (2009)	Modified stimuli in our study
Words	
bə-tɪ-gu	bə-tɪ-gu
sɪ-tʃə-vi	sɪ-tʃə-vi
vʊ-bo-sæ	vʊ-bo-sa
to-gʊ-tʃa	to-gʊ-tʃa
Part words	
sæ-to-gʊ	sa-to-gʊ
gu-vʊ-bo	gu-vʊ-bo
tʃa-sɪ-tʃə	tʃa-sɪ-tʃə
vi-bə-tɪ	vi-bə-tɪ
Stimuli for Language 2	
Words	
gu-pæ-tə	gu-pa-tə
dʒi-ga-pʊ	dʒi-ga-pʊ
sæ-dʒu-bo	sa-dʒu-bo
ta-bi-sɪ	ta-bi-sɪ
Part words	
bo-dʒi-ga	bo-dʒi-ga
sɪ-gu-pæ	sɪ-gu-pa
pʊ-ta-bi	pʊ-ta-bi
tə-sæ-dʒu	tə-sa-dʒu

Table 2. *Acoustic Measurement of the English- and French-Accented Speech.*

	<u>English- accented speech</u>		<u>French- accented speech</u>		<u>Paired t test analysis between English- and French-accented speech</u>
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	
Mean duration (ms)	247.65	89.56	207.63	55.30	$t(55) = 4.14, p < 0.001$
Mean fundamental frequency (Hz)	249.96	32.47	250.93	39.17	$t(55) = -0.14, p = 0.889$
Mean intensity (dB)	80.92	1.98	79.93	1.96	$t(55) = 3.46, p = 0.001$

Note. The mean and standard deviations of the acoustic measurements were computed across syllables.

Table 3. *Hierarchical Linear Regression Models.*

Parameters	<u>Model 1</u>			<u>Model 2</u>		
	β	<i>S.E.</i>	<i>p</i>	β	<i>S.E.</i>	<i>p</i>
Intercept	0.49	0.051	<0.001	0.34	0.063	<0.001
Number of Trials (Level 1)	0.02	0.008	0.012	0.032	0.012	0.009
Infants' language background (Level 2)	0.02	0	0.77	0.003	0.066	0.961
Infants' segmentation ability (Level 2)	-0.001	0.012	0.928	0.005	0.012	0.66
Number of trials X Infants' language background (Cross-level interaction)	0.02	0.019	0.28	-0.016	0.018	0.375

Note. Model 1 analyzed infants' performance during the pre-switch phase and Model 2 analyzed infants' performance during the post-switch phase.

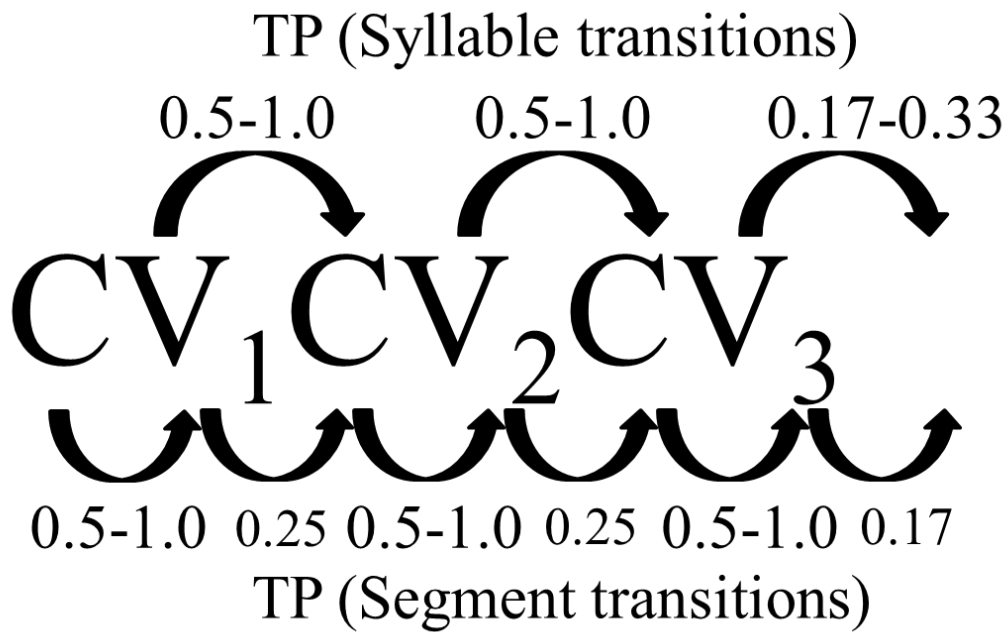


Figure 1. Transitional probabilities (transitional probability) of the underlying structures (i.e., inconsistent statistical cues) of the two languages when combined.

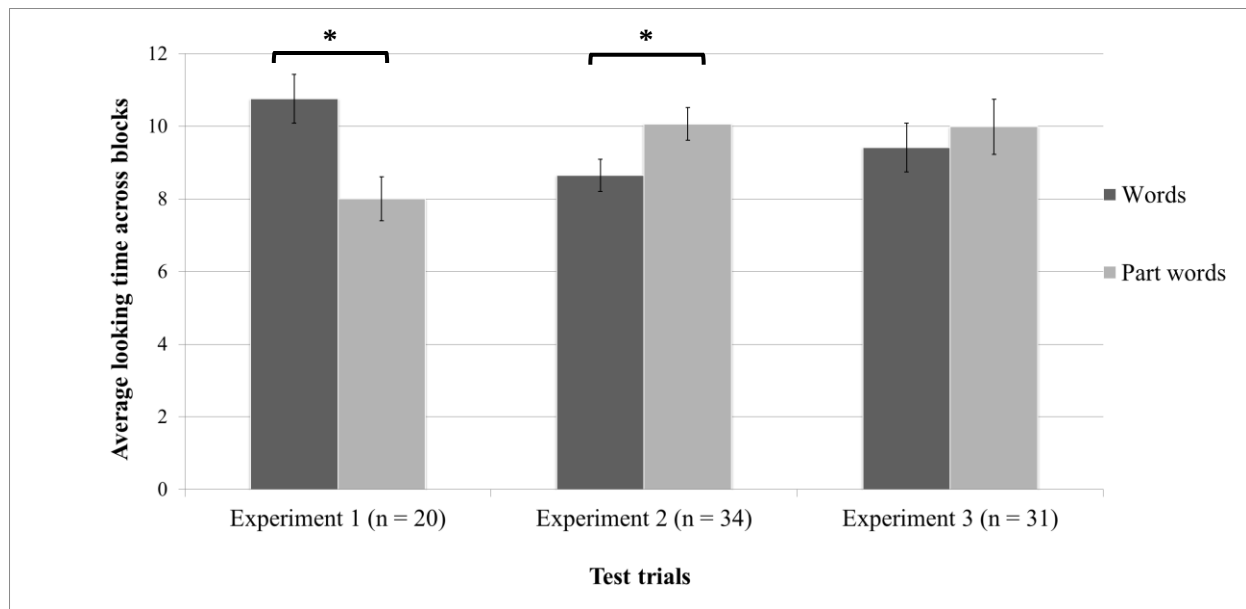


Figure 2. Infants' mean looking time to words and part-words across different experiments. Note that outliers were removed in Experiment 1 and Experiment 3, thus changing the number of infants entered in the analyses and those reported in this figure. Asterisk indicates significant difference ($p < 0.05$) between mean looking time to words and part-words in a particular test block. Error bar represents the standard error of the mean.

