

**EXAMINING THE FRENCH AND ENGLISH LANGUAGE PROFICIENCY OF  
GRADE 11 FRENCH IMMERSION STUDENTS IN BRITISH COLUMBIA**

by

Anahita Afrin Rustom

B.A. (Hons), The University of British Columbia, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES  
(Audiology and Speech Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2015

© Anahita Afrin Rustom, 2015

## **Abstract**

The primary objective of the present study was to compare the French and English language proficiency of Early French Immersion (EFI) and Late French Immersion (LFI) students nearing the time of high-school graduation. This research was undertaken as a means of exploring the effect of age on second language proficiency, in the context of French immersion. Four indicators of language proficiency were examined: receptive vocabulary knowledge, grammar knowledge, listening comprehension, and pronunciation. Participants were evaluated in French and English. Results demonstrated no significant differences between EFI and LFI groups on the French language measures, suggesting evidence against the notion of a sensitive period for language learning in the context of French immersion.

Subsequent analyses were conducted in which the sample was divided by home-language into bilinguals (participants who spoke only English at home) and multilinguals (participants who spoke a language other than English at home). These analyses revealed that multilinguals performed as well as, and on some measures better than bilinguals in both languages, with the exception of English pronunciation. Higher language aptitude and motivation to learn a foreign language observed in the multilingual group were factors that could explain these results. Due to the better performance of multilinguals and the uneven distribution of multilinguals across EFI and LFI groups, it is unclear whether this study provides evidence against age effects in formal language education. The results emphasize the success of multilinguals in EFI and LFI, but also highlight the need to account for the changing demographics of students in French immersion programs in future studies.

## **Preface**

My contribution to this thesis was in the formulation of the research questions, the design of the study, the data collection, the data analyses, and the interpretation of the results. Dr. Stefka Marinova-Todd advised me on the formulation of research questions, the design of the study, the interpretation of the results, and preparation for all chapters of this thesis.

This project received ethics approval at UBC from the Behavioural Research Ethics Board under the title “Language Achievement in Early and Late French immersion Programs” (H14-03048).

## Table of contents

<b>Abstract</b> .....	<b>ii</b>
<b>Preface</b> .....	<b>iii</b>
<b>Table of contents</b> .....	<b>iv</b>
<b>List of tables</b> .....	<b>v</b>
<b>List of abbreviations</b> .....	<b>vi</b>
<b>Acknowledgements</b> .....	<b>vii</b>
<b>Dedication</b> .....	<b>ix</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1 Overview.....	1
1.2 Literature review .....	3
1.2.1 Critical period hypothesis .....	3
1.2.2 Making sense of mixed evidence.....	7
1.2.3 A review of research comparing EFI and LFI across Canada .....	10
1.3 The present study .....	16
1.3.1 Central research question and hypotheses .....	17
<b>Chapter 2: Methods</b> .....	<b>18</b>
2.1 Participants.....	18
2.1.1 French immersion students .....	18
2.1.2 Raters of pronunciation.....	21
2.2 Procedures.....	21
2.3 Materials .....	22
2.3.1 French language proficiency measures .....	23
2.3.2 English Language Proficiency Measures.....	26
2.3.3 Background questionnaire & language history questionnaire .....	29
<b>Chapter 3: Results</b> .....	<b>31</b>
3.1 Data analyses .....	31
3.1.1 EFI and LFI performance on French language proficiency measures .....	32
3.1.1.1 Post-hoc analyses on French grammar knowledge task .....	33
3.1.2 EFI and LFI performance on English language proficiency measures .....	35
3.2 Dividing the sample based on home-language .....	36
3.2.1 Post-hoc analyses on language aptitude and motivation.....	39
3.3 Summary of results .....	41
<b>Chapter 4: Discussion</b> .....	<b>42</b>
4.1 French language proficiency outcomes .....	42
4.2 English language proficiency outcomes .....	46
4.3 Revisiting the notion of a sensitive period for language learning .....	49
4.4 Strengths and limitations.....	51
4.5 Future directions .....	53
4.6 Conclusion .....	54
<b>References</b> .....	<b>56</b>
<b>Appendices</b> .....	<b>66</b>
Appendix A French grammaticality judgement task (adapted from Birdsong, 1992).....	66
Appendix B English grammaticality judgement task (adapted from McDonald, 2006) .....	69
Appendix C Material used for speech production tasks .....	73
C.1 French paragraph (O’Shaughnessy, 1984).....	73
C.2 English paragraph (Speech Accent Archive, Weinberger, 2015).....	73

## List of tables

<b>Table 1:</b> Demographic information for French immersion participants .....	20
<b>Table 2:</b> Examples of stimuli for the French grammaticality judgement test (adapted from Birdsong, 1992).....	25
<b>Table 3:</b> Examples of stimuli for the English grammaticality judgement test (adapted from McDonald, 2006) .....	28
<b>Table 4:</b> Mean scores (and standard deviations) for EFI and LFI on French language proficiency measures with independent samples t-test results.....	32
<b>Table 5:</b> EFI and LFI performance on each sentence-type within the French grammar knowledge task .....	34
<b>Table 6:</b> Mean scores (and standard deviations) for EFI and LFI on English language proficiency measures with independent samples t-test results.....	36
<b>Table 7:</b> Mean scores (and standard deviations) for bilinguals and multilinguals on French language proficiency measures with independent samples t-test results .....	37
<b>Table 8:</b> Mean scores (and standard deviations) for bilinguals and multilinguals on English language proficiency measures with independent samples t-test results .....	38
<b>Table 9:</b> Post-hoc analyses on bilinguals and multilinguals for language aptitude and motivation .....	40

**List of abbreviations**

<b>CPH</b> .....	<b>Critical Period Hypothesis</b>
<b>EFL</b> .....	<b>Early French Immersion</b>
<b>LFI</b> .....	<b>Late French Immersion</b>
<b>L1</b> .....	<b>First (native) Language</b>
<b>L2</b> .....	<b>Second Language</b>
<b>L3</b> .....	<b>Third Language</b>
<b>MFI</b> .....	<b>Middle French Immersion</b>
<b>RT</b> .....	<b>Reaction Time</b>
<b>SES</b> .....	<b>Socio-Economic-Status</b>

## **Acknowledgements**

This master's thesis would not have been possible without the help of many individuals who have contributed their valuable time and assistance. It is a pleasure to convey my appreciation and gratitude to all of them in my humble acknowledgment.

First, I would like to express my gratitude to my research supervisor, Dr. Stefka Marinova-Todd, for her expertise and support during the research process, and for her insightful feedback on previous versions of this work. In part, the research reported in this thesis was funded by the VPRI Research Grant for Non-Affiliated SSRHC Faculty awarded to Dr. Marinova-Todd.

I would also like to sincerely thank the members of my supervisory committee, Dr. Jeff Small and Dr. Richard Hodgson, for their valuable time, helpful suggestions, stimulating discussions, and encouragement throughout this academic journey.

It is with immense gratitude that I acknowledge the teachers and principals who graciously allowed me to conduct this research in their schools. I would also like to thank the participants for their time and willingness to support this research.

I wish to acknowledge my colleague, Katrina Kwan, and the research assistants who kindly facilitated with data collection. As well, I would like to thank Dr. Paola Colozzo, who generously offered the use of her laptop for data collection when technical difficulties arose. I also wish to recognize the helpful staff at the UBC School of Audiology and Speech Sciences and the UBC Library, as well as the kind stranger who taught me how to use a microfiche machine in order to access some older research articles. I would also like to thank my colleague,

Hillary Stahl, and my sister, Nazneen Rustom, for helping to demystify the statistical software, and my fiancé, Narius Dastur, for his technical expertise.

Most of all, I am grateful to God, and I wish to thank my dear friends, family, and fiancé, for their unwavering love, kindness, and support. Thank you for always believing in me and continuously lifting my spirits!

*This work is dedicated to my beloved family,  
my parents, Mahrukh & Sarosh,  
and my sister, Nazneen,  
all of whom inspire me each and every day!*

## **Chapter 1: Introduction**

### **1.1 Overview**

This study compared the French and English language proficiency of students in early and late French immersion programs, nearing the time of high-school graduation. Over the 40 years that these two French immersion programs have been implemented in Canada, only two studies (viz., Morrison & Pawley, 1986; Turnbull, Lapkin, Hart, & Swain, 1998) have compared the language outcomes of early and late French immersion students nearing the end of high-school. Prior to the current study, no research comparing the language outcomes of early and late French immersion students in high school had been undertaken exclusively in the Canadian province of British Columbia. This province differs from provinces in the eastern regions of Canada, where French is used more widely outside of the classroom and within the community. Thus, the present study aimed to examine the French language outcomes of early and late French immersion students based solely upon program completion.

The first French immersion program in Canada was introduced in 1965, within the bilingual community of St. Lambert, Quebec. The purpose of this program was to help monolingual Anglophone children acquire French, in order to help them communicate with the French-speaking members of their community. Rather than being taught about the French language, French became the medium of instruction, starting at the kindergarten level. Shortly after this time, the recommendation that French become one of Canada's two official languages, along with English, was made by the Canadian Royal Commission on Bilingualism and Biculturalism in 1969, leading to the proliferation of French immersion programs in schools across the country.

As of 2011, 17.5% of Canadians, approximately 5.8 million people, reported being bilingual in English and French, which is more than double from 6.9% in 1961 (Lepage, & Corbeil, 2013). Much of this increase may be attributed to French immersion programs, which produce thousands of graduates year after year. The benefits of bilingualism reach far greater lengths than the ability to communicate in two languages. French immersion students, among other kinds of bilinguals, have also been found to benefit from linguistic, academic and cognitive skills, such as heightened mental flexibility and a greater communicative sensitivity (Lazaruk, 2007).

Over the past five decades, variations of French immersion programs have been employed across the county. The original program designed in St. Lambert, in which students began receiving French instruction in Kindergarten or Grade 1, at approximately the age of 5 or 6, is known today as Early French Immersion (EFI). In addition to EFI, two other French immersion programs with differing entry-points are currently available. In select school-districts across Canada, Middle French Immersion (MFI) is offered; this program starts in Grade 3 or 4, when students are approximately the age of 8 or 9. Late French Immersion (LFI) is a more common program which is offered in all Canadian provinces. LFI typically starts in Grade 6 or 7, when students are approximately the age of 12 or 13.

The present study aimed to reveal any differences in language proficiency between students of EFI and LFI programs. EFI and LFI programs contrast on age of exposure to the foreign language (French), and consequently, on the amount of time learning French. If factors such as these influence language proficiency, then the outcomes of the two programs are potentially unequal and need to be recognized. As the first study to compare EFI and LFI closer

to the time of high-school graduation in nearly 20 years, and the only study of this kind in British Columbia, the results of this study will provide information regarding the language proficiency attainable upon completion of each respective program.

## **1.2 Literature review**

Research comparing the language proficiency of students from the different immersion programs (EFI, MFI and LFI) has often been an attempt to compare the language proficiency levels of second languages learners of different ages of exposure. The hypothesis that younger language-learners will be more proficient than older language-learners is commonly known as the Critical Period Hypothesis (Lenneberg, 1967).

### **1.2.1 Critical period hypothesis**

If language achievement is a function of age of exposure, then students who entered French immersion at a younger age (i.e. EFI) would achieve greater French language proficiency than students who entered French immersion at an older age (i.e. LFI). The idea of a maturational decline in neural plasticity as a cause for increasing difficulties with language learning was first postulated by Penfield and Roberts (1959). Lenneberg further introduced the Critical Period Hypothesis (CPH), according to which the ability to fully master a language is limited to those who acquire the language before the onset of puberty. Lenneberg believed that puberty marked the time when brain lateralization (i.e. the process by which the right and left hemispheres of the brain develop specialised functions) occurred.

Although the idea of a CPH for first language (L1) acquisition has not been rejected (Long, 2013a), the extension of the CPH to second language (L2) learning is the subject of long-standing debate. There has also been disagreement with respect to why age of exposure may affect second language learning (Long, 2013a). Explanations have varied from brain lateralization (Lenneberg, 1967), to the presence of a neural language acquisition device or universal grammar (Chomsky, 1965), and more detailed theories based on the interaction of experience, environment and genetics (e.g. Purves et al., 2001).

As stated by Long (2013a), even after hundreds of studies investigating the CPH, researchers are unable to provide a clear biological explanation for the connection between age of onset and language proficiency. There has also been disagreement in the field regarding the age at which the critical period ends. Various researchers have proposed that the critical period culminates at a specific age, such as 5 years (Krashen, 1973), 6 years (Pinker, 1994), 9 years (Penfield & Roberts, 1959), 12 years (Abrahamson, 2012; Lenneberg, 1967; Singleton, 1995), 15 years (Johnson, & Newport, 1989; Patkowski, 1980) or over 16 years (DeKeyser, 2000; DeKeyser, Alfi-Shabtay, & Ravid, 2010; Weber-Fox, & Neville, 1999).