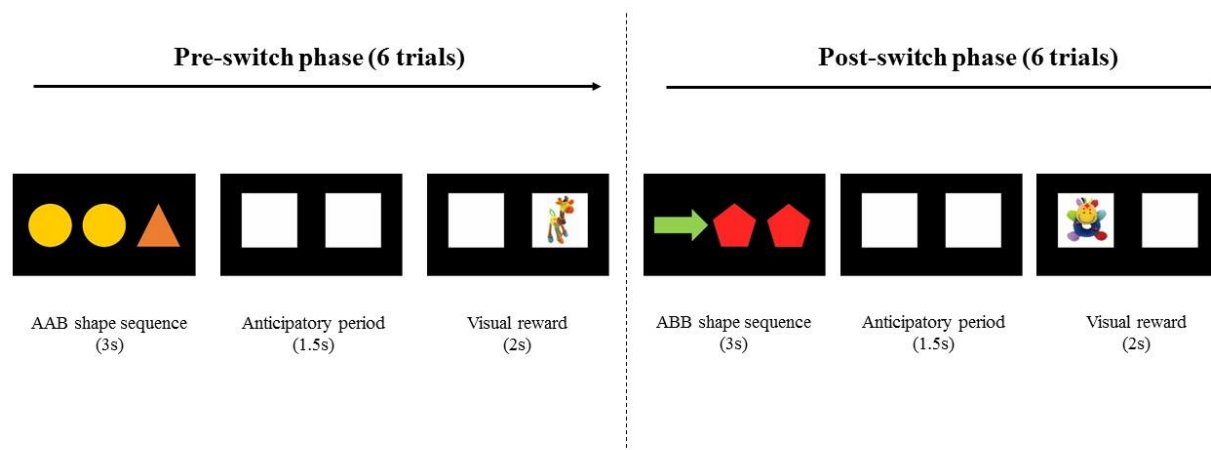


Figure 3. Procedure of the anticipatory eye movement paradigm in the experiment. Infants first entered the pre-switch phase. During the pre-switch phase, infants were trained that one shape sequence pattern (e.g., AAB) predicted a visual reward on the right. If infants learned the association between the shape pattern and the visual reward, they should begin to predict the location of the visual reward and look at the right side during the anticipatory period. After receiving the six trials of the pre-switch phase, infants entered the post-switch phase. During the post-switch phase, infants were trained with another association. This new association involved a pairing between a new shape pattern (e.g., ABB) and a new visual reward on the left. If infants were able to inhibit the previously learned association, they will predict the location of new visual reward and look at the left side during the anticipatory period in the post-switch phase. The shape patterns (AAB or ABB), locations of the visual rewards (right or left) and the pairings between shape patterns and visual rewards, were counterbalanced across infants.

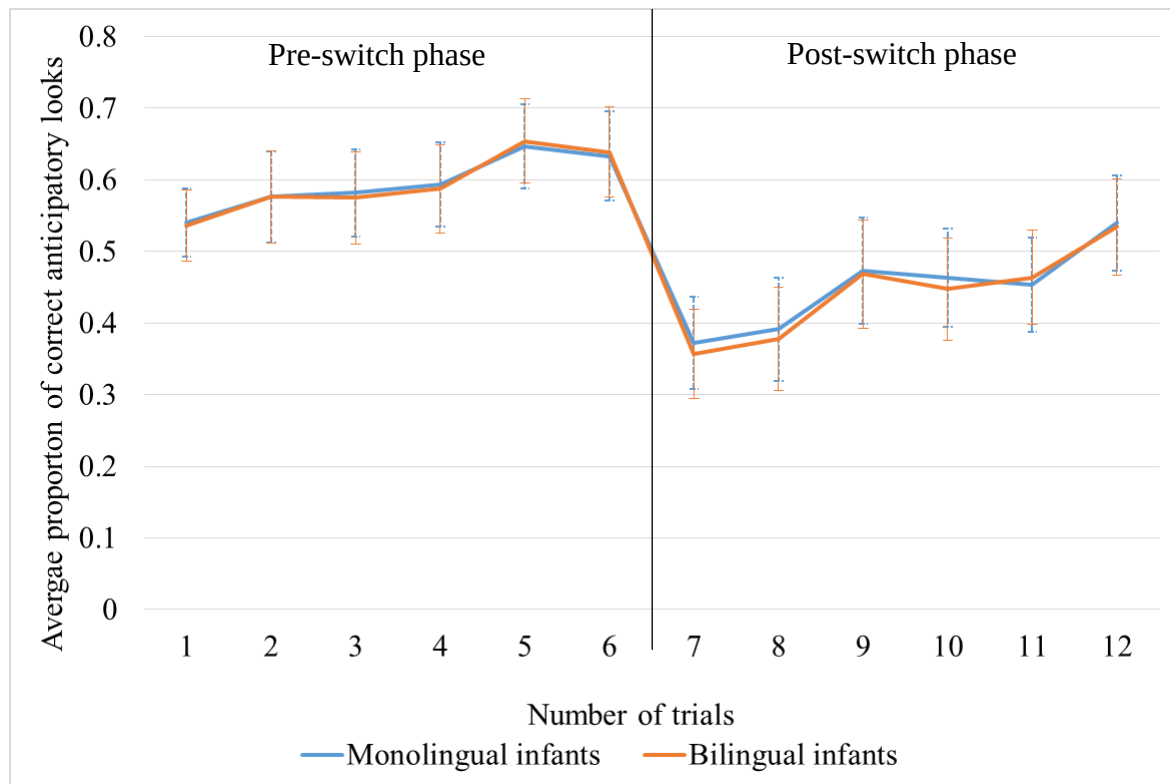


Figure 4. Infants' average proportion of correct anticipatory looks across trials during pre-switch and post-switch phases.