Given the lack of clear support for the CPH, the term “sensitive period”, originally introduced by Lamendella (1977), has now gained popularity in order to emphasize that later language-learning may still be attainable at a lower rate of efficiency (Lamendella, 1977). Researchers now agree that the ability to learn a second language decreases gradually over time, rather than decreasing abruptly as the CPH suggested (Ardila & Ramos, 2007; Dillon, Lamb & Grayson, 2010; Birdsong & Molis, 2001; Birdsong, 2005; Long, 2013a). Furthermore, Long (2013a), among others, suggested the possibility of multiple sensitive periods for different

domains of language learning (e.g. for pronunciation or morpho-syntax), and argued that multiple sensitive periods are better able to account for individual variation in second language acquisition.

The existence of a sensitive period or age effects in language learning has remained under active investigation. Hundreds of studies within the field of L2 acquisition have compared the language proficiency of L2 learners falling within or outside the sensitive period, in an attempt to test the bounds beyond the putative period. Some studies and reviews challenge the concept of a sensitive period for second language learning (e.g. Birdsong, 2005, 2006, 2009; Herschensohn, 2007; Marinova-Todd, 2003; Marinova-Todd, Marshall, & Snow, 2000; Muñoz, & Singleton, 2011; Singleton, 2005), while others appear to support the claim for a sensitive period (e.g. Abrahamsson, & Hyltenstam, 2009; DeKeyser, 2012; Granena & Long, 2013).

While studies have shown inconsistent results concerning a sensitive period for early and late L2 learners in naturalistic settings (e.g. Marinova-Todd, 2003 cf. Abrahamsson & Hyltenstam, 2009), evidence in support of the notion of a sensitive period has been even less clear within the context of formal L2 learning environments (i.e. school settings). To my knowledge, only one study (Larson-Hall, 2008) has provided some evidence in support of the sensitive period in the context of formal education. Participants in Larson-Hall's (2008) study were L1-Japanese university students who began studying English as a foreign language between the ages of 3 and 12. They were tested on two measures: (1) a grammaticality judgement task, and (2) a phonemic discrimination task in English. Within this group, it was found that an earlier starting age was significantly correlated with higher scores on the grammaticality judgement test, after controlling for language aptitude and amount of language input. However, starting age was

not correlated with performance on the phonemic discrimination task. Yet, when the results of this group of participants were compared to that of a second group, who began learning English in middle school at the age of 12 or 13, the opposite trend was found. There was a significant effect of age found for performance on the phonemic discrimination task, such that the first group of younger-learners performed better than the older learners in the second group, after controlling for hours of study input. However, no age effects were found for performance on the grammaticality judgement test, even after controlling for hours of study input. Larson-Hall (2008) concluded that age effects may only be perceived for some linguistic domains in an L2 instructional context after controlling for amount of input.

In contrast to Larson-Hall's (2008) study, the majority of studies comparing early and late learners in a formal education setting have suggested evidence against the notion of a sensitive period (e.g. Al-Thubaiti, 2010; Cenoz, 2003; Muñoz, 2011). In a recent study by Al-Thubaiti (2010), early and late learners of English were tested on four measures of morpho-syntax and semantics. The early-learner group was comprised of 50 L1-Arabic university students who learned English as an L2 in kindergarten or elementary school, between the ages of 3 and 11. In contrast, the late-learner group was comprised of 82 L1-Arabic university students who began learning English in middle school, between the ages of 12 and 13. The early-learner group had studied English for an average of 15.2 years, with an average of 846 hours of instruction, while the late-learner group had studied English for an average of 10.2 years, with an average of 576 hours of instruction. No differences between early and late learners of English as a foreign language were found on any of the measures, suggesting evidence against the putative sensitive period in an L2 education context.

Muñoz, (2008; 2010) argued that the extension of the CPH/sensitive period from more naturalistic L2 learning contexts to more formal contexts should be done with caution, due to differences in the amount and quality of input between the two settings. However, Long (2013b) argued that researchers investigating age effects in formal language learning contexts have been unable to identify effects of the putative sensitive period because of the language outcomes measured in these studies. According to Long (2013b), pronunciation should be measured, as strong negative-correlations between age of onset and global pronunciation ratings in the L2 have been noted in naturalistic settings (e.g. Patkowski, 1980; Yeni-Komshian, Flege & Liu, 2000; Granena & Long, 2013). In addition, Long (2013b) argued that listening comprehension skills should also be measured, as this linguistic skill has also been linked with an early-learner advantage in naturalistic contexts (e.g. Oyama, 1978).

Giving consideration to Long's arguments, the present study aimed to compare the language proficiency of early and later language learners with varied language measures, including pronunciation and listening comprehension. Thus, if there is a sensitive period for language learning in a formal education L2 context, then one would expect that the EFI group would perform better than the LFI group on these measures. However, if Muñoz, (2008; 2010) is correct in differentiating between formal and naturalistic contexts, then there should be no differences between EFI and LFI groups on measures of French language proficiency.

### **1.2.2 Making sense of mixed evidence**

Many researchers have attempted to make sense of the mixed evidence regarding the critical/sensitive period for L2 acquisition (e.g. Huang, 2014; DeKeyser, 2013). These

researchers have shed light on the differences in conceptual understandings and methodological obscurities which have likely contributed to the heterogeneity in the research evidence to date. For example, researchers have employed a variety of tests and testing batteries when assessing language proficiency outcomes. Even studies which have appeared to use a same-named measure (e.g. a grammaticality judgement test) might differ on modality (e.g. if the stimuli are presented aurally or on paper) or format (e.g. if the test is timed or not), which are aimed to draw on different aspects of language knowledge (e.g. implicit or explicit), and processing (Huang, 2014; Murphy, 1997). CPH researchers have also addressed differences in the sensitivity of the stimuli or measures employed (e.g. Abrahamsson, & Hyletenstam, 2009; Birdsong, & Molis, 2001). For instance, it has been shown that the sensitivity of a grammaticality judgement task may be elevated by measuring reaction time and accuracy, rather than accuracy alone. In order to address these methodological concerns, the present study used a combination of standardized measures, and experimental measures which were previously employed in other studies, including a timed measure of grammar knowledge.

Another important issue that has been debated is the practice of comparing the language proficiency of L2-learners with that of native (L1) learners, who serve as ‘controls’. As stated by Grosjean (1989), “the bilingual is not two monolinguals in one person!” (p.3), and thus, bilinguals should not be compared against monolinguals. In the same vein, Cook (2002) asserted that “ultimate attainment is a monolingual standard, rather than an L2 standard” (P.6). In other words, there is no reason why an L2 learner’s language proficiency should be the same as that of a monolingual native speaker, and that the focus should be on L2 learners in their own right. This rationale has led to what is known as the “bilingual turn” (Ortega, 2013). In this view, it is

argued that comparing the language proficiency of monolinguals to that of L2 learners/bilinguals in only one of the languages they speak created a monolingual bias and a deficit view of L2 acquisition from the outset. Moreover, Ortega (2013) argued that comparing bilinguals and L2-learners to one another (rather to monolinguals), will enhance the field of bilingualism and impact language learning in education. Accordingly, in the present study, the language outcomes of EFI and LFI students are compared to one another, and not against monolingual L1 speakers of French and/or English.

Furthermore, many studies in L2 acquisition only evaluate learners in the L2, and neglect to evaluate language learners in their L1 or in other languages they may speak. Research undertaken in this way is unable to account for potential language transfer effects from L1 to L2, even though it has been accepted that language transfer plays a role in L2 development (Treffers-Daller & Sakel, 2015). Therefore, research in the field of L2 acquisition should consider attributes of the L1 (e.g. aspects of grammar, pronunciation, etc.) and how they might transfer to the L2 or subsequent languages learned. Testing skills in the L1 and being aware of the linguistic features of the L1 could help explain language performance in the L2; further, testing participants in both languages could also provide insight regarding expected and unexpected language outcomes.

In addition to language transfer, studies which only test performance in the L2 are unable to account for degree of proficiency in the L1, and how it might affect proficiency in the L2. A recent study by Sparks, Patton and Ganschow (2012) found that students' level of achievement in L1 prior to L2 exposure was strongly related to their language learning aptitude and degree of proficiency achieved in the L2. Their research underlined the importance of evaluating bilinguals

and L2 learners in both languages known, and that this practice is essential for the identification of individual differences in L2 proficiency. As a result, the students in the present study were tested in both languages of their formal education (French and English), for the majority of whom, English was a native language.

### **1.2.3 A review of research comparing EFI and LFI across Canada**

When comparing EFI and LFI populations, the effect of age of entry into immersion is evaluated. An earlier start in language learning also results in more cumulative time in a language learning environment. Thus, EFI students have an earlier age of exposure and more cumulative time in a French immersion environment. The notion of a sensitive period would suggest that EFI students would perform better than LFI students on measures of French language proficiency. However, studies comparing same-grade EFI and LFI students have revealed mixed results for different domains of language proficiency.

In a cross-sectional study, Genesee (1981) compared the French language outcomes of two groups of LFI students (one-year LFI and two-year LFI) with EFI students in Grades 7, 8 and 9. The one-year LFI group consisted of students who had been in an LFI program starting in Grade 7 in which 80% instruction was in French, with 40% instruction in French in Grades 8 and 9. The two-year LFI group consisted of students in a Grade 7-8 late-immersion program in which 80% instruction was in French, with 40% instruction in French in Grade 9. Reading ability, reading comprehension, dictation, writing, listening comprehension and speaking ability were measured. In Grade 7, the EFI group performed better than the one-year LFI group on every language measure, except speaking ability, for which no significant group differences were

found. In Grade 8, no differences were found between the EFI and two-year LFI students, except on the written dictation task, on which the EFI group scored higher; in addition, EFI and two-year LFI groups performed better than the one-year LFI group on all measures, except speaking ability, on which no significant differences were found. In Grade 9, the EFI group and two-year LFI group performed equivalently on all measures, except for listening comprehension; in addition, EFI scored higher than the one-year LFI group on the dictation task, listening comprehension and speaking ability. Focussing on the comparisons between EFI and two-year LFI groups, Genesee (1981) concluded that there was only a slight EFI advantage after two years of LFI studies (i.e. for listening comprehension), but cautioned that his conclusion may have been contingent upon the particular program structure of the EFI and LFI programs in the Ottawa-Carleton and Montreal regions.

In contrast, other researchers have found that EFI and LFI groups achieved equivalent language proficiency levels at the Grade 7 level (Bruck, Lambert, & Tucker, 1975). Though, subsequent evaluations on the following cohort completed one year later revealed an EFI-advantage within all domains of language proficiency: listening, speaking, reading and writing (Bruck, Lambert, & Tucker, 1976). Bruck et al. (1976) stated that the discrepancies between the findings of their two studies were likely due to differences in the attitude towards language testing between the EFI cohorts in each respective study. Namely, Bruck et al. (1976) suggest that the EFI group in their first study could have cultivated negative attitudes toward language testing as a result of repeated testing over the years, and thus, that their language data may not have been a reliable representation of language performance in EFI (Bruck et al., 1976).

Further research conducted by Cziko, Holobow and Lambert (1977), which compared EFI and LFI students in Montreal at the end of Grade 7, illustrated a similar result to Bruck et al. (1976): students in EFI performed better on most measures of French language proficiency, including writing, reading and comprehension. However, this study also found that the students in the LFI program performed better than students in the EFI group on measures of vocabulary and grammar in a French speaking task, in the context of a recorded semi-directed interview. Though, the authors also pointed out that the poor quality of some of the audio-recordings of the speaking task may have inadvertently led to the higher scores achieved by some students in LFI.

On the one hand, studies comparing EFI and LFI students from Grades 7-9 show equal performance levels on most indicators of language proficiency (Bruck et al., 1975; Genesee, 1981). These results are supportive of the notion that that age effects may be mediated by learning context, such that starting age may play a less influential role in formal education settings (Muñoz, 2008; 2010). On the other hand, some studies showed advantages for the early programs (Bruck et al., 1976; Cziko et al., 1977) and possibly for late programs (Cziko et al., 1977). The discrepancies between the findings of these studies could be due to several factors, including differences in methodology (e.g. use of different language outcome measures, quality/validity of the measures), differences in sample characteristics (e.g. age at the time of testing, socio-economic status), or differences between the EFI and LFI programs across school districts. Further, since this time-frame is still early on in the learning process for the LFI group, it may be too early to draw conclusions about the language outcomes attainable by LFI students. Research comparing EFI and LFI students closer to the time of graduation could help determine if the differences between programs found in earlier grades resolve over time. To the best of my

knowledge, only two such studies have been undertaken thus far (Morrison & Pawley, 1986; Turnbull et al., 1998).

Morrison and Pawley (1986) studied the French language proficiency levels attained by EFI and LFI groups at the Grade 12 level. This research took place in the Ottawa/Carleton region. Participants were tested on measures of listening and reading comprehension, writing skills in a dictation task, and speaking, in the context of a semi-directed interview, which was further evaluated on aspects of fluency (i.e. speech rate and rhythm) and communicative competence (e.g. use of vocabulary and syntax). Results demonstrated that the EFI and LFI groups performed comparably on the listening comprehension, reading comprehension, and writing measures. However, the EFI group scored significantly higher on the French speaking measure, suggesting evidence in support of a sensitive period for the development of oral communication.

Similarly, Turnbull et al. (1998) conducted a study comparing the French language proficiency of EFI, LFI and MFI students in Grade 12, across both eastern and western regions of Canada. Researchers reported that EFI students had accumulated approximately 6000 hours of French instruction, whereas MFI students accumulated 2000 hours and LFI students accumulated 1200 hours of French instruction prior to Grade 8, the point at which students from all programs were joined together. Language proficiency was tested with the following measures: listening (listening comprehension and sentence-repetition), reading (oral reading and reading comprehension), writing (a cloze-task and a written opinion measure), and speaking (sentence-repetition and oral opinion). The results showed that the EFI group performed better than the LFI and MFI groups only on the sentence-repetition task, which involved components of listening

and speaking. The EFI group also performed better than the MFI group on the cloze-task, but no other group differences were found on any of the other language measures.

On the whole, the findings of the study by Turnbull et al. (1998) corroborate the findings of Morrison and Pawley (1986). However, the sentence repetition task used by Turnbull et al. (1998) drew on aspects of listening and speaking which were not measured by Morrison & Pawley (1986), who used distinct tasks to measure speaking ability (i.e. semi-directed interviews) and listening (i.e. a listening comprehension task). Given the nature of a repetition task, it is likely that this task drew on skills involving auditory-verbal working memory, in addition to comprehension. Thus, the differing results between Turnbull et al. (1998) and Morrison and Pawley (1986) concerning EFI and LFI performance on the listening tasks, could be due to differences in task-demands.

With the exception of the study by Turnbull et al. (1998), which included some participants from western regions of Canada, all of the studies discussed thus far involved EFI and LFI participants in eastern regions of Canada, in areas of Ontario and Quebec, which have the highest numbers of native-French and native-French-English bilingual speakers in Canada (Statistics Canada, 2012a). In contrast, British Columbia has a much lower number of native-French or native English-French bilingual speakers (Statistics Canada, 2012a). Therefore, research on French immersion students from the eastern regions of Canada may represent language outcomes which are reflective of French immersion education aided by living in an environment which is more likely to facilitate community interaction in French. Thus, studies which evaluate the French language outcomes of EFI and LFI students in British Columbia

would be more likely to represent the levels of French attainable almost exclusively within the context of French immersion education.

To date, only one study has compared the performance of EFI and LFI students in the province of British Columbia. In their study, Day and Shapson (1988) investigated EFI and LFI students in Grade 7, in three school districts across British Columbia. Students were tested with the *test de rendement*, which included different aspects of language proficiency (vocabulary, spelling, reading comprehension and grammar), in addition to a listening comprehension task. Analyses revealed no statistically significant differences between the scores of EFI and LFI groups in two out of the three districts on the *test de rendement*; however, in one of the three districts, the EFI group performed better than the LFI group. In addition, it was found that the LFI students scored significantly lower on the listening comprehension measure across all three districts. In their conclusion, Day and Shapson (1988) argued that listening comprehension could continue to develop over time, and that future research performed closer to the time of graduation could determine whether equivalency between the two groups of immersion students is achievable.

In summary, research comparing the language outcomes of EFI and LFI students from Grade 7 to Grade 9 has demonstrated inconsistent results among the language skills tested, and the comparative language outcomes of EFI and LFI students. Some studies have suggested that even in the first year of French immersion, LFI students are able to achieve language outcomes which are equivalent to EFI students. However, other studies have demonstrated an EFI advantage. Differences in the participant data set and/or between the language outcomes measured may account for the discrepancies between the results of the different studies.

Research comparing EFI and LFI groups closer to the time of graduation has been limited, but more consistent, than research in earlier grades. These studies showed that students in EFI and LFI programs demonstrated comparable levels of language achievement on most language proficiency measures.

### **1.3 The present study**

The main objective of the present study was to compare the language proficiency of EFI and LFI students nearing the time of graduation, within the province of British Columbia. This study aimed to provide further evidence supporting or refuting the putative sensitive period in the context of French immersion. Given the varied findings in a formal education context, including EFI and LFI, recent work has called for greater methodological rigor, and sensitivity to context when interpreting research findings (e.g. DeKeyser, 2013; Muñoz, 2008; Long, 2013b).

In line with the recommendation of Sparks et al. (2012), EFI and LFI groups were tested in French and English, in order to establish outcomes in both languages of their formal education. Further, the language outcomes of the two groups of L2 learners (EFI and LFI) were compared with one another, rather than against monolingual/L1 speakers of French or English, which addressed concerns regarding a monolingual-bias previously raised by Cook (2002) and Ortega (2013).

### 1.3.1 Central research question and hypotheses

The main research question was: *Are there differences in the French and/or English language proficiency of EFI and LFI students nearing the time of graduation?*

In addition, the following hypotheses were addressed:

1) *Hypothesis #1: French language proficiency*

Based on previous work comparing EFI and LFI at the Grade 12 level (i.e. Morrison & Pawley, 1986; Turnbull et al., 1998), it was hypothesized that EFI and LFI groups would perform equivalently on the receptive vocabulary knowledge, grammar knowledge and listening comprehension tasks, but that EFI participants would perform better on pronunciation in French.

2) *Hypothesis #2: English language proficiency*

It was expected that EFI and LFI groups would perform equivalently on all measures of English language proficiency, provided that the groups did not differ in age of exposure to English. This hypothesis is consistent with the concept that both French immersion programs foster additive-bilingualism, and in no way impede learning English (Lazaruk, 2007).

## **Chapter 2: Methods**

### **2.1 Participants**

Two sets of participants took part in this study. The first set of participants involved French immersion students, in either EFI or LFI. The second set of participants included groups of native-English-speakers and native-French speakers, who rated the pronunciation of the French immersion students in their respective native language. French immersion students and raters of pronunciation are described in the following sections.

#### **2.1.1 French immersion students**

A total of 83 French immersion students in Grade 11 (56 EFI, 27 LFI) participated in this study. EFI and LFI participants were recruited from two public secondary schools within a western Canadian school district. Within this school district, students enrolled in EFI started receiving French instruction in Kindergarten. From Kindergarten to Grade 2, teaching was delivered entirely in French. From Grades 3 to 7, children were taught in French for 80% of the time, as they were also enrolled in an English language arts class. From Grade 8 to 10, 50% of their course-load was offered in French, and in Grade 11, 25% of courses were offered in French.

In comparison, students enrolled in LFI began in Grade 6, in which the first year was taught entirely in French. In Grade 7, LFI students were taught in French 80% of the time, as they were enrolled in an English language arts class. LFI and EFI students were joined together as a single cohort in secondary school, from Grade 8 to 12. Upon completing Grade 12, students in both programs would receive a French immersion designation on their high school transcripts.

At the time of testing, all French immersion participants were in Grade 11. The average age of participants in the EFI group was 16;8 (range: 16;3-17;3), and the average age of participants in the LFI group was 16;7 (range: 16;2-17;1). As the city in which the school district was located is culturally and linguistically diverse, information about the participants' language backgrounds and socio-economic status (SES) were also collected through a paper-based background questionnaire and an online language history questionnaire (LHQ 2.0, Li, Zhang, Tsai & Puls, 2013). As shown in Table 1, many participants in the sample were native-bilinguals, 37.35% (i.e. speaking two languages from birth), or L1 speakers of languages other than English, 19.27%. Participants who were L1 speakers of other languages reported that their first exposure to English was between the ages of 3 and 6. For both EFI and LFI groups, the average household income was reported to be between \$75,000 and \$100,000. A summary of the demographic information and details regarding participants' language backgrounds, for both EFI and LFI students is presented in Table 1.

**Table 1: Demographic information for French immersion participants**

	<b>Early French Immersion</b> <i>(n = 56)</i>	<b>Late French Immersion</b> <i>(n = 27)</i>
<b>Age, mean (range)</b>	16;8 (16;3-17;3)	16;7 (16;2-17;1)
<b>Sex (<i>n</i>)</b>		
Male	20	8
Female	36	19
<b>Annual household income (M)</b>	\$75,000-100,000	\$75,000-100,000
<b>First Language (<i>n</i>)</b>		
Cantonese	4	1
English	30	6
Mandarin	4	2
Russian	1	2
Serbian	1	0
Spanish	1	0
<b>Native Bilinguals (<i>n</i>)</b>		
Amharic-English	1	0
Bengali-English	0	1
Cantonese-English	4	1
Cantonese-Mandarin	2	6
Hindi-English	0	1
Mandarin-English	3	5
Punjabi-English	0	1
Russian-English	2	0
Spanish-English	2	0
Tagalog-English	1	1

### **2.1.2 Raters of pronunciation**

A second set of participants was recruited for the purpose of rating the pronunciation of the French immersion participants. Ten native-French speakers and ten native-English speakers were recruited to evaluate French and English pronunciation, respectively. The mean age of the French rater group was 40, and the mean age of the English rater group was 32. All raters had resided in a large urban city in British Columbia for at least five years. Each rater group comprised of an equal number of males ( $n=5$ ) and females ( $n=5$ ). In addition, each rater group had equal proportions of raters of a language-related profession (e.g. speech-language pathologists, language teachers) ( $n=5$ ), and from various non-language-related professional backgrounds (e.g. nursing, computing science, etc.) ( $n=5$ ).

## **2.2 Procedures**

Prior to commencement of the study, classroom teachers sent information packages home with the prospective French immersion participants. The information packages included a form for parents/guardians to consent to their child's participation, along with a language background questionnaire, which were to be returned to the classroom teacher. Information about participants' language experience with French was collected to ensure that no significant exposure to French outside of the French immersion programs had taken place. Participants were excluded if they spoke French natively, if they reported speaking French at home, or if they had a history of speech, language, hearing or cognitive disorders. Preceding the first individual testing session, participants provided verbal assent to participate in the study. Only students who assented and whose parents consented were included in the study.

Trained research assistants conducted a total of two individual testing sessions (one in French and one in English) with each participant. Testing sessions lasted between 30-45 minutes, and were separated by one to two weeks. Within each testing session, five measures of language proficiency were collected: receptive vocabulary knowledge, grammar knowledge (accuracy and reaction time), listening comprehension, and a speech production task. The speech production task was audio-recorded. All measures were typically completed in a single session, and were administered in no particular order within a session. After completing individual testing sessions, French immersion students were also asked to fill out a language history questionnaire.

Subsequently, the audio-recordings of the French immersion students' speech production task were evaluated by native-French speakers and native-English speakers, in individual sessions. To avoid confounding due to order-effects, the audio-recordings were presented to each rater in a randomized order. Verbal instructions to raters were supplemented by written instructions on the top of each rating sheet, which included a visual depiction of the numerical rating-scale. Pronunciation ratings were made on a 9-point continuous scale, from 1 (least like a native speaker) to 9 (most like a native speaker), which is standard practice in research investigating pronunciation (Isaacs & Thomson, 2013).

### **2.3 Materials**

The materials used in the present study involved a combination of standardized and experimental measures of language proficiency. French language proficiency measures and English language proficiency measures are described in sections 2.3.1 and 2.3.2, respectively.

In addition to these measures, information about French immersion students' language learning experience and SES was collected with a paper-based language background questionnaire and a computer-based language history questionnaire, which are described in section 2.3.3.

### **2.3.1 French language proficiency measures**

#### *Receptive vocabulary knowledge*

The *Échelle de Vocabulaire en Images Peabody (ÉVIP-III)* (Dunn, Dunn & Thériault-Whalen, 1993) is a standardized test of French receptive vocabulary knowledge. In this task participants were asked to identify the picture that correctly corresponded to a word presented orally by the test administrator, among four picture alternatives. The test-retest reliability coefficient for this age group in the manual was .73 (Dunn et al., 1993).

#### *Grammar knowledge*

A grammaticality judgement test adapted from Birdsong (1992) was implemented to assess participants' grammar knowledge. This test was selected because of the comprehensive number and variety of morpho-syntactic structures assessed. A sentence was either completely correct or incorrect, in terms of either word-structure (morphology) or sentence structure (syntax). Of the 76 sentences originally included in the test by Birdsong (1992), 4 grammatical sentences were removed as they posed problems for interpretation. The French grammaticality judgement test implemented in this study consisted of a total of 72 sentences, of which 44 were grammatical and 28 were ungrammatical. Examples of stimuli of each of the seven sentence

construction types can be found in Table 2, with a complete list of sentences included on the test in Appendix A. A comprehensive description of each sentence construction type can be found in Birdsong (1992).

The grammaticality judgement test was administered on a laptop computer, using E-Prime® Experimental Software (Schneider, Eschman, & Zuccolotto, 2002). Written instructions were provided on the computer screen, and participants had the opportunity to ask the experimenter for clarification. A set of 6 practice-sentences were provided to ensure that participants knew how to complete the task. Participants were presented with one written sentence at a time on a computer screen. They were instructed to judge whether each sentence was grammatical or not, by pressing designated computer keys. In addition, they were instructed to respond as quickly and accurately as possible. Sentences were presented in a random order to control for order effects. Both the accuracy and reaction time were recorded with the E-Prime Software.