General Discussion

A fundamental step of word learning is to find words in the fluent speech stream. However, this is a daunting task because words are not reliably separated by clear and standard word boundaries, like pauses, in continuous speech. Statistical learning, a learning mechanism that allows learners to keep track of regularities in ambient environments, has been hypothesized as a possible mechanism that helps learners to segment words from fluent speech. A number of prior studies have suggested that learners can segment words from continuous speech on the basis of statistical regularity, such as syllable-level transitional probabilities in the language (e.g., Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996a). However, most studies to date have been focused on how learners keep track of transitional probabilities in order to segment words from a single language. Given the global prevalence of bilingualism (Grosjean, 2010), the present thesis aimed to examine whether learners can also track transitional probability across syllables when segmenting words from two different languages.

This thesis consisted of two large studies, involving multiple experiments each, to address a series of particular research questions arising from the above aim. These studies involve both adult and infant learners, to first model acquisition and then apply this research to the population who are in the process of establishing word segmentation skills. The first research question explored whether contextual cues within a single person's speech can facilitate word segmentation from two overlapping languages. In particular, I investigated whether phonetic cues associated with particular languages, which I term as accented speech, can facilitate segmentation of two overlapping languages. The second question examined whether bilinguals outperform monolinguals when they segment words from two languages. In addition, I explored whether factors, such as specific bilingual experience (i.e., age of acquisition) or cognitive ability, affected

bilinguals' performance in the segmentation task. In this section, I will discuss the results associated with each of the above research questions and address the resultant implications. Finally, I will discuss the limitations of the current thesis and propose how future studies can address these limitations.

Can Phonetic Cues (i.e., Accents) help Learners Segment Words from Two Overlapping Languages?

The first goal of this thesis was to examine whether learners can make use of contextual cues within a single individual's speech to segment words from two overlapping languages. This goal builds upon research showing that adults learned best and successfully segmented two overlapping language when they were provided with salient contextual cues, such as different gendered voices (Weiss, Gerken, & Mitchel, 2009) and different visual faces (Mitchel & Weiss, 2010). However, these previous studies have mainly focused on speaker-specific contextual cues in which the two languages were always spoken by two different individuals, thus limiting our understanding of how learners segment two languages from a single individual's speech. Indeed, exploring whether learners can segment words from two languages from a single individual's speech is an important question in bilingual research, as it would be the typical learning scenario for infant bilinguals. It would be extremely unnatural for one parent to only speak French and never English and the other to speak English and never French. Bilingual individuals necessarily speak two languages. Beyond this, bilingual infants often need to learn words in a language mixing environment where speakers code switch between two languages (Poplack, 1980) or borrow words from another language (Myers-Scotton, 1992) in one conversation. Thus, I extended the literature by exploring how learners make use of cues that present in naturalistic bilingual environments to segment words from two overlapping languages.

In the present thesis, I examined whether adults and infants can make use of the differences in phonetic cues (accents) to represent two different language inputs. This would, in turn, help the learners to track the transitional probabilities within the individual languages, leading to correct word segmentation in each language. I presented participants with a statistical segmentation task involving two overlapping languages in which the differences between languages were marked by French and English accents. Overall, the results in Study 1 and Study 2 indicated that adults and infants are able to segment words from two overlapping languages, supporting my hypothesis that accented speech can be effective in helping learners to segment two overlapping languages. However, it is important to note that the efficacy of phonetic cues for language separation varies across presentation conditions and language backgrounds.

Beginning with the variation discovered across presentation conditions, adults in Study 1 failed to segment words from one of the languages (namely AL2) when this language was presented in a French accent. I hypothesized that the phonetic properties of this particular language in the French accent condition may drive these difficulties. Follow-up acoustic analysis has suggested that participants could have difficulty in differentiating between two French-realized phonemes - /b/ and /p/. Failure in differentiating the /b/ and /p/ phonemes in the language would lead participants to confuse certain syllables, resulting in incorrect computations of transitional probability in that language and failing to segment words accurately. This finding shows that certain phonetic realizations pose additional challenges for learners to segment words accurately. However, it is possible that learners may be able to overcome this perceptual challenge if they receive an extensive exposure to this language (i.e., AL2 in a French accent). For example, Gebhart, Aslin and Newport (2009) found that participants were able to learn a more difficult language when this language was tripled in length. Thus, future research can explore whether a longer

exposure to the language may help learners to overcome any possible perceptual difficulties arising from the phonetic realizations and therefore succeed in word segmentation. Although participants failed to segment AL2 when it was presented in a French accent, participants in Study 1 still succeeded in segmenting words from the other language (AL1) in the face of the difficulties with AL2. Thus, I can conclude that accented speech is a useful tool for learners to separately represent two overlapping languages, as the statistics from the second language were not interfering with the first.

The findings that learners can segment words from a single individual's speech have an implication of whether the one-parent-one-language approach is a necessary approach for successful bilingual language acquisition. The one-parent-one-language approach suggests that each parent should only stick to one language when speaking to the bilingual children (Ronjat, 1913). For example, father only speaks French to the infant and mother only speaks English to the infant. The logic of the approach is that if each language is consistently associated with a particular person, this will prevent children from being confused about the two languages and help children to differentiate between two languages. However, past studies casted doubt on whether this one-parent-one-language approach is necessary for children to distinguish between two language inputs. For example, children whose parents often mixed two languages at home successfully learned the two languages (De Houwer, 2007; Place & Hoff, 2011). Other researchers have shown that even bilingual children mix their languages in their conversations, these children were able to speak the correct language (e.g., English) in accordance with the interlocutor's language background (e.g., monolingual English speaker) (Genesee, Nicoladis, Paradis, 1995; Lanza, 1992). These behaviours clearly demonstrate that young bilingual children are highly sensitive to the context and speakers' language background when using their languages. The fact that young children can effectively use

their languages to interact with different speakers reflects that they are capable of differentiating between their two languages. Together, these studies provide evidence that children can successfully learn two languages even though their parents mix two languages often in their environments, supporting the idea that one-parent-one-language approach is not a necessary approach for young bilingual children's language acquisition. However, these studies are indirect evidence because they are based on observational data. In contrast, my thesis provided direct evidence. I discovered that both adults and infants were capable of finding words in a simulated language-mixing environment where one individual speaker alternated two languages in a continuous speech stream. As such, the current thesis is the first to show that adults and infants can rely on subtle cues (i.e., phonetic cues) to separate two languages with conflicting language structures. This experimental evidence shows that learners can cope with the challenges of language differentiation in a language-mixing bilingual environment.

Do Bilinguals Outperform Monolinguals when Segmenting Words from Two Overlapping Languages? If so, What Explains this Bilingual Advantage?

The second goal of the present thesis was to investigate whether bilingual speakers performed better in the segmentation task. Prior findings have shown that adult and infant bilinguals outperform monolinguals in several statistical learning tasks (e.g., Antovich & Graf Estes, 2018; Bartolotti, Marian, Schroeder, & Shook, 2011; Wang & Saffran, 2014), it is thus possible that bilingual speakers may perform better when segmenting words from two overlapping languages.