**Table 2: Examples of stimuli for the French grammaticality judgement test (adapted from Birdsong, 1992)**

Grammatical Construction	Number of sentences	Ungrammatical & Grammatical Examples
Adjacency	15	*Diane a placé dans sa chambre des fleurs. 'Diane put in her room some flowers.' Diane a placé des fleurs dans sa chambre. 'Diane put some flowers in her room.'
<i>En-avant</i>	7	Elle a lu ce livre. *Elle en téléphone a l'auteur. 'She read this book. She-of it-phones the author.' Elle a lu ce livre. Elle en aime l'auteur. 'She read this book. She - of it - likes the author.'
Prenominal Past Participle	6	*Le connu romancier vient d'arriver. 'The known novelist just arrived.' Le très connu Marcel Proust vient d'arriver. 'The well-known Marcel Proust just arrived.'
(Qui/Que) <i>That-Trace</i>	10	*Qui disais-tu qui a épousé Laure? 'Who did you say that married Laure?' Qui disais-tu a épousé Laure? 'Who did you say married Laure?'
<i>De+</i> Modifier	16	*Elle a les lettres d'écrites. 'She has the letters written.' Elle a deux lettres d'écrites. 'She has two letters written.'
Ce vs. il (elle)	10	*Sophie nous a dit que c'était malade hier soir. 'Sophie told us that she/it was sick last night.' Sophie nous a dit qu'elle était malade hier soir.' 'Sophie told us that she was sick last night.'
Middle-voice	8	*Cette maison a été vendue d'elle-même. 'This house was sold by itself.' Cette maison s'est vendue d'elle-même. 'This house sold itself.'
Total: 72		

*Note: Sentences marked with an asterisk (\*) are ungrammatical.*

### *Listening comprehension*

The *Compréhension orale* subtest from the *Test de rendement individuel de Wechsler (WIAT-II, Wechsler, 2007)* was used to measure listening comprehension in French. This subtest used a picture elicitation task to assess sentence comprehension, and included items that measured expressive and receptive vocabulary knowledge. The test-retest reliability coefficient in the manual was .88 for this age group (Wechsler, 2007).

### *Speech production*

The prompt for the speech production task was a typed and printed paragraph (See Appendix C) adapted from O'Shaughnessy (1984). This paragraph was chosen as it contained all consonant and vowel sounds in the French language, in addition to a wide range of phonetic sequences and syntactic constructions, which were representative of French speech. Participants were given one minute to prepare, prior to reading aloud and being audio-recorded.

## **2.3.2 English Language Proficiency Measures**

### *Receptive vocabulary knowledge*

The *Peabody Picture Vocabulary Test-4 (PPVT-4, Dunn & Dunn, 2007)* is a standardized test of English receptive vocabulary knowledge. Consistent with the task in the ÉVIP-3, participants were asked to identify one picture from a selection of four, which correctly corresponded to a word presented orally by the test administrator. The mean test-retest reliability coefficient for the PPVT-4 is .92 (Dunn & Dunn, 2007).

### *Grammar knowledge*

A grammaticality judgement test adapted from McDonald (2006) was used to assess participants' English grammar knowledge. Similar to the French grammaticality judgement test, this test was selected because of the comprehensive number and variety of morpho-syntactic structures assessed. This test included fifty pairs of sentences, in which each pair included a grammatical and ungrammatical version. Sentences were either completely correct or incorrect, in terms of either morphology or syntax. Ten different sentence constructions were tested. Examples of stimuli used in the grammaticality judgement test can be found in Table 3, while a complete list of the sentences used in this test can be found in Appendix B.

The grammaticality judgement test was administered on a laptop computer, using E-Prime Software. Written instructions were provided on the laptop screen and a set of 6 practice-sentences were provided. Participants were presented with one written sentence at a time on a computer screen. They were instructed to judge whether each sentence was grammatical or not, by pressing designated computer keys. In addition, they were instructed to respond as quickly and accurately as possible. Sentences were presented in a random order to control for order effects. Both the accuracy and reaction time were recorded with the E-Prime software.

**Table 3: Examples of stimuli for the English grammaticality judgement test (adapted from McDonald, 2006)**

Grammatical Construction	Number of Sentences	Ungrammatical & Grammatical Examples
Articles	8	*The lady drove same car for past 20 years. The lady drove <u>the</u> same car for the past 20 years.
Regular past tense	16	*Last night my friend walk home after dark. Last night my friend <u>walked</u> home after dark.
Third person singular	16	*The boy jump whenever he is startled. The boy <u>jumps</u> whenever he is startled.
Regular plural	16	*There are 20 flute in our marching band. There are 20 <u>flutes</u> in our marching band.
Present progressive	8	*The little girl is play with her dolls. The little girl is <u>playing</u> with her dolls.
Irregular past tense	8	*Last week the pilot flied to Paris. Last week the pilot <u>flew</u> to Paris.
Irregular plural	4	*Several of the mans decided not to go to the football game. Several of the <u>men</u> decided not to go to the football game.
Wh-questions	8	*What you think about the new coach? What <u>do</u> you think about the new coach?
Yes/no questions	8	*Drives the teacher a really fancy red car? <u>Does</u> the teacher drive a really fancy car?
Word order	8	*The teacher the tests graded. The teacher <u>graded</u> the tests.
Total: 100		

*Note: Sentences marked with an asterisk (\*) are ungrammatical.*

### *Listening comprehension*

The *Listening Comprehension* subtest from the *Woodcock Language Battery Proficiency-Revised (WLPB-R)* (Woodcock, 1991) was used to measure listening comprehension. In this standardized task, the experimenter read short sentences aloud, with each sentence missing a single word. This task required participants to say one word, which would appropriately complete the sentence both in terms of structure (syntax) and meaning (semantics). The test-retest reliability coefficient for this assessment was .75 (Woodcock, 1991).

### *Speech production*

The prompt for the speech production task was a short paragraph, which was typed and presented to participants on paper. The English paragraph (See Appendix C) was obtained from The Speech Archive website (Weinberger, 2015), and was previously used by Huang (2014) to study the effect of age on speech production. This paragraph was chosen because it contained the full inventory of English vowels and consonants.

### **2.3.3 Background questionnaire & language history questionnaire**

The background questionnaire gathered information about the students' experiences of French immersion, languages spoken in the home, parents' English proficiency levels, parents' current occupations, and annual household income as an indication of SES. In addition to this paper-based questionnaire, the LHQ 2.0 (Li, Zhang, Tsai & Puls, 2013) collected information about students' current level of education, country of origin and of residence, self-rated language abilities, and years of language use. For the purpose of the present study, information about the

participants' socio-demographic history (e.g. age, gender and socio-economic status), and language background (e.g. languages known, years of language exposure, languages spoken in the home), were included.

## Chapter 3: Results

### 3.1 Data analyses

To test my hypotheses, each indicator of language proficiency was analyzed with an independent samples t-test. For the standardized language measures, receptive vocabulary knowledge and listening comprehension, the t-tests were performed on the standardized scores. For the experimental measures, grammar knowledge and pronunciation, the t-tests were performed on the raw scores. To correct for multiple comparisons, the Holm-Bonferroni method was used (Holm, 1979).

Two-way random intra-class correlations (ICC) for average measures on the pronunciation ratings indicated a high level of consistency between the pronunciation raters. For the group of native-English raters, the average measure ICC was high ( $r = .824$ ) with a 95% confidence interval from .762 to .875 ( $F(82,738) = 5.679$ ,  $p < .001$ ). For the group of native-French raters, the average measure ICC was high ( $r = .849$ ) with a 95% confidence interval from .795 to .893 ( $F(82,738) = 6.607$ ,  $p < .001$ ). Given that inter-rater reliability was high ( $r > .75$ ), an average pronunciation rating for each participant was used in the calculation of the group mean-score for each language (English and French).

Consistent with standard practice in reaction time (RT) research, only RTs for correct responses in the grammar knowledge task were included in the analyses. Prior to completing all t-tests, outliers were assessed by inspection of a boxplot, normality was assessed using Shapiro-Wilk's test, and homogeneity of variances was assessed by Levene's test. A single outlying score was identified in the data, as assessed by inspection of a box plot for extreme values. In this case, one participant in the EFI group scored 78 (out of 100) on the English grammar knowledge task,

which was lower than the first quartile minus three times the interquartile range ( $Q1 - 3 * IQR$ ). In order to eliminate the influence of this outlier on the data set, the score was removed. Further, the assumption of homogeneity of variances was violated for the French grammar knowledge RT measures. Accordingly, the results on this measure underwent the Welch-Satterthwaite correction for degrees of freedom (Satterthwaite, 1946; Welch, 1947). In the following sections, results are reported before and after the Holm-Bonferroni adjustment was applied.

### 3.1.1 EFI and LFI performance on French language proficiency measures

The mean scores and standard deviations for EFI and LFI groups, along with the results of the t-tests for the French measures are presented in Table 4.

**Table 4: Mean scores (and standard deviations) for EFI and LFI on French language proficiency measures with independent samples t-test results**

	<b>EFI</b> <i>M</i> ( <i>SD</i> )	<b>LFI</b> <i>M</i> ( <i>SD</i> )	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>
<b>French Measures</b>						
Receptive vocabulary knowledge	88.46 (15.27)	83.52 (18.37)	81	1.293	.200	0.29
Grammar knowledge	44.02 (3.19)	43.44 (3.56)	81	.738	.463	0.17
- RT (m/s)	4644.93 (1392.40)	4248.74 (984.79)	69.98	1.325	.140	0.32
Listening comprehension	88.13 (9.54)	89.00 (10.22)	81	-.383	.703	- 0.09
Pronunciation	4.84 (0.65)	4.81 (0.76)	81	.209	.835	0.04

For all measures of French language proficiency, there were no significant differences found in group performance between EFI and LFI.

### **3.1.1.1 Post-hoc analyses on French grammar knowledge task**

In addition to the overall grammar knowledge and RT measures, a post-hoc comparison between EFI and LFI groups on each sentence construction type and RT for each sentence construction type, was completed. Analyses on these data were carried out by means of a series of non-parametric Independent Sample Mann-Whitney U tests, as it was found that the scores and RTs for each sentence construction type was not normally distributed. Distributions of accuracy scores and RTs between EFI and LFI groups were similar for each sentence construction, as assessed by visual inspection. The mean accuracy scores, RTs, and standard deviations for all sentence construction types, along with results of the Independent Mann Whitney U tests, are presented in Table 5.

**Table 5: EFI and LFI performance on each sentence-type within the French grammar knowledge task**

	<b>EFI</b>	<b>LFI</b>	<i>Mann-Whitney</i>		
	<i>M</i>	<i>M</i>	<i>U</i>	<i>p</i>	<i>d</i>
	( <i>SD</i> )	( <i>SD</i> )			
<b>Accuracy</b>					
Adjacency ( <i>n</i> =15)	12.63 (1.30)	11.93 (1.49)	550.00	.040	0.50
En-avant ( <i>n</i> =7)	2.98 (1.19)	3.18 (1.11)	829.00	.462	- 0.17
Prenominal past participle ( <i>n</i> =6)	3.29 (1.11)	3.19 (0.88)	696.00	.542	0.10
<i>That</i> -Trace ( <i>n</i> =10)	5.95 (1.70)	6.48 (1.32)	886.00	.198	- 0.35
De + modifier ( <i>n</i> =16)	8.52 (1.53)	8.67 (1.80)	813.50	.567	- 0.09
Ce vs. il (elle) ( <i>n</i> =10)	5.70 (1.32)	5.44 (1.69)	691.50	.522	0.17
Middle-voice ( <i>n</i> =8)	4.96 (1.31)	4.56 (1.03)	590.00	.096	0.34
<b>Reaction Time</b>					
Adjacency RT (m/s)	3942.57 (1261.2)	3741.42 (977.38)	713.00	.676	0.17
En-avant RT (m/s)	6539.79 (2638.53)	6069.64 (3042.56)	682.00	.472	0.17
Prenominal past participle RT (m/s)	5048.92 (2047.87)	4041.84 (1420.92)	540.00	.036	0.57
<i>That</i> -Trace RT (m/s)	4841.69 (1760.77)	4301.91 (1098.04)	594.00	.115	0.37
De + modifier RT (m/s)	4092.90 (1432.97)	3656.57 (918.25)	630.00	.221	0.36
Ce vs. il (elle) RT (m/s)	5663.34 (2067.60)	4907.70 (1208.38)	618.00	.180	0.45
Middle-voice RT (m/s)	4441.96 (1533.22)	4361.49 (1319.28)	762.00	.953	0.06

Results revealed that the EFI group performed significantly better than the LFI group on the Adjacency sentence-construction, and that the LFI group responded significantly faster than the EFI group on the prenominal past-participle sentence construction. However, after the Holm-Bonferroni adjustment for alpha was applied to each family of tests (accuracy or RT), no statistically significant differences between EFI and LFI groups were detected ( $p > .007$ ).

An additional post-hoc analysis was carried out to compare the performance of EFI and LFI groups on the French grammar knowledge task for only incongruent trials (i.e. identification of sentences which were ungrammatical,  $n = 22$ ). Similar to the comparative result concerning the grammar knowledge scores across all trials, the result of this analysis revealed no significant difference between EFI ( $M=13.80$ ,  $SD = 3.80$ ) and LFI ( $M=13.26$ ,  $SD = 3.88$ ) groups,  $t(1, 81) = .608$ ,  $p = .545$ ,  $d = 0.14$ .