Overall, I did find that bilinguals and monolinguals performed differently in the dual-input segmentation task in adulthood and infancy, however bilinguals did not necessarily outperform monolinguals across the lifespan. In Study 1, I found that bilingual adults, specifically

simultaneous bilinguals, segmented words better than monolingual adults. On the contrary, in Study 2, monolingual infants, but not bilingual infants, successfully segmented words from two overlapping languages. Interestingly, as hinted above, the age of acquiring the second language plays a significant role in bilinguals' performance in the segmentation task. For example, in Study 1, I only found that simultaneous bilinguals who learned two languages since birth segmented more correct words than monolinguals. Sequential bilinguals who learned their second language after age 1 did not differ from monolinguals in terms of the segmentation accuracy. Interestingly, my finding in Study 2 appears to directly contrast with the finding in Study 1 as monolingual infants performed better than simultaneous bilingual infants. I will return to this point later to give potential explanations as to why bilingual infants performed poorer than monolinguals in the segmentation task. But first, I will discuss possible reasons as to why age of acquisition influenced adult learners' segmentation performance.

I suggest that acquiring the second language at an early age may alter how learners interpret the variability of sensory inputs in new environments. In this thesis, I have emphasized that learners' expectations of input changes are the key to success in a dual-language segmentation task (Basseville & Nikiforov, 1993; see Gebhart, Aslin & Newport, 2009; Qian, Jaeger, & Aslin, 2012 for reviews). When the two languages share common syllables, the transitional probabilities between syllables would be largely incompatible across the two languages and thus learners would receive unstable statistical information. Let us consider an earlier example in the thesis, *moto*, in this situation. *Moto* is a word in French and the transitional probability between these two syllables is therefore high in that language, whereas *moto* is not a word in English and the transitional probability between these two syllables is low in that language. Thus, the transitional probability varies widely when the language input changes from French to English. When learners receive

unstable statistical cues, they need to interpret whether such variability is noise or if it indicates a change in the underlying structure of the inputs. When contextual cues are provided in the speech stream, learners have a greater chance to separately represent the two language inputs and treat the variability as a change in underlying structure. After separating the two language inputs, learners would start tracking transitional probability from each individual language and eventually succeed in segmenting words from two overlapping languages.

For simultaneous bilinguals, they need to differentiate between two parallel language inputs before they can establish full knowledge of their languages. In order to learn two different languages successfully, simultaneous bilinguals must be flexible when interpreting variability in language inputs and they need to form an expectation that the underlying structure of the language input may change. For example, a bilingual parent may mix two languages together when speaking to her infant. When a bilingual mom mixes two languages together when speaking to her infant, this infant needs to be flexible when interpreting the input variability in this conversation. This infant must be open to the possibility that two input sets of statistics are in the input and open to the cues that may mark those set (e.g., phonetic sound variability across languages). If this infant treats the input variability as simple noise, she would not be able to separately represent two languages. As such, it is possible that simultaneous bilinguals may have developed an expectation that new environments could be subject to change, leading them to be more flexible when interpreting variability in sensory inputs in novel language environments. In addition, simultaneous bilinguals may be well aware to look for subtle contextual cues (e.g., phonetic cues across languages, rhythm of the languages) when they need to differentiate between the two languages. They may be quick to identify these contextual cues as markers of two language inputs, and therefore better at statistically segmenting words from two language inputs.

However, it is important to note that the benefit of early bilingualism is different across adults and infants. So how can we reconcile the observation that there is a simultaneous bilingual advantage in statistical segmentation in adults but not infants? One possibility is that bilingual infants are in the midst of learning when phonetic cues matter in their language environments. Phonetic cues in the bilingual language environments do not always signal a change of language inputs. For example, when a single speaker alternates between two languages within one conversation, phonetic cues signal a change of languages in the conversation. In contrast, in a situation where a group of multiple speakers were involved in one conversation, a change of phonetic cues may not signal a change in languages. Consider a conversation where there is a mix of Anglophone and Francophone speakers, an Anglophone speaker may speak French with an English accent whereas other Francophone speaker speak French with a French accent, but all speakers speak in one language, thus French and English phonetic cues in this language context do not signal a change in languages. For simultaneous adult bilinguals, they have extensive experience in communicating with other bilingual speakers, including within a language-mixing context (i.e., they often hear two languages mixing together within a conversation). Thus, they would be experienced in handling the variability of phonetic cues in their language environments. This experience arguably helps bilingual adults to identify the context where phonetic cues signal a change in language inputs. Here, I argue that extensive bilingual experience may help adults to note that phonetic cues are important cues in my experiment and thus the adult learners utilize phonetic cues when segmenting words from two languages.

However, young bilingual infants may not have sufficient experience to infer when phonetic cues matter in their language environments, and therefore they mistakenly interpret phonetic cues as non-meaningful signals for a change of language inputs. Indeed, prior studies

reported that young bilingual infants may suffer temporary confusion when determining meaningful phonological cues across different language contexts. For example, Bosch and Sebastián-Gallés (2003a) tested Spanish-Catalan bilingual infants with a close vowel contrast, /e/-/ɛ/, that is only meaningful in Catalan, but not Spanish. While 4- and 12-month-old bilinguals successfully discriminated the vowel contrast, bilinguals at 8 months failed (see similar results for a close consonant contrasts /s/-/z/ in Bosch & Sebastián-Gallés, 2003b). Interestingly, this U-shaped developmental pattern was also found in 8-month-old bilingual infants when they were tested with an acoustically close vowel contrast that exists in both Spanish and Catalan (Sebastián-Gallés & Bosch, 2009). Taken together, these studies have demonstrated that bilingual infants around 8 months of age may have a temporary confusion of acoustically close contrasts. The hypothesized explanation for this confusion is that bilinguals are trying to determine which close phonological cues are meaningful within and between language contexts (i.e., Spanish and Catalan in this case).

Other research has also reported that young bilingual infants have difficulties determining when a phonemic cue is meaningful in their environments when said cue is incompatible across the two languages. For example, Singh and Foong (2012) reported that young bilingual Mandarin-English infants were confused when a change of tones reflected a change of word meaning across Mandarin and English contexts. Tone variations in Mandarin signal changes in word meanings whereas tone variations in English did not reflect a change in word meanings. The syllable “ma” produced in a higher tone (i.e., meaning “mother”) or in dipping tone (i.e., meaning “horse”) represents two different words in Mandarin, but it would represent the same word in English (i.e. a diminutive form of the word mother: “Ma.”). Successful English-Mandarin bilinguals need to demonstrate selective sensitivity to tones in one language context (Mandarin) but not the other

(English). Singh and Foong (2012) found that bilingual Mandarin-English infants of 7.5 months demonstrated heightened sensitivity to tones regardless of language context (both Mandarin and English). Further, bilingual infants of 9 months did not show any sensitivity to tones in either Mandarin or English language contexts. It is not until 11 months of age that bilingual infants showed selective sensitivity to tones in different language contexts. That is, 11-month-old bilingual infants were sensitive to tones in Mandarin context but not in English context. This study suggests that bilingual infants between 7.5 to 10 months of age may not have sufficient linguistic experience to correctly determine meaningful cues in different language contexts, but this ability appears to improve by their first birthday.