### **3.1.2 EFI and LFI performance on English language proficiency measures**

The mean scores and standard deviations for EFI and LFI groups, along with the results of the t-tests for the English measures are presented in Table 6.

**Table 6: Mean scores (and standard deviations) for EFI and LFI on English language proficiency measures with independent samples t-test results**

	<b>EFI</b> <i>M</i> ( <i>SD</i> )	<b>LFI</b> <i>M</i> ( <i>SD</i> )	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>
<b>English Measures</b>						
Receptive vocabulary knowledge	108.57 (10.81)	108.85 (11.38)	81	-.109	.914	- 0.03
Grammar knowledge	93.00 (3.29)	94.11 (3.04)	80	-1.47	.146	- 0.35
- RT (m/s)	3205.56 (792.06)	2968.65 (621.23)	81	1.361	.177	0.33
Listening comprehension	525.14 (9.99)	525.48 (9.37)	81	-.148	.883	- 0.04
Pronunciation	8.00 (0.55)	7.64 (0.80)	81	2.385	.019	0.53

A statistically significant difference was found only on the English pronunciation measure, such that the EFI group sounded more native-like than the LFI group. For all other English language proficiency measures, no significant differences in performance were found.

After the Holm-Bonferroni correction for alpha was applied to the English measures ( $p = .01$ ), no statistically significant differences were detected between EFI and LFI groups.

### 3.2 Dividing the sample based on home-language

The lower pronunciation score in English of the LFI group was surprising and prompted a further examination of the sample. Responses on the language background questionnaire revealed that the EFI group contained proportionally more participants whose home-language (and L1) was English only (53.5%, or 30 out of 56), when compared with the LFI group (22.2%, or 6 out of 27).

As a result, the sample was divided into two groups based on home-language: bilinguals (participants who spoke only English at home), and multilinguals (participants who spoke a language other than English at home). In total, there were 36 bilinguals (30 EFI, 6 LFI) and 47 multilinguals (26 EFI, 21 LFI) within the sample. In order to determine if there were differences in the language outcomes of bilinguals and multilinguals, a series of independent samples t-tests were carried out on the mean scores for each group. The mean scores, standard deviations, and results of the t-tests comparing bilinguals and multilinguals on the French and English measures are presented in Tables 7 and 8, respectively.

**Table 7: Mean scores (and standard deviations) for bilinguals and multilinguals on French language proficiency measures with independent samples t-test results**

	<b>Bilinguals</b>	<b>Multilinguals</b>				
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>
<b>French Measures</b>						
Receptive vocabulary knowledge	84.08 (15.95)	88.98 (16.58)	81	-1.355	.179	- 0.30
Grammar knowledge	43.94 (3.26)	43.75 (3.37)	81	0.271	.787	0.06
- RT (m/s)	4735.04 (1309.35)	4348.31 (1248.08)	81	1.370	.175	0.30
Listening comprehension	85.61 (9.27)	90.55 (9.59)	81	-2.361	.021	- 0.53
Pronunciation	4.65 (0.59)	4.96 (0.73)	81	-2.057	.043	- 0.46

**Table 8: Mean scores (and standard deviations) for bilinguals and multilinguals on English language proficiency measures with independent samples t-test results**

	<b>Bilinguals</b>	<b>Multilinguals</b>				
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>
<b>English Measures</b>						
Receptive vocabulary knowledge	107.61 (9.91)	109.47 (11.70)	81	-0.765	.447	- 0.17
Grammar knowledge	92.43 (3.34)	94.06 (3.02)	80	-2.320	.023	- 0.52
- RT (m/s)	3348.74 (945.44)	2940.43 (588.67)	55.17	2.597	.006	0.70
Listening comprehension	526.72 (10.33)	524.13 (9.20)	81	1.207	.231	0.26
Pronunciation	8.18 (0.37)	7.66 (0.74)	70.93	4.117	.000	0.98

On the French listening comprehension task, multilinguals scored significantly higher than bilinguals. Based on the French pronunciation ratings, it was revealed that both groups tended to score toward the non-native-like end of the scale, but the multilinguals' pronunciation of French was significantly closer to native-like than the bilinguals' pronunciation of French. After the Holm-Bonferroni adjustment for alpha was implemented ( $p = .01$ ), none of these results were statistically significant.

On the English grammar knowledge measure, multilinguals scored significantly higher than bilinguals. In addition, multilinguals had significantly faster RTs than bilinguals. Further, bilinguals' pronunciation of English sounded significantly more native-like than that of the multilinguals. After the Holm-Bonferroni adjustment for alpha was implemented ( $p = .01$ ), only two of these results, pronunciation and RT, remained statistically significant.

### 3.2.1 Post-hoc analyses on language aptitude and motivation

In order to further investigate the differences in French language outcomes between bilinguals and multilinguals, a series of post-hoc t-tests on language aptitude and motivation were carried out. More detailed analyses on language aptitude and motivation were conducted and reported in a related study by Kwan (2015), which investigated the relationship between language aptitude and motivation on the language outcomes with the same sample of students. In this study, the Pimsleur Language Aptitude Battery (PLAB, Pimsleur, Reed & Stansfield, 2004) tested aspects of language aptitude, and the Attitude Motivation Test Battery (AMTB, Gardner, 1985) evaluated motivation for foreign language learning.

From the PLAB, three aspects of language aptitude were analyzed: verbal ability (i.e. knowledge of words and ability to reason analytically with verbal materials), auditory ability (i.e. reception and processing of information via the ear), and a total aptitude score. From the AMTB, four aspects of attitude/motivation were measured: integrativeness (i.e. willingness to identify with a different language community), motivation (i.e. goal-directed behavior), attitude toward the learning situation, and an attitude/motivation index score. An in-depth explanation of each measure can be found in Kwan (2015). The results of these post-hoc analyses are summarized in Table 9.

**Table 9: Post-hoc analyses on bilinguals and multilinguals for language aptitude and motivation**

	<b>Bilinguals</b>	<b>Multilinguals</b>				
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>
<b>Language Aptitude Measure</b>						
Verbal ability	24.17 (5.40)	27.87 (5.17)	81	-3.162	.002	-0.70
Auditory Ability	39.61 (4.90)	46.02 (5.91)	81	-5.267	<.001	-1.17
Total Aptitude	81.47 (9.33)	93.85 (10.25)	81	-8.668	<.001	-1.93
<b>Attitude/Motivation Measure</b>						
Integrativeness	159.86 (24.51)	179.23 (17.46)	81	-4.204	<.001	-0.93
Communicative motivation	86.39 (18.44)	100.68 (15.05)	81	-3.887	<.001	-0.86
Attitude toward learning situation	109.35 (17.56)	115.28 (16.16)	81	-1.564	.122	-0.35
Attitude/Motivation index	400.03 (49.77)	442.43 (46.04)	81	-3.935	<.001	-0.87

The results of the t-tests revealed statistically significant differences between bilinguals and multilinguals for all measures of language aptitude. After the Holm-Bonferroni adjustment was applied ( $p = .016$ ), all of these differences remained statistically significant.

In addition, statistically significant differences between bilinguals and multilinguals were found on all measures of attitude/motivation, except attitude towards the learning situation. After the Holm-Bonferroni adjustment was applied ( $p = 0.0125$ ), all of these differences remained statistically significant.

### 3.3 Summary of results

Analyses revealed that EFI and LFI students performed equivalently on all language outcome measures of French and English, except for pronunciation in English, on which the EFI group performed significantly better than the LFI group. When dividing the sample based on home-language, multilinguals performed better than bilinguals on certain measures of French and English language outcomes. In French, the multilinguals performed better than the bilinguals on measures of listening comprehension and pronunciation. In English, multilinguals performed significantly faster and more accurately than the bilinguals on the English grammar knowledge task. However, the bilinguals sounded more native-like than the multilinguals on the English pronunciation measure. Post-hoc analyses further revealed that multilinguals had higher language aptitude and motivation to learn a foreign language, than bilinguals.

However, after conducting the Holm-Bonferroni adjustment for alpha, only two of the results concerning language outcomes still held: bilinguals performed better than multilinguals on the English pronunciation measure, and multilinguals responded faster than bilinguals on the English grammar knowledge task. The application of the adjustment resulted in a loss of statistical power, and an increased likelihood of Type II error (i.e. accepting the null-hypothesis when it is actually false). Due to the exploratory nature of the study, the Holm-Bonferroni adjustment for alpha was judged to be too conservative; thus, the uncorrected results will be discussed in the next chapter.

## **Chapter 4: Discussion**

This study compared the language proficiency of EFI and LFI students nearing the end of high school. The main purpose of this investigation was to determine if there were differences between EFI and LFI groups on French and English language outcomes. This study was also intended to provide more general information about the existence of a sensitive period for language acquisition in the context of formal education. In the following sections, the findings of the present study are discussed in relation to the original hypotheses, and within the framework of the central research question. The implications of these results and directions for future research are also discussed.

### **4.1 French language proficiency outcomes**

Firstly, it was expected that EFI and LFI groups would perform equivalently on the measures of French receptive vocabulary knowledge, grammar knowledge and listening comprehension. As predicted, no significant differences were found between EFI and LFI groups on these measures. This finding corroborates the results reported by Morrison and Pawley (1986) and Turnbull et al. (1998), suggesting that performance on these measures may not be affected by starting-age or by a greater amount of exposure to French in the context of school immersion. However, the presence of multilingual students in the current study, who generally performed better than the bilingual students, may provide an alternative explanation for the apparent lack of age effects.

Of particular note is the finding that EFI and LFI groups performed equivalently on the French listening comprehension task, which was the only language measure that previously revealed an EFI-advantage in Grade 7 (Day & Shapson, 1988) and in Grade 9 (Genesee, 1981). Day and Shapson (1988) speculated that listening comprehension may be acquired more slowly than other language skills and “if so, late immersion students may attain equivalence with their early immersion counterparts in future grades” (p. 302). The results of the present study support this hypothesis, and highlight the importance of providing sufficient time for LFI students to develop listening comprehension skills in French, prior to evaluating them comparatively with EFI students.

However, when the sample was divided by home-language, into groups of multilinguals and bilinguals, it was found that multilinguals performed better than bilinguals on the listening comprehension measure. The large presence of multilinguals within the sample (56.6%) distinguishes the sample in the present study from that of Day and Shapson’s (1988) study, in which the sample reported that English was the language most often spoken at home. Further, since there were proportionally more multilinguals in the LFI group (77.8%) than in the EFI group (46.4%), it is possible that any EFI-advantage on this measure may have been washed out by a multilingual-advantage. Thus, differences between sample characteristics (viz., home-language) of this study in contrast to Day and Shapson’s study may explain the discrepant findings concerning an EFI-advantage on the listening comprehension measure.

The finding that multilinguals performed better than bilinguals on the listening comprehension task is explained by their heightened aptitude for processing auditory information and motivation to learn French, when compared to bilinguals. In addition, it is possible that

multilinguals demonstrated a form of additive multilingualism, which helped them excel on this task; that is, a positive transfer from second-language learning, to learning French as an L3. This line of thinking is consistent with previous studies on multilingual students in French immersion programs, which indicated advantages for both general language proficiency, and specific language skills, such as listening comprehension (e.g. Edwards et al., 1977; Wightman, 1981) and perceptual discrimination in a third-language (e.g. Enomoto, 1994).

Secondly, it was expected that the EFI group would perform better than the LFI group on the French pronunciation measure. Long (2013b) recently argued that evidence for a sensitive period in a formal education context may be more apparent if pronunciation was assessed. However, despite an earlier start, the EFI group in the present study did not appear to have superior French pronunciation than the LFI group. The contrast between the findings of the present study and that of previous work comparing EFI and LFI groups on French speaking ability (Morrison & Pawley, 1986; Turnbull et al., 1998) are likely due to disparities between the nature of the speaking tasks and their evaluation. The present study utilized a paragraph to prompt oral output. By using this method, differences in pronunciation were better isolated than they would have been in a spontaneous speech sample, which is often characterized by aspects of language that could directly or indirectly confound ratings of pronunciation (e.g. grammatical errors, use of vocabulary).

However, Morrison and Pawley (1986) evaluated French speaking ability using semi-directed interviews, which simulate spontaneous speech. Further, their speech samples were evaluated on aspects of language that are outside of the scope of pronunciation/accent (e.g. communicative effectiveness, vocabulary, and complexity of spoken language). Similarly, the

sentence-repetition task used by Turnbull et al. (1998), in which participants would have to repeat a sentence verbatim, likely drew on aspects of auditory-verbal working memory and speech-imitation ability, which are also linguistic skills that are beyond the scope of pronunciation.

Further, it is possible that participant home-language affected EFI and LFI comparisons on the French pronunciation task. Though it is not explicitly stated in either study (i.e. Morrison & Pawley, 1986; Turnbull et al., 1998), it is assumed that English was the home-language of most participants in these studies. However, as mentioned previously, there were a large number of participants who were multilingual in the present study. By virtue of the finding that multilinguals performed better than bilinguals on the French pronunciation task, the possible interference of this factor cannot be ruled out.