Of course, the bilingual infants tested in my experiments were at 9.5 months of age. It is possible that these young bilinguals do not have sufficient linguistic experience to determine when phonetic cues are meaningful cues for language differentiation. In my experiment, young bilingual infants may be unsure when phonetic cues signal a change of language inputs and may not be able to determine the language context they are in (i.e., a fluent speaker alternating between two languages). Thus, they did not treat the change of phonetic cues as meaningful for language differentiation. This argument suggests that when older bilingual infants gain more linguistic experience, they would succeed in segmenting words from two overlapping languages in my experiment. Future work could explore this possibility by testing both younger and older bilingual infants to examine the developmental transition where bilingual infants start correctly determining that within-speaker phonetic cues are meaningful in the dual-input segmentation task.

Implication of the Relationship Between Cognitive Ability and Statistical Segmentation

Another central finding in my thesis is that the difference in segmentation performance between bilinguals and monolinguals is not related to the individuals' cognitive ability, at least

those tested in our tasks. This rules out the possibility that bilinguals' better performance in statistical segmentation (in Study 1) was driven by enhanced executive processing skills.

First, bilinguals in our studies did not differ from monolinguals in terms of their executive processing skills, which is consistent to recent studies reporting no bilingual cognitive advantages (e.g., Gathercole et al., 2014; Kousaie & Philips, 2012; Paap & Greenberg, 2013; Paap & Sawi, 2014). However, it is important to note that the null effect in bilingual cognitive advantage in the current thesis may be related to the bilingual contexts present in the Ottawa area. According to the Adaptive Control Hypothesis (Green & Abutalebi, 2013), different interactional bilingual contexts place different demands on language control, which in turn adaptively change bilinguals' cognitive control ability. They identified three different bilingual contexts: dense code-switching, single-language, and dual-language. The first two of these contexts implicate few cognitive demands, but the last involves heavy switching demands.

In a dense code-switching context, bilinguals can routinely switch between two languages within an utterance due to the presence of other bilinguals. For instance, a bilingual mom may say to her bilingual children "*Regarde! Look at le chien!*" In this context, bilinguals can freely use their two languages as all speakers in this context are assumed to understand these two languages. Thus, the language control demand in the dense code-switching context is low because bilinguals do not need to inhibit any languages at all. In a single-language context, bilinguals only speak one language in one environment and the other language in another environment (e.g., speaking French only at home and speaking English only at work). As there is no switching between two languages in either language environment, bilinguals living in a single-language context simply need to inhibit the irrelevant language for successful communication.

Green and Abutalebi (2013) proposed that dual-language context has the highest switching demands. In the dual-language context, bilinguals would use both of their languages in one environment (e.g., work). However, bilinguals use their languages with different speakers such that bilinguals must be careful in selecting the target language for successful communication (e.g., speak English only to an Anglophone co-worker but speak French only to a Francophone co-worker). Therefore, switching between languages would not occur within an utterance because the bilinguals are restricted by interlocutors in their use of the target language. In addition to inhibiting the irrelevant language, bilinguals must actively monitor which language is the target language according to the interlocutor's language background. Thus, the language control demand in the dual-language context is arguably the highest among all three contexts listed in Green and Abutalebi (2013). Bilinguals who primarily live in the dual-language context are therefore most likely to have developed enhanced cognitive ability. In the case of Ottawa, Vinerte and Sabourin (2015) have suggested that the local bilingual context is one of dense code-switching, as bilinguals frequently mix and code-switch between two languages within one environment. Therefore, it is possible that the bilingual participants in the current thesis do not live in a linguistic context that encourages the development of the enhanced cognitive abilities seen in some other work.

The lack of correlations between statistical learning and executive processing skills across both monolinguals and bilinguals seems to challenge the Extraction and Integration Framework (Thiessen et al., 2013). The Extraction and Integration Framework predicts that working memory and attention are the underlying processes for extracting words from fluent speech, thus we should see positive relationships between learners' working memory, attention and their segmentation performance. One possible explanation for the lack of correlations in the current thesis is that working memory and attention are necessary, but not sufficient, for the extraction processes of the

word segmentation task. Perhaps working memory and attention are important to help learners extract potential words from the speech stream, but other cognitive processes (e.g., transferring newly extracted words to long-term memory storage) are needed to represent these words in the lexicon. However, the way that we measured learners' segmentation performance cannot capture the sub-components of the whole process of word segmentation. For example, we measured adults' accuracy of identifying words in Study 1 and infants' ability to discriminate words from part-words in Study 2. These measures only provide one single score that reflects all these cognitive processes in the word segmentation task. As such, these measures may not be able to reflect how working memory and attention individually contribute in the process of word segmentation, resulting in a lack of relationships between working memory, attention and segmentation performance. Future work is needed to develop a more sensitive measure that examines participants' responses throughout the statistical learning process. Such work can provide a more comprehensive picture of the links between different cognitive skills and the learning process of statistical learning, thus offering a better test of the predictions generated by the Extraction and Integration Framework.

Limitations and Future Directions

It is important to note that bilingualism is a continuum, not a simple binary variable. Although I have tried to rigorously measure different aspects of bilingualism in the participants within this thesis, there are some aspects that I was not able to control and/or measure. For example, infants were classified as bilingual merely based on parental reports. I conducted an interview with the parent by asking the parent to estimate their and other's language output to the infant, such as estimates of the number of hours that the infant was listening to French or English across different contexts and speakers. This self-reported measure may not accurately reflect infants' actual

exposure to French and English. It is possible that parents may have over- or under-estimated infants' exposure to the particular languages. To address this problem, additional work is needed to investigate the correlation between parental report and the actual infants' exposure to different languages. For example, researchers can record infants' language input in daily bilingual environments via a long-battery recorder, and later compute the correlation between the infants' actual exposure to the two languages and the parental report. These studies can shed light on whether the use of parental report can provide researchers with a good estimate of bilingual infants' actual language exposure. This could in turn help researchers to develop more fine-grained interview questions for measuring infants' language exposure in the parental report.

Moreover, the parental interview does not provide a measure of how often bilingual infants are exposed to a language-mixing environment: there is no question related to mixing on the form. Measuring whether the bilingual infants are learning in a language-mixing environment may provide a better understanding of why bilingual infants do not efficiently process phonetic cues as indicators of two separate languages. Future work can include a questionnaire about language mixing of infants' daily language environments (Byers-Heinlein, 2013) when conducting parental interviews. This questionnaire can provide additional information about the nature of infants' bilingual environments (e.g., whether they often hear parents mixing the two languages when parents speak to the infants), offering a more sensitive measure of bilingual infants' language environments.