The most likely reason that multilinguals sounded more native-like in French than the bilinguals is that they had higher integrative motivation and oral willingness to communicate. Post-hoc analyses on measures of language aptitude demonstrated that multilinguals had a heightened aptitude score for verbal ability, in addition to integrative motivation and goal-directed behavior in general. This argument and finding is congruent with recent work by Mady (2015), which demonstrated that immigrants in French immersion who performed better than their Canadian-born peers on a speaking task, also demonstrated higher levels of integrative motivation and willingness to communicate.

Alternatively, it is possible that multilinguals achieved higher ratings of pronunciation than bilinguals due to an accent-bias. Namely, it is possible that the native-French raters judged the pronunciation of French by native-English-speakers (bilinguals) more harshly than they did

for speakers of other languages (multilinguals), possibly because the raters were more familiar with English accented French than the accents of other languages. When asked about the most salient characteristics of accented-speech, the native-French raters would often emulate anglicized French. Further, some raters specifically mentioned aspects of pronunciation which led to lower native-like ratings, such as aspirated voiceless stops [p t k]<sup>1</sup> and vowel distortions, which are documented features of English accented French (Gavac, 2006).

In sum, despite the EFI group having six more years of French instruction early on, no significant differences were found between EFI and LFI groups on all measures of French language proficiency, suggesting evidence against the notion of a sensitive period. However, these results are interpreted with caution as it is possible that the EFI and LFI comparisons were influenced by participant language-background (i.e. home-language). Analyses revealed that participants who were multilingual performed better than participants who were bilingual on measures of French listening comprehension and pronunciation. Taken together, it appears that French language achievement is affected more by previous language-learning experience and less by age of entry into French immersion.

## **4.2 English language proficiency outcomes**

Provided that there were no differences in age of exposure to English between the two groups, it was expected that EFI and LFI groups would achieve parity on all English language proficiency measures. This hypothesis was supported for all measures, except pronunciation, on which the EFI group performed better than the LFI group. This prompted a further exploration of

---

<sup>1</sup> In French, these sounds are unaspirated, which means that they are not produced with an audible puff of breath.

the sample, which revealed that the EFI group contained proportionally more native-English speakers (bilinguals), while the LFI group contained more individuals whose L1 included languages other than English (multilinguals). As reported on the language history questionnaire, multilinguals were first exposed to English between 0 and 6 years of age, which is considered to be within the sensitive period for L2 pronunciation (Long, 1990, 2013a, Granena & Long, 2013). However, as stated by Long (2013a), exposure to the L2 during the sensitive period does not necessarily guarantee that one will achieve native-like levels of pronunciation. Evidence from previous studies on L2 pronunciation by early L2-learners (e.g. Flege, 1999, Flege, Frieda, & Nozawa, 1997, Yeni-Komishian, Flege & Liu, 2000) has shown that accented speech is more likely to appear in cases where the young learners have attained high levels of proficiency in the L1 and/or continue to use the L1 in significant proportions in their everyday lives (Long, 2013a). Although the proficiency level of the non-English L1 was not tested in the present study, responses on the language background questionnaire revealed that multilinguals continued to use their non-English L1 regularly and on a daily basis at home. Thus, the finding that multilinguals sounded less native-like than bilinguals on the English pronunciation task is in line with the aforementioned studies.

In addition, it was found that multilinguals performed better than bilinguals on the English grammaticality judgment test. Albeit highly speculative, it is possible that the superior performance of multilinguals on this task is attributed to cognitive advantages gained as a result of their greater language experience. These cognitive advantages may include (a) heightened

metalinguistic awareness,<sup>2</sup> which would likely elevate accuracy scores, and (b) greater executive control<sup>3</sup>, which would lead to faster processing time. This notion and interpretation stems from a large body of research which has shown that bilinguals demonstrate linguistic and cognitive advantages relative to monolinguals, including metalinguistic awareness and executive control (e.g. Bialystok & Barac, 2012; Bialystok, Barac, Blaye & Poulin-Dubois, 2010; Carlson & Meltzoff, 2008; Lazaruk, 2007; Galambos & Goldin-Meadow, 1990; Reder, Marec-Breton, Gombert & Demont, 2013). However, as we did not test the L1 proficiency of the multilinguals, we cannot be certain that they were balanced bilinguals at time of first exposure to French.

Alternatively, it is possible that multilinguals responded faster than bilinguals due to differences in a sociolinguistic capacity (i.e. culture). As stated by Petrova, Wentura and Fu (2013), the notion that cultural background influences executive control and attentional processing is not new. Since bilinguals reported speaking only English at home, it is inferred that this group consisted largely of persons of a Western Canadian culture. Correspondingly, responses on the background questionnaire revealed that multilinguals were predominantly (68.08%) of an East Asian culture. Previous comparisons of these two cultural groups have suggested that children of an East Asian cultural background develop and master executive control earlier than children of a Western culture (e.g. Lewis et al. 2009; Sabbagh et al. 2006; Oh & Lewis, 2008; Liu, Wellman, Tardif & Sabbagh, 2009).

---

<sup>2</sup> The ability to distance one's self from the content of language, in order to pay attention to the structural features and properties of words and sentences (Reder et al., 2013)

<sup>3</sup> An umbrella term for the cognitive processes involved in working memory, shifting of mental states (cognitive flexibility in task-monitoring, task execution, response selection) and inhibition (Cattaneo et al., 2014; Bialystok, 2009; Miyake et al., 2000)

This phenomenon in East Asian populations has primarily been associated with cultural values and linguistic practices. In line with their Confucian heritage, parents of East Asian culture, including those living in Western societies, place greater importance on obedience, respect for others, behavioral self-control, and emotional regulation than parents in Western cultures (Chen et al. 1998; Park & Cheah, 2005; Oh & Lewis, 2008), which consequently create more opportunities to practice executive control (Oh & Lewis, 2008; Lewis et al., 2009). Further, the speech of East Asian parents to preschoolers contain more action verbs (Choi & Bowerman, 1991), which may implicitly emphasize the role of behavioral control early on in life (Kim, McGregor & Thompson, 2000; Lewis et al., 2009; Oh & Lewis, 2008). Therefore, bearing in mind the dominant cultures of the respective multilingual and bilingual groups, a viable explanation for the faster response times by multilinguals is an effect of culture.

### **4.3 Revisiting the notion of a sensitive period for language learning**

By investigating the French language outcomes of EFI and LFI students, one of the aims of the present study was to investigate the effect of age of entry into French immersion on French language outcomes. As mentioned previously, results revealed no significant differences on French language outcomes between EFI and LFI groups. This finding may be interpreted in two ways concerning the notion of a sensitive period. On the one hand, this finding may indicate that the age of entry into EFI (approximately 5 or 6) and LFI (approximately 12 or 13) are within the putative sensitive period. Thus, it would appear that this period culminates beyond the age of 12 or 13 (i.e. when LFI students enter French immersion). On the other hand, this finding may

suggest evidence against the notion of a sensitive period or age effects in the context of French immersion.

Although the age at which the critical/sensitive period culminates is a source of continual debate, the second interpretation is more convincing as it is in line with the more recent findings comparing early and late L2 learners in the broader field of foreign language acquisition in an educational setting (i.e. Al-Thubaiti, 2010; Cenoz, 2003; Muñoz, 2011). As mentioned by Muñoz (2008), language learning in the context of education is subject to many factors which may better explain ultimate attainment in an L2 than the age of first exposure. These factors include the limited amount of exposure/instruction (i.e. only within school-hours), the limited source of input (i.e. the teachers and peers in classroom), and inconsistency of the quality of input (i.e. as teachers' and peers' French proficiency varies). Although French immersion fosters richer and more natural input than more traditional approaches to foreign language instruction (i.e. where language itself is the topic of instruction, rather than the medium of instruction), it does not appear to be sufficient to trigger early advantages observed more readily, though still inconsistently, in naturalistic language-learning contexts.

However, it is also possible that age effects were concealed by the large proportion of multilingual participants within the sample. In recent years, the language background of French immersion student populations has become increasingly heterogeneous, and the assumption that French immersion students are monolingual-English speakers outside of the classroom is no longer feasible (Swain & Lapkin, 2005). This was certainly the case in the present study, in which the majority of participants spoke languages other than English at home. Considering that it was found that these multilingual participants performed as well as, if not better than,

bilinguals on all French language measures, it is possible that the heightened proportion of multilinguals in the LFI group positively influenced the scores of the LFI group to achieve parity with those of the EFI group.

While the findings from this study suggest that age effects may be less robust than previously thought, it is not possible to conclude that age of entry into French immersion bears no effect on language outcomes. This is primarily because the disproportion of multilinguals across EFI and LFI groups may have obscured an EFI-advantage, should one exist. Nevertheless, the results of this study have shed further light on the impacts and influences of factors other than age, on language proficiency. These would include one's linguistic and cultural background, language aptitude, and motivation to learn a foreign language. Altogether, this study underlines the complex nature of language proficiency, and gives rise to further study on the impact of these factors in addition to age on second or subsequent language acquisition, in the context of formal education.

#### **4.4 Strengths and limitations**

The present study was the first in nearly 20 years to compare the language outcomes of EFI and LFI groups nearing the time of graduation, and one which reflected the changing demographics of enrollment in French immersion. Further, it was the only study which compared EFI and LFI groups nearing the time of graduation within the province of British Columbia. In line with recent methodological recommendations, students in this study were tested in both languages of their formal education: French and English, rather than only the target language

(French). Finally, this study compared the French language outcomes of early and late learners of French to one another, rather than to the performance of monolingual native-French speakers.

However, this work was not without limitation. The present research took place within a single school district in Western Canada, and French immersion programs differ in program structure (i.e. concerning percentage of French instructional time, subjects offered in French) across districts and provinces. Thus, it is possible that the generalizability of findings is limited to students in French immersion programs that are similar in the matter of program structure.

Further, in the interest of not using too much of the class-time of students and teachers, only a limited number of language outcomes were measured. For the same reason, multilingual students were tested only in the two languages they were educated in, English and French, even though several participants reported being users of other languages at home. Furthermore, it was not possible to control for L1-typologies when comparing EFI and LFI groups because of the heterogeneity across the sample, which resulted in small numbers of each L1-type.

The presence of multilinguals in this study, and their uneven distribution across the EFI and LFI groups, is another limitation of this study. However, the large number of multilinguals across the sample is representative of the linguistic and cultural diversity of the population within the city where this research took place, in which approximately 63% of residents reported having an L1 other than English in the latest Canadian Census (Statistics Canada, 2012b). There are also several practical reasons which could explain the heightened proportion of multilinguals in an LFI program. For instance, multilingual students are often immigrants, and they may have entered the Canadian school-system at an age which is beyond EFI entry in kindergarten or Grade 1. Another reason could be that parents who speak languages other than English at home

may have chosen not to enroll their child in an EFI program, possibly because they felt it might not give their child sufficient opportunity to learn English (Genesee & Jared, 2008). For these reasons, among others, it may be more appropriate to consider the greater proportion of multilinguals to be an inherent factor of LFI.

#### **4.5 Future directions**

The growing number of students from culturally and linguistically diverse backgrounds in French immersion deserves careful consideration in future studies. In urban areas, where immigrant populations represent over 40% of the total population (Statistics Canada, 2012a), increasing numbers of students with bilingual and multilingual backgrounds are likely to enroll in French immersion programs. Educators, parents and students would benefit from information guiding their understanding of the various French immersion programs, which would further enable students to consider their options for success in language learning and academic achievement.

In light of the results of the present study, further exploration comparing the language outcomes of EFI and LFI students on a larger-scale is needed in order to fully examine potential effects of age of entry into French immersion. Future studies investigating age effects in the context of French immersion should gather comprehensive language background data, and control for L1 by only including participants who are speakers of the same L1 or by categorizing speakers of the same L1 *a priori*. Further research of a longitudinal nature would also allow researchers to examine language achievement for persons of different linguistic profiles and L1-typologies over time. Since language achievement also appears to be linked with language

aptitude and motivation to learn a foreign language, these factors should not be ignored in future studies.

French immersion programs have significantly evolved over time in terms of the program delivery, the curriculum used, and the demographics of the student body. While the present study shed greater light on demographic aspects, it would be beneficial for future work to focus on pedagogical aspects of French immersion. An investigation of the quantitative and qualitative features of French input from teachers could mutually inform theories of second or subsequent language acquisition, and promote our understanding of language development through French immersion programs.

#### **4.6 Conclusion**

This study was the first to compare the language outcomes of EFI and LFI students nearing the time of high-school graduation in the province of British Columbia. In contrast with eastern regions of Canada, French is spoken less widely within the communities of British Columbia. Thus, the language achievement of EFI and LFI students from this province is more likely to represent the French language skills gained exclusively from the respective French immersion programs. Since EFI and LFI students enter French immersion at different ages (i.e. approximately 5 and 12, respectively), this comparative examination of EFI and LFI language outcomes aimed to shed further light on the potential effect of age on language achievement in the context of French immersion.

No differences were found between EFI and LFI groups on measures of French language proficiency. Thus, the findings do not appear to support the notion of a critical or sensitive period

within the context of L2 acquisition in French immersion. While these results are congruent with findings in the broader field of foreign language education, no conclusive statements regarding the putative sensitive period can be made from this study, since age effects may have been masked by the better performance of multilinguals, the majority of whom were in LFI.

In the context of French immersion, it appears that factors other than starting-age, such as language background, language aptitude, and motivation, are more influential on language achievement. This study contributes new evidence to a growing body of research in support of the language achievement of multilingual students in French immersion classrooms. Considering the growth in the number of multilingual students in French immersion programs and the contrastive performance on language outcomes between bilinguals and multilinguals in the present study, the cultural and linguistic diversity of French immersion students deserves careful consideration in future research.

## References

- Abrahamsson, N. (2012). Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition*, 34(02), 187-214.
- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249-306.
- Al-Thubaiti, K. A. (2010). *Age effects in a minimal input setting on the acquisition of English morpho-syntactic and semantic properties by L1 speakers of Arabic* (Doctoral dissertation). University of Essex.
- Ardila, A., & Ramos, E. (2007). *Speech and language disorders in bilinguals*. Nova Publishers.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(01), 3-11.
- Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, 122(1), 67-73.
- Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*, 11(4), 485-508.
- Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 68, 706-755.
- Birdsong, D. (2005). Nativelikeness and non-nativelikeness in L2A research. *International Review of Applied Linguistics in Language Teaching*, 43(4), 319-328.
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56, 9.
- Birdsong, D. (2009). Age and the end state of second language acquisition. *The new handbook of*

*second language acquisition. Amsterdam: Elsevier.*

- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and language, 44*(2), 235-249.
- Bruck, M. (1975). Assessing functional bilingualism within a bilingual program: The St. Lambert project at grade eight. In *TESOL Convention, Los Angeles, California.*
- Bruck, M., Lambert, W. E., & Tucker, G. R. (1976). Alternative forms of immersion for second language teaching. *NABE, 1*(3), 33-48.
- Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental science, 11*(2), 282-298.
- Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code mixing. *Age and the acquisition of English as a foreign language, 77-93.*
- Chen, X., Hastings, P. D., Rubin, K. H., Chen, H., Cen, G., & Stewart, S. L. (1998). Child-rearing attitudes and behavioral inhibition in Chinese and Canadian toddlers: a cross-cultural study. *Developmental psychology, 34*(4), 677.
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*(1), 83-121.
- Chomsky, N. (1965). Aspects of the theory of syntax Cambridge. *Multilingual Matters: MIT Press.*
- Cook, V. (2002). Background to the L2 user. *Portraits of the L2 user, 1*, 1-28.
- Cziko, G. A. (1977). Early And Late French Immersion: A Comparison of Children at Grade Seven.
- Day, E. M., & Shapson, S. (1988). A comparison study of early and late French immersion programs in British Columbia. *Canadian Journal of Education/Revue canadienne de l'education, 290-305.*

- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499-533.
- DeKeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, 62(s2), 189-200.
- DeKeyser, R. (2013). Age effects in second language learning: stepping stones toward better understanding. *Language Learning*, 63(1), 52-67.
- DeKeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, 31(03), 413-438.
- Dillon, G., Lamb, S.J. & Grayson, A., 2010. Atypical child development. In: Banyard, P., Davies, M.N.O., Norman, C., & Winder, B. (Eds), *Essential psychology: a concise introduction*. London: Sage, 2010, pp. 296-313.
- Dunn, L., & Dunn, D. (2007). *PPVT-4: Peabody picture vocabulary test – (4th ed.)*. Minneapolis, MN.: Pearson Assessments.
- Dunn, L., Dunn, L., & Theriault-Whalen, C. (1993). *Échelle de vocabulaire en images Peabody – (3rd ed.)*. Toronto: PSYCAN.
- Edwards, H.P., Doutriaux, C.W., McCarrey, H., & Fu, L. (1977) *Evaluation of the federally and provincially funded extensions of the second language programs in the school of the Ottawa Roman Catholic Separate School Board*. Ottawa: Ottawa Roman Catholic Separate School Board.
- Flege, J. E. (1999). Age of learning and second language speech. *Second language acquisition and the critical period hypothesis*, 101-131.
- Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the

- pronunciation of an L2. *Journal of Phonetics*, 25(2), 169-186.
- Galambos, S. J., & Goldin-Meadow, S. (1990). The effects of learning two languages on levels of metalinguistic awareness. *Cognition*, 34(1), 1-56.
- Gardner, R.C. (1985). *The Attitude Motivation Test Battery: Technical Report*. London, Ontario, Canada: University of Western Ontario, Department of Psychology.
- Genesee, F. (1981). A comparison of early and late second language learning. *Canadian Journal of Behavioural Science/Revue canadienne des Sciences du comportement*, 13(2), 115-128.
- Genesee, F., & Jared, D. (2008). Literacy development in early french immersion programs. *Canadian Psychology/Psychologie Canadienne*, 49(2), 140-147. doi:10.1037/0708-5591.49.2.140
- Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311-343.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and language*, 36(1), 3-15.
- Herschensohn, J. R. (2007). *Language development and age*. Cambridge: Cambridge University Press.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65-70.
- Huang, B. (2014). The effects of age on second language grammar and speech production. *Journal of Psycholinguistic Research*, 43(4), 397-420
- Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2

pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, 10(2), 135-159.

Johnson, J. S., & Newport, E. (1989). Critical period effects in second language learning: The influence of the maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English-and Korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language*, 27(02), 225-254.

Kwan, K. (2015) *Success in early and late French Immersion in British Columbia: a study on language aptitude and motivation* (Master's thesis). University of British Columbia.

Lamendella, J. (1977). General principles of neurofunctional organization and their manifestation in primary and nonprimary language acquisition. *Language Learning*, 27, 155-196.

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research*, 24(1), 35-63.

Lazaruk, W. (2007). Linguistic, Academic, and Cognitive Benefits of French Immersion. *Canadian Modern Language Review*, 21(6), 605-628.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Lepage, J., & Corbeil, J. (2013). The evolution of English–French bilingualism in Canada from 1961-2011. Insights on Canadian Society, 75 006 X. <http://www.statcan.gc.ca/pub/75-006-x/2013001/article/11795-eng.pdf>

Lewis, C., Koyasu, M., Oh, S., Ogawa, A., Short, B., & Huang, Z. (2009). Culture, executive function, and social understanding. *New directions for child and adolescent*

*development*, 2009(123), 69-85.

- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(03), 673-680.
- Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: a meta-analysis of false-belief understanding across cultures and languages. *Developmental psychology*, 44(2), 523.
- Long, M. (2013a). Maturation constraints on child and adult SLA. *Sensitive periods, language aptitude, and ultimate L2 attainment*, 35, 3-41.
- Long, M. (2013b). Some implications of research findings on sensitive periods in language learning for educational policy and practice. *Sensitive periods, language aptitude, and ultimate L2 attainment*, 35, 259-271.
- Mady, C. (2015). Immigrants outperform Canadian-born groups in French Immersion: Examining factors that influence their achievement. *International Journal of Multilingualism*, 12(3), 298-311. doi:10.1080/14790718.2014.967252
- Marinova-Todd, S. (2003). *Comprehensive analysis of ultimate attainment in adult second language acquisition* (Doctoral dissertation). Harvard University.
- Marinova-Todd, S., Marshall, D., & Snow, C. (2000). Three misconceptions about age and second language acquisition. *TESOL Quarterly*, 34, 9-34.
- McDonald, J. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381-401.
- 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(1), 49-100.

Morrison, F., & Pawley, C. (1986). *Evaluation of the Second Language Learning (French) Programs in the Schools of the Ottawa and Carleton Boards of Education, Volume I. French Proficiency of Immersion Students at the Grade 12 Level*. Ontario Dept. of Education, Toronto, 1, 1-40.

Muñoz, C. (2008). Age-related differences in foreign language learning. Revisiting the empirical evidence. *International Review of Applied Linguistics*, 46, 197-220.

Muñoz, C. (2010). On how age affects foreign language learning. *Advances in Research on Language Acquisition and Teaching*, 39-49.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44(01), 1-35.

Muñoz, C. (2011). Input in foreign language learning: more significant than starting age. *International Review of Applied Linguistics (IRAL)*.

Murphy, V. A. (1997). The effect of modality on a grammaticality judgement task. *Second Language Research*, 13(1), 34-65.

O’Shaughnessy, D. (1984). A multispeaker analysis of durations in read French paragraphs. *The Journal of the Acoustical Society of America*, 76(6), 1664-1672.

Oh, S., & Lewis, C. (2008). Korean preschoolers’ advanced inhibitory control and its relation to other executive skills and mental state understanding. *Child Development*, 79(1), 80-99.

Ortega, L. (2013). SLA for the 21st century: Disciplinary progress, transdisciplinary relevance, and the Bi/multilingual turn. *Language Learning*, 63(1), 1-24. doi:10.1111/j.1467-

9922.2012.00735.x

- Oyama, S. (1978). The sensitive period and comprehension of speech. *NABE journal*, 3(1), 25-40.
- Park, S. Y., & Cheah, C. S. (2005). Korean mothers' proactive socialisation beliefs regarding preschoolers' social skills. *International Journal of Behavioral Development*, 29(1), 24-34.
- Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning*, 30, 449-472.
- Penfield, W. & Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton, NJ: Princeton University Press.
- Petrova, K., Wentura, D., & Fu, X. (2013). Cultural influences on oculomotor inhibition of remote distractors: Evidence from saccade trajectories. *Vision research*, 84, 43-49.
- Pimsleur, P., Reed, D. J. and Stansfield, C.W. (2004). *Pimsleur Language Aptitude Battery: Manual 2004 Edition*. Bethesda, MD: Second Language Testing, Inc., 2004.
- Pinker, S. (1994). *The language instinct*. New York: W. Morrow and Co.
- Rabinowitz, J.C. & Craik, F.I.M. & Ackerman, B. P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology*, 36, 325-344.
- Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A. S., McNamara, J. O., & Williams, S. M. (2001). The development of language: A critical period in humans.
- Reder, F., Marec-Breton, N., Gombert, J. E., & Demont, E. (2013). Second-language learners' advantage in metalinguistic awareness: A question of languages' characteristics. *British Journal of Educational Psychology*, 83(4), 686-702.
- Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind a comparison of Chinese and US

- preschoolers. *Psychological science*, 17(1), 74-81.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin*, 110-114.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime: User's Guide, version 1.0, Psychology Software Tools.
- Singleton, D. (1995). "A critical look at the critical period hypothesis in second language acquisition research". In D. Singleton, & Z. Lengyel (Eds.), *The Age Factor in Second Language Acquisition: A critical look at the critical period hypothesis* (pp. 129). Clevedon: Multilingual Matters.
- Singleton, D. (2005). The Critical Period Hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching*, 43(4), 269-285.
- Sparks, R. L., Patton, J., & Ganschow, L. (2012). Profiles of more and less successful L2 learners: A cluster analysis study. *Learning and Individual Differences*, 22(4), 463-472.
- Statistics Canada, (2012a). The Canadian Population in 2011: Population Counts and Growth. Analytical document. Population and dwelling counts, 2011 Census. Statistics Canada Catalogue no. 98-310-X-2011001. Ottawa.
- Statistics Canada. (2012b). Focus on Geography Series, 2011 Census. Statistics Canada Catalogue no. 98-310-XWE2011004. Ottawa, Ontario. Analytical products, 2011 Census.
- Swain, M., & Lapkin, S. (2005). The evolving sociopolitical context of immersion education in Canada: Some implications for program development. *International Journal of Applied Linguistics*, 15(2), 169-186.
- Taylor, S. (1992). Victor: A case study of a Cantonese child in early French Immersion. *The*

*Canadian Modern Language Review*, 48, 736-759.

- Treffers-Daller, J., & Sakel, J. (2012). Why transfer is a key aspect of language use and processing in bilinguals and L2-users. *International Journal of Bilingualism*, 16(1), 3-10.
- Turnbull, M., Lapkin, S., Hart, D., & Swain, M. (1998). Time on task and immersion graduates' French proficiency. *French second language education in Canada: Empirical studies*, 31-55.
- Weber-Fox, C. M., & Neville, H. J. (1999). Functional neural subsystems are differentially affected by delays in second language immersion: ERP and behavioral evidence in bilinguals. *Second language acquisition and the critical period hypothesis*, 2338.
- Wechsler, D. (2007). *Wechsler individual achievement test – (2nd ed.)*. San Antonio, TX: Psychological Corp.
- Weinberger, S. (2015). *Speech Accent Archive*. George Mason University. Retrieved from <http://accent.gmu.edu>
- Welch, B. L. (1947). The generalization of student's' problem when several different population variances are involved. *Biometrika*, 28-35.
- Woodcock, R. W. (1991). *Woodcock Language Proficiency Battery—Revised*. Itasca, IL: Riverside Publishing Company.
- Wightman, M. (1981). The French Listening Comprehension Skills of Grade 6 English Program Students; Second Year of Testing. Research Report 81-10.
- Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism Language and Cognition*, 3(2), 131-149.