The difficulty of measuring bilingualism is not the only limitation. Indeed, the nature of monolingual population may also be a confounding factor in the study. It is easier for researchers to identify purely monolingual infants because their primary language input comes from their parents. If the home is monolingual, this clearly suggests that the infant's exposure to the second

language is limited. However, controlling for the exposure of second language is very difficult in adult population. For adult participants, their language learning and exposure history can be much more complex. In Canada, a majority of adult learners have been exposed to or learned French across all of their schooling. Of course, the degree of exposure and frequency of using the second language vary substantially across these adult learners. However, they nevertheless have been exposed to French at some points in their life. Further, living in Ottawa where there is a strong emphasis of both French and English in the community, it is highly possible that the adult participants who are classified as English monolinguals are “contaminated” by some consistent exposure to French. Thus, they are not purely English monolingual participants. The “impurity” of monolingual English speakers may limit the generalization of my findings. In particular, one of the major findings is that monolinguals can also make use of French and English phonetic cues to separate and segment words from two overlapping languages. However, my monolingual adult speakers may have some familiarity to the French phonetic cues and this familiarity may facilitate their segmentation performance. Therefore, I may only be able to conclude that familiar phonetic cues would be effective for word segmentation in dual-input language task and whether unfamiliar phonetic cues are equally effective is an unknown question.

A possible way to address the above question is to test the participants in the current thesis with unfamiliar phonetic cues. For example, we can test monolingual English participants with Cantonese phonetic cues, which are not consistently present in their environment. If the monolingual English participants can make use of the Cantonese phonetic cues to segment words from two overlapping languages, it means that unfamiliar phonetic cues are also effective in facilitating word segmentation. We could also test both simultaneous and sequential bilingual English-French participants with Cantonese phonetic cues to determine if simultaneous bilingual

English-French participants still outperform sequential French-English (and monolingual English) participants. If we use Cantonese phonetic cues in the task and still find a simultaneous bilingual advantage in segmentation, this evidence will be a confirmation of my proposition that early bilingual experience during infancy may help learners to be more flexible in interpreting input variability and thus facilitate segmentation of two overlapping languages. Together, the proposed future studies can provide a fuller picture of whether monolingual and bilingual speakers can make use of phonetic cues to separate and segment words from two overlapping languages.

Concluding Statement

The findings of the current thesis expand our understanding of how learners statistically segment words from two overlapping language in a bilingual environment. Both adults and infants are capable of utilizing phonetic cues within a single individual's speech to segment words from each language successfully. This suggests that learners generally can rely on naturalistic phonetic cues to differentiate between two conflicting language inputs within a single person's speech stream. One implication of this result is that learners can overcome the challenge of learning two languages in a language-mixing context, and thus the one-parent-one-language strategy is unnecessary for young children to learn two languages successfully. In addition, other findings have suggested that bilinguals performed differently in comparison to monolinguals. Crucially, this difference is not driven by individual's cognitive abilities (at least those tested in the thesis), but appear to be related to early bilingualism. In Study 1, only simultaneous bilingual adults who learned two languages since birth performed differently from monolingual adults and other bilingual adults. This result lends support to the hypothesis that early bilingual experience may have prepared learners to be more flexible when processing information in new environments, thus helping these learners to detect subtle changes of the underlying structures of language inputs. In

contrast, monolingual infants in Study 2 surprisingly performed better than bilingual infants when segmenting words from two languages. Monolingual infants, but not bilingual infants, succeeded in utilizing phonetic cues to segment two overlapping language. It is proposed that bilingual infants during the first year of life may be in the process of learning when phonetic cues matter across different language contexts. Consequently, they do not make efficient use of the phonetic cues to separate two language inputs in the segmentation task. It is possible that longer bilingual experience may help learners to correctly identify the language context they are in and determine when phonetic cues are meaningful in differentiating languages in the environments. Future work therefore needs to investigate whether older bilingual infants can successfully make use of phonetic cues to segment words from two language inputs. Together, the studies presented in the current thesis add to the literature on dual-input statistical learning by demonstrating that learners can rely on subtle phonetic cues to segment words from two overlapping languages and early bilingualism may facilitate segmentation performance.

References

- Antovich, D. M., & Graf Estes, K. (2018). Learning across languages: Bilingual experience supports dual language statistical word segmentation. *Developmental Science*, 21(2), e12548. doi: 10.1111/desc.12548
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321-324). doi:10.1111/1467-9280.00063a
- Basseville, M., & Nikiforov, I. (1993). *Detection of abrupt changes – Theory and application*. Englewood Cliffs, NJ: Prentice-Hall.

- Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, 2, 324. doi: 10.3389/fpsyg.2011.00324
- Bosch, L., & Sebastián-Gallés, N. (2003a). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46(2-3), 217-243. doi: 10.1177/00238309030460020801
- Bosch, L., & Sebastián-Gallés, N. (2003b). Language experience and the perception of a voicing contrast in fricatives: Infant and adult data. In *Proceedings of the 15th International Conference of Phonetic Sciences* (pp. 1987-1990). Barcelona, Spain: UAB/Casual Prods.
- Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition*, 16(1), 32-48. doi: 10.1017/S1366728912000120
- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, 28(03), 411-424. doi:10.1017/S0142716407070221
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., ... Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, 5, 11. doi: 10.3389/fpsyg.2014.00011

- Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science*, 33(6), 1087-1116. doi: 10.1111/j.1551-6709.2009.01041.x
- Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, 22(3), 611-631. doi: 10.1017/S0305000900009971
- Grosjean, F. (2010). *Bilingualism: Life and reality*. Cambridge, MA: Harvard University Press.
- Kousaie, & Phillips. (2012). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research*, 1446, 71-90. doi: 10.1016/j.brainres.2012.01.052
- Lanza, E. (1992). Can bilingual two-year-olds code-switch? *Journal of Child Language*, 19(3), 633-658. doi: 10.1017/S0305000900011600
- Mitchel, A. & Weiss, D.J. (2010). What's in a face? Visual contributions to speech segmentation. *Language and Cognitive Processes*, 25(4), 456-482. doi: 10.1080/01690960903209888
- Myers-Scotton, C. (1992). Comparing codeswitching and borrowing. *Journal of Multilingual & Multicultural Development*, 13(1-2), 19-39. doi: 10.1080/01434632.1992.9994481
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232-258. doi: 10.1016/j.cogpsych.2012.12.002

- Paap, K., & Sawi, O. (2014). Bilingual advantages in executive functioning: Problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Frontiers in Psychology*, 5, 962. doi: 10.3389/fpsyg.2014.00962
- Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Development*, 82(6), 1834–1849. doi:10.1111/j.1467-8624.2011.01660.x
- 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
- Qian, T., Jaeger, T. F., & Aslin, R. N. (2012). Learning to Represent a Multi-Context Environment: More than Detecting Changes. *Frontiers in Psychology*, 3, 228. doi: 10.3389/fpsyg.2012.00228
- Ronjat, J. (1913). *Le développement du langage observé chez un enfant bilingue*. Paris: Champion.
- 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
- Sebastián-Gallés, N., & Bosch, L. (2009). Developmental shift in the discrimination of vowel contrasts in bilingual infants: Is the distributional account all there is to it?. *Developmental Science*, 12(6), 874-887. doi: 10.1111/j.1467-7687.2009.00829.x
- Singh, L., & Foong, J. (2012). Influences of lexical tone and pitch on word recognition in bilingual infants. *Cognition*, 124(2), 128-142. doi: 10.1016/j.cognition.2012.05.008

- Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin*, 139(4), 792–814. doi: 10.1037/a0030801
- Vinerte S., & Sabourin, L. (2015). Bilingualism and cognitive control: The ANT in a Canadian context. *Proceedings of the Annual Conference of the Canadian Linguistics Association*, Ottawa, Canada, 1-15.
- Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, 5, 953. doi: 10.3389/fpsyg.2014.00953
- Weiss, D.J., Gerfen, C. & Mitchel, A.D. (2009) Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development* 5, 30-49. doi: 10.1080/15475440802340101

Appendix A: Adult Language Background Questionnaire

Language Background Questionnaire - Short Version



Participant information – To be filled out by the researcher

Project code: _____ Today's date: _____ Participant code: _____

1. Biographical information

Month and year of birth: _____ Current Age: _____ Gender: _____

Your native language(s)? _____

Mother's native language(s)? _____

Father's native language(s)? _____

Other caregivers' native language(s)? _____

Are you left- or right-handed? _____

Do you have any known visual impairments (including colour-blindness)? _____

Do you have any known hearing impairments? _____

Have you ever had a serious head injury? _____

What is your current profession? _____

Place of birth (city, country): _____

Please list all the places where you have lived, when you lived there, and for how long. Please list these in chronological order. _____

2. Basic language information

For an example, see page 3. Please ask if you have any questions.

In the **first row**, please list all the spoken or signed languages of which you have ANY current or previous knowledge in order of **DOMINANCE**, i.e. how comfortable you are using them.

In the **second row**, please write the age at which you were **first exposed** to this language, i.e. when it first became present in your environment.

In the **third row**, please indicate whether you were ever **IMMERSED** in this language, i.e. if you received significant exposure to this language in either a community/home setting **or** a school (greater than 20%). If so, please also indicate the age or grade at which this began.

Dominance	Most ← → Least				
Language:	1.	2.	3.	4.	5.
Age/grade of 1 st exposure:					
Immersed:	YES / NO	YES / NO	YES / NO	YES / NO	YES / NO
Age/grade:	_____	_____	_____	_____	_____

Language Background Questionnaire - Short Version

For each of the above-listed languages, please provide your current and highest ever attained level of proficiency. This consists of your **general** proficiency. If you feel that you are between categories, simply indicate where along the scale you feel your proficiency is best represented.

Language: _____	Current level of general proficiency
_____	Highest level of proficiency ever attained

Language: _____	Current level of general proficiency
_____	Highest level of proficiency ever attained

Language: _____	Current level of general proficiency
_____	Highest level of proficiency ever attained

Language: _____	Current level of general proficiency
_____	Highest level of proficiency ever attained

Language: _____	Current level of general proficiency
_____	Highest level of proficiency ever attained

Language Background Questionnaire - Short Version

EXAMPLE 1: Julia's mother is a native speaker of Spanish, and her father is a native speaker of English. However, Julia's mother passed away when she was four years old, and Julia therefore did not get much exposure to Spanish after that age. Julia spoke Spanish natively until the age of four (highest level of proficiency ever attained) and considers her current level of proficiency to be advanced. Her current and highest ever level of proficiency in English is native. Furthermore, Julia was in French immersion all throughout high school. When she graduated, her level of proficiency was advanced, but she now feels that it is closer to intermediate. She also took a beginner-level course in Japanese when she was in her first year of university.

Dominance:	Most ←————→ Least				
Language:	1. English	2. Spanish	3. French	4. Japanese	5.
Age/grade of 1 st exposure:	Birth (age=0)	Birth (age=0)	First grade (age=6)	First year univ. (age=18)	
Immersion:	YES / NO	YES / NO	YES / NO	YES / NO	YES / NO
Age/grade:	Age: Birth	Age: Birth	Age: 12	Age: _____	Age: _____

Language: <u>English</u>	Current level of general proficiency					
	very low	low	intermediate	advanced	near native	native
	Highest level of proficiency ever attained					
	very low	low	intermediate	advanced	near native	native

Language: <u>Spanish</u>	Current level of general proficiency					
	very low	low	intermediate	advanced	near native	native
	Highest level of proficiency ever attained					
	very low	low	intermediate	advanced	near native	native

Language: <u>French</u>	Current level of general proficiency					
	very low	low	intermediate	advanced	near native	native
	Highest level of proficiency ever attained					
	very low	low	intermediate	advanced	near native	native

Language: <u>Japanese</u>	Current level of general proficiency					
	very low	low	intermediate	advanced	near native	native
	Highest level of proficiency ever attained					
	very low	low	intermediate	advanced	near native	native

Language Background Questionnaire – Extended Version



Participant information – To be filled out by the researcher

Project code: _____

Today's date: _____

Participant code: _____

1.1 Parent/Caregivers' information

Parent/Primary caregiver (from birth to 24 months): _____

Native language(s): _____ Other language(s): _____

During your infancy, their language(s) of communication:

- with you: _____
- in the home: _____
- with other family members: _____
- with people outside the home: _____

Place of birth: _____ Current residence: _____

Please list all places where this caregiver has lived, when they lived there, and for how long, in chronological order:

Parent/Other primary caregiver (from birth to 24 months): _____

Native language(s): _____ Other language(s): _____

During your infancy, their language(s) of communication:

- with you: _____
- in the home: _____
- with other family members: _____
- with people outside the home: _____

Place of birth: _____ Current residence: _____

Please list all places where this caregiver has lived, when they lived there, and for how long, in chronological order:

Other caregiver(s) (from birth to 24 months): _____

Native language(s): _____ Other language(s): _____

During your infancy, their language(s) of communication:

- with you: _____
- in the home: _____
- with other family members: _____
- with people outside the home: _____

Place of birth: _____ Current residence: _____

Please list all places where this caregiver has lived, when they lived there, and for how long, in chronological order:

Please list these caregivers in order of amount of time you spent with them. Also indicate approximately how much time you spent with each of them during a typical week (from birth to 24 months).