## Appendices

### Appendix A French grammaticality judgement task (adapted from Birdsong, 1992)

The following sentences were presented individually on a computer screen using E-Prime<sup>®</sup> Experimental Software (Schneider, Eschman, & Zuccolotto, 2002). Participants judged the sentences as either grammatical or not grammatical by pressing keys on the lap-top keyboard.

*(\*) indicates an ungrammatical sentence.*

Sentence Construction	Grammatical/Ungrammatical
<b>Adjacency</b>	
Albert finira son travail bientôt.	Grammatical
Lucie a donné à Henri des fleurs.	Ungrammatical
Thomas a pris avant la fête une douche.	Ungrammatical
Diane a placé des fleurs dans sa chambre.	Grammatical
Antoinette a traversé rapidement la rue.	Grammatical
Diane a placé dans sa chambre des fleurs.	Ungrammatical
Lucie a donné des fleurs à Henri.	Grammatical
Les garçons regardent la télévision avec intérêt.	Grammatical
Marie a descendu prudemment les marches.	Grammatical
Antoinette a traversé la rue rapidement.	Grammatical
Albert finira bientôt son travail.	Grammatical
Jeanne mange de la crème glacée souvent.	Grammatical
Les garçons regardent avec intérêt la télévision.	Grammatical
Marie a descendu les marches prudemment.	Grammatical
Thomas a pris une douche avant la fête.	Grammatical
<b>En-Avant</b>	
Maurice a vu un certain lit -> Maurice en a vu un certain.	Ungrammatical
Charles a vu un grand lit -> Charles en a vu un grand.	Grammatical
Elle a lu ce livre. Elle en téléphone à l'auteur.	Ungrammatical
La préface de ce livre est flatteuse → La préface en est flatteuse.	Grammatical
Elle a lu ce livre. Elle en aime l'auteur.	Grammatical
Le président de cette compagnie se paie un bon salaire → Le président s'en paie un bon salaire.	Ungrammatical
La préface de ce livre m'agace → La préface m'en agace.	Ungrammatical

<b>Prenominal Past Participle</b>	<b>Grammatical/Ungrammatical</b>
C'était un étonné candidat qui a perdu aux élections.	Grammatical
Le connu romancier vient d'arriver.	Ungrammatical
Le très connu Marcel Proust vient d'arriver.	Grammatical
C'était un étonné Giscard qui a perdu aux élections.	Grammatical
C'était un étonné candidat.	Ungrammatical
Le très connu romancier vient d'arriver.	Grammatical
<b>That-Trace</b>	
Qui crois-tu qui rendra visite à Marc?	Ungrammatical
Qui crois-tu que rendra visite à Marc?	Ungrammatical
Que dis-tu que qui a acheté?	Grammatical
Qui dis-tu qui a acheté quoi?	Ungrammatical
Qui disais-tu qu'a épousé Laure?	Ungrammatical
Qui disais-tu qui a épousé Laure?	Ungrammatical
Que dis-tu qu'a acheté qui?	Ungrammatical
Qui disais-tu a épousé Laure?	Grammatical
Qui crois-tu rendra visite à Marc?	Grammatical
Que dis-tu que Marie a acheté?	Grammatical
<b>De + Modifier</b>	
Elle a oublié deux matinées de libre.	Grammatical
Elle a les lettres d'écrites.	Ungrammatical
Elle a une matinée de libre.	Grammatical
J'ai trouvé un problème de résolu.	Ungrammatical
Elle a les matinées de libre.	Grammatical
Elle a deux lettres d'écrites.	Grammatical
Elle a la matinée de libre.	Grammatical
Deux matinées de libre étaient marquées sur son calendrier.	Grammatical
Elle a oublié une lettre d'écrite.	Ungrammatical
Deux lettres d'écrites étaient retrouvées dans sa chambre.	Ungrammatical
Encore un problème de résolu.	Grammatical
Elle a une lettre d'écrite.	Grammatical
Elle a oublié deux lettres d'écrites.	Ungrammatical
Elle a deux matinées de libre.	Grammatical
Elle a oublié une matinée de libre.	Grammatical
Elle a la lettre d'écrite.	Ungrammatical

<b>Ce/il (elle)</b>	<b>Grammatical/Ungrammatical</b>
Qui est Victor Hugo? C'est un grand écrivain du XIXième siècle.	Grammatical
Bernard Pivot, c'est un intellectuel qui n'a peur de rien.	Grammatical
Marie voulait dire à Jean qu'il était un génie que tout le monde respecte.	Grammatical
Sophie nous a dit qu'elle était malade hier soir.	Grammatical
Marie a dit de Jean que c'est un génie.	Grammatical
Marie voulait dire à Jean que c'était un génie que tout le monde respecte.	Grammatical
Bernard Pivot, il est un intellectuel qui n'a peur de rien.	Ungrammatical
Sophie nous a dit que c'était malade hier soir.	Ungrammatical
Marie a dit de Jean qu'il est un génie.	Grammatical
Qui est Victor Hugo? Il est un grand écrivain du XIXième siècle.	Ungrammatical
<b>Middle-Voice</b>	
Cette maison a été vendue d'elle-même.	Grammatical
Cette pièce se balaie d'elle-même.	Ungrammatical
Cette voiture se lave d'elle-même.	Ungrammatical
Ce tire-bouchon s'utilise de lui-même.	Ungrammatical
Ces formules se mémorisent d'elles-mêmes.	Ungrammatical
Les langues romanes s'acquièrent d'elles-mêmes.	Grammatical
Jeanne mange souvent de la crème glacée.	Grammatical
Cette maison s'est vendue d'elle-même.	Grammatical

## Appendix B English grammaticality judgement task (adapted from McDonald, 2006)

The following sentences were presented individually on a computer screen using E-Prime<sup>®</sup> Experimental Software (Schneider, Eschman, & Zuccolotto, 2002). Participants judged the sentences as either grammatical or not grammatical by pressing keys on the lap-top keyboard.

*(\*) indicates an ungrammatical sentence.*

Sentence Structure	Grammatical/Ungrammatical
<b>Articles</b>	
The lady drove the same car for the past 20 years.	Grammatical
The lady drove same car for the past 20 years.	Ungrammatical
I saw you at the hotel, but you did not see me.	Grammatical
I saw you at hotel, but you did not see me.	Ungrammatical
I jumped in the pool and started swimming very quickly.	Grammatical
I jumped in pool and started swimming very quickly.	Ungrammatical
I have never had a cat, but I have had two dogs.	Grammatical
I have never had cat, but I have had two dogs.	Ungrammatical
<b>Regular Past Tense</b>	
Last night my friend walked home after dark.	Grammatical
Last night my friend walk home after dark.	Ungrammatical
Earlier today, the girl mailed a letter to her uncle.	Grammatical
Earlier today, the girl mail a letter to her uncle.	Ungrammatical
Yesterday morning the man cooked breakfast for his wife.	Grammatical
Yesterday morning the man cook breakfast for his wife.	Ungrammatical
Last semester my roommate bragged about his grades.	Grammatical
Last semester my roommate brag about his grades.	Ungrammatical
Last night the show ended a half hour early.	Grammatical
Last night the show end a half hour early.	Ungrammatical
Last summer my friend rented a beach house.	Grammatical
Last summer my friend rent a beach house.	Ungrammatical
Yesterday afternoon the housekeeper folded laundry for 2 hours.	Grammatical
Yesterday afternoon the housekeeper fold laundry for 2 hours.	Ungrammatical
Last week the child tasted ice cream for the first time.	Grammatical
Last week the child taste ice cream for the first time.	Ungrammatical

<b>Sentence Structure</b>	<b>Grammatical/Ungrammatical</b>
<b>Third person singular</b>	
The boy jumps whenever he is startled. The boy jump whenever he is started.	Grammatical Ungrammatical
The girl likes pancakes for breakfast when she is hungry. The girl like pancakes for breakfast when she is hungry.	Grammatical Ungrammatical
My cat climbs that tree when it is frightened. My cat climb that tree when it is frightened.	Grammatical Ungrammatical
The woman prints her name when she is signing checks. The woman print her name when she is signing checks.	Grammatical Ungrammatical
The mother kisses her daughter good night after she is in bed. The mother kiss her daughter good night after she is in bed.	Grammatical Ungrammatical
The doctor watches television while he is eating. The doctor watch television while he is eating.	Grammatical Ungrammatical
Our friend fixes cars when he is low on money. Our friend fix cars when he is low on money.	Grammatical Ungrammatical
The girl brushes her teeth after she is through with lunch. The girl brush her teeth after she is through with lunch.	Grammatical Ungrammatical
<b>Regular plural</b>	
This meal takes more than four pans to prepare. This meal takes more than four pan to prepare.	Grammatical Ungrammatical
His uncle has over 20 pigs at his ranch. His uncle has over 20 pig at his ranch.	Grammatical Ungrammatical
There are 20 flutes in our marching band. There are 20 flute in our marching band.	Grammatical Ungrammatical
He visited three parks on his summer vacation. He visited three park on his summer vacation.	Grammatical Ungrammatical
I need some matches to light the fire. I need some match to light the fire.	Grammatical Ungrammatical
The realtor sold four houses last month. The realtor sold four house last month.	Grammatical Ungrammatical
The waitress brought four glasses to the table. The waitress brought four glass to the table.	Grammatical Ungrammatical
There are 12 inches in one foot. There are 12 inch in one foot.	Grammatical Ungrammatical

<b>Present Progressive</b>	<b>Grammatical/Ungrammatical</b>
The little girl is playing with her dolls. The little girl is play with her dolls.	Grammatical Ungrammatical
She is leaving to go on her vacation tomorrow. She is leave to go on her vacation tomorrow.	Grammatical Ungrammatical
I am beginning to understand what you are saying. I am begin to understand what you are saying.	Grammatical Ungrammatical
My father is fixing the light that is hard to reach. My father is fix the light that is hard to reach.	Grammatical Ungrammatical
<b>Irregular past tense</b>	
Yesterday morning he took the car for a test drive. Yesterday morning he taked the car for a test drive.	Grammatical Ungrammatical
Last week the pilot flew to Paris. Last week the pilot flyed to Paris.	Grammatical Ungrammatical
Last night the man's car broke down on the interstate. Last night the man's car breaked down on the interstate.	Grammatical Ungrammatical
Last year your team held the record for the most victories. Last year your team holded the record for the most victories.	Grammatical Ungrammatical
<b>Irregular plurals</b>	
After dancing for 2 hours, my feet began to hurt. After dancing for 2 hours, my foots began to hurt.	Grammatical Ungrammatical
Several of the men decided not to go to the football game. Several of the mans decided not to go to the football game.	Grammatical Ungrammatical
<b>Wh-Questions</b>	
What do you think about the new coach? What you think about the new coach?	Grammatical Ungrammatical
When does your father's plane arrive? When your father's plane arrive?	Grammatical Ungrammatical
Why did he come home last night? Why he come home last night?	Grammatical Ungrammatical
Where did you go last night? Where you go last night?	Grammatical Ungrammatical

<b>Yes/No Questions</b>	<b>Grammatical/Ungrammatical</b>
Does the teacher drive a really fancy red car? Does drives the teacher a really fancy red car?	Grammatical Ungrammatical
Does the boy see the new book? Does sees the boy the new book?	Grammatical Ungrammatical
Do you like oranges or apples better? Do like you oranges or apples better?	Grammatical Ungrammatical
Do you go to the store every day? Do go you to the store every day?	Grammatical Ungrammatical
<b>Word Order</b>	
The young child likes carrots. Likes carrots the young child.	Grammatical Ungrammatical
The busy girl finished her homework. Finished her homework the busy girl.	Grammatical Ungrammatical
The cat chased the mouse. The cat the mouse chased.	Grammatical Ungrammatical
The teacher graded the tests. The teacher the tests graded.	Grammatical Ungrammatical

## **Appendix C Material used for speech production tasks**

### **C.1 French paragraph (O'Shaughnessy, 1984)**

"La bise et le soleil se disputaient, chacun assurait qu'il était le plus fort, quand ils virent un voyageur s'avancer enveloppé dans son manteau. Ils tombèrent d'accord que celui qui arriverait le premier à lui faire enlever son manteau serait le plus fort. Alors la bise se mit à souffler de toutes ses forces; mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui, et à la fin, la bise renonça à le lui faire enlever. Alors le soleil se mit à briller et au bout d'un moment le voyageur, réchauffé, enleva son manteau. Ainsi la bise dû reconnaître que le soleil était le plus fort des deux."

*English translation: "The wind and the sun fought, each asserting that he was the strongest, when they saw a traveler advancing, wrapped in his cloak. They were in agreement that the one to first make him take off his coat would be the strongest. Then the wind began blowing with all his might; but the more it blew, the more the traveler clutched his coat around him and in the end, the wind gave up and stopped. Then the sun began to shine and after a while the traveler, warmed up, took off his coat. And the wind had to recognize that the sun was the stronger of the two."*

### **C.2 English paragraph (Speech Accent Archive, Weinberger, 2015)**

"Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station."