1. _____ Approx. time spent per week: _____
2. _____ Approx. time spent per week: _____
3. _____ Approx. time spent per week: _____

Language Background Questionnaire – Extended Version

1.2 Languages in your environment during infancy (0 to 24 months of age inclusively)

Please list all the languages that you were exposed to **during your infancy** and, for each language, the approximate percentage of the time that you heard it on a weekly basis. Note: This should add up to 100%.

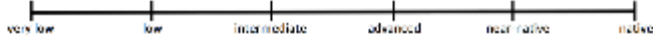
Language(s)	Percentage of the time that you heard this language on a weekly basis DURING INFANCY:

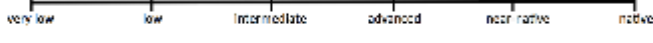
2. Current Language Proficiency

Please evaluate your **current** level of proficiency for **all** languages that you have **ever** been exposed to:

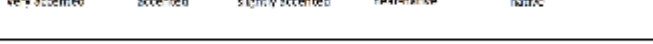
Language: _____

Oral comprehension: 

Oral production: 

Writing proficiency: 

Reading proficiency: 

Pronunciation: 

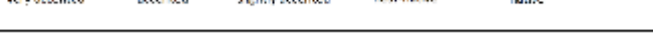
Language: _____

Oral comprehension: 

Oral production: 

Writing proficiency: 

Reading proficiency: 

Pronunciation: 

Language Background Questionnaire – Extended Version

Language: _____

Oral comprehension: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Oral production: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Writing proficiency: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Reading proficiency: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Pronunciation: ☐ very accented ☐ accented ☐ slightly accented ☐ near-native ☐ native

Language: _____

Oral comprehension: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Oral production: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Writing proficiency: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Reading proficiency: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Pronunciation: ☐ very accented ☐ accented ☐ slightly accented ☐ near-native ☐ native

Language: _____

Oral comprehension: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Oral production: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Writing proficiency: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Reading proficiency: ☐ very low ☐ low ☐ intermediate ☐ advanced ☐ near-native ☐ native

Pronunciation: ☐ very accented ☐ accented ☐ slightly accented ☐ near-native ☐ native

If you have been exposed to more than the above 5 languages, please list the others here and comment on your general proficiency for each:

Language Background Questionnaire – Extended Version

3. Evolution of Language Use

1. In the first row, fill in the ages, grades, or calendar years corresponding to the education level specified on top of each column. **Please ask the researcher if you need help.**
2. In each cell, use **percentages** to indicate your usage of/exposure to English, French, and other languages (combined) for the corresponding context and age.

Note: If your language use changed within these age groups, or if the age group is inaccurate with respect to education level, please specify by writing it inside the box or explain in the comments sections below. If you have any questions, please speak to your experimenter.

	Age 2 to kindergarten	Kindergarten	Elementary school 1	Elementary school 2
Ages/grades/calendar years				
School: i.e. language of instruction. If French immersion, please specify in comments below which grades and number of hours/week	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
At home: interactions with immediate and extended family, significant other (if you lived with them) and roommates	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Friends: interactions with friends and significant other (if you did not live with them)	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Media use: social media, leisurely reading, television, cinema, radio, internet, music, etc.	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Extracurricular activities: sports, hobbies, work (if part-time), etc.	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Daily activities in the community: grocery store, shopping mall, restaurants, gas station, etc.	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Other: _____	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:

Comments: _____

Language Background Questionnaire – Extended Version

	High school	College/Cégep	University	Other: _____
Ages/grades/calendar years				
School: i.e. language of instruction. If French immersion, please specify in comments below which grades and number of hours/week	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
At home: interactions with immediate and extended family, significant other (if you lived with them) and roommates	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Friends: interactions with friends and significant other (if you did not live with them)	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Media use: social media, leisurely reading, television, cinema, radio, internet, music, etc.	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Extracurricular activities: sports, hobbies, work (if less than 20h/week), etc.	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Daily activities in the community: grocery store, shopping mall, restaurants, gas station, etc.	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Work: (if over 20h/week)	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:
Other: _____	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:	Eng: Fr: other:

Comments: _____

Thank you!

Appendix B: Parental Questionnaire about Infant Language Exposure

Language Development Lab: Bilingual Questionnaire

Subject ID: _____ Birthdate: _____ Age in months: _____ Test Date: _____

Language environment: Language in the family

Have you or your baby's mother/father ever had a communication disorder?

Do you have any reason to believe that your baby has a hearing problem?

What language(s) do you speak to your baby? When you were growing up, what language(s) did your parents speak to you?

What language(s) does your baby's mom/dad speak to your baby? When he/she was growing up, what language(s) did his/her parents speak to him/her?



Circles represent mothers and squares represent fathers.

Details of language exposure

1. How many hours per day is your baby awake?

- This question operates on the assumption that for most of the time that the baby is awake, someone is talking to him/her. Thus, this information allows us to see how many hours of language exposure per day the baby is receiving.

2. When your baby is awake, are you always with him/her? If yes, proceed to 3. If no, ask a), b), and c).

- Do you work? If yes, how many days per week do you work?
- On a working day, how many hours do you spend with your baby, during which your baby is awake.
- On a weekend/non-working day, how many hours do you spend with your baby, during which your baby is awake?

3. If the interviewee does not work, ask a). If the interviewee works, ask b) and c).

- How many hours per day do you speak English/French/other languages to your baby?
- During the X hours that you spend with your baby on a working day, how much time do you spend speaking English/French/other languages?
- During the X hours that you spend with your baby on a non-working day, how much time do you spend speaking English/French/other languages?

BILINGUAL STATISTICAL LEARNING

4. Has this situation been the same since birth? If yes, proceed to 5. If no, ask the following:
- For how many months has this situation has been the same?
 - Confirm the amount of time before this, and ask the same questions about the period of time before this.
 - e.g. If baby is 16 months old and the mom says she started working when the baby was 12 months old, confirm that the above situation has been the same for 4 months, and ask questions 2 and 3 about the first 12 months of the baby's life.
5. Does anyone else talk to your baby on a regular basis (more than once a week)? Ask questions 2-4 regarding each person the parent lists.
- Examples: other parent, grandparents, siblings, parents' friends, day care, etc.

Person	Time period	Language	Hours/day	Days/Week	Calculation

Note: the Time Period column is only relevant if the situation has not been the same since birth. If this is the case, write the elapsed time for each situation in the time period column, and divide the time period by the age of the baby, then multiply the calculation of that row by the age of the baby. If the situation has been the same since birth for all language contributors, this division is not necessary.

Video conferencing (e.g. Skype): only include interactive conferencing; exclude TV.

Person	Time period	Language	Hours/day	Days/week	Calculation

Note: include video conferencing in the total calculation.

Global estimate

What is the percentage of time your baby is hearing English/French/other languages?

Average of parents' global estimate and calculation: