EARLY AND EMERGENT BILINGUALS: THE ROLE OF COGNITIVE CONTROL IN THE PROCESSING OF LINGUISTIC CONFLICT

A Dissertation submitted to the Faculty of the Graduate School of Arts and Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Spanish

By

Timothy James McCormick, M.S.

Washington, D.C.
February 14, 2020
ABSTRACT

Language processing requires frequent resolution of conflict (e.g. temporary ambiguities, conflicting parsing principles, see Jegerski, 2012). This conflict triggers cognitive control, which has been shown to be a major player in disambiguation among monolinguals (Botvinick, Braver, Barch, Carter & Cohen, 2001; January, Trueswell & Thompson-Schill, 2009). Cognitive control is also a variable of considerable interest in research on early bilingualism and on emergent (second language) bilingualism, as it is responsible for non-active language suppression (Kroll, Dussias, Bogulski, & Valdes Kroff, 2012). For example, many studies reveal cognitive control advantages for early bilinguals in conflict resolution tasks like the flanker task (see Kroll & Bialystok, 2013). Meanwhile, prefrontal brain regions associated with cognitive control are highly active during non-native, non-highly proficient language processing, but that activity diminishes for proficient or native processing (Abutalebi, 2008). However, further research is needed to understand how cognitive control resources are allocated when both disambiguation and non-active language suppression are necessary.

To that end, this dissertation investigated linguistic conflict resolution by early bilinguals and by emergent bilinguals. The participants were 29 Catalan-Spanish early bilinguals (Experiment 1) and 40 L2 Spanish learners at varying proficiency levels (Experiment 2). Following recent research on monolinguals (Hsu & Novick, 2016), the experimental design leveraged the Gratton effect (sustained cognitive control following
conflict detection) in order to compare processing speeds after conflictive vs. non-conflictive flanker trials (i.e., in a cognitively neutral vs. engaged state). Mixed-effects models in Experiment 1 showed that subject-object ambiguities presented the same difficulties for early bilinguals that Jegerski (2012) observed of Spanish monolinguals. However, the introduction of cognitive control engagement in the design revealed that ambiguity resolution is faster in the engaged state, in line with Hsu and Novick (2016). In Experiment 2, emergent bilinguals revealed a non-native sentence-final processing effect, indicative of either a delayed garden path or shallow second language processing (cf., Clahsen & Felser, 2006). Nonetheless, this effect was also neutralized when the linguistic conflict was preceded by a conflictive flanker. These two experiments buttress Hsu and Novick’s (2016) finding that dynamic cognitive control engagement facilitates recovery from linguistic conflict.
Acknowledgments

“Nanos gigantium humeris insidentes”

This dissertation would not have been possible with the guidance of my committee, Dr. Chandan Vaidya, Dr. Jill Jegerski, and especially Dr. Cristina Sanz, whose guidance throughout the doctoral program has shaped my approach to science and research.

I would also like to thank my family and friends for their endless support. I have been fortunate enough to count on a wide and sturdy safety net throughout my doctoral program:

To all my colleagues at Georgetown, both in and beyond the Department of Spanish and Portuguese, for your help and solidarity.

To my best friends and cousins, Sam and Anna, you gave me perspective on my most difficult days and laughter and companionship on my lost lonely days.

To my brothers and sister, if there are shoulders on which I stand, they are yours. Your passion, knowledge and curiosity have always inspired me.

And of course, those virtues are nurtured by parents who allow their children’s wonder to run free. Good science is not mechanical but depends on an unfettered curiosity. To my mom and dad, thank you for inspiring me to love the questions.

A Daniel y a Matilda, por su apoyo incondicional. Me alivian mi carga mental cada día con su cariño. Gracias por recordarme lo más importante.
# TABLE OF CONTENTS

Chapter 1: Introduction ........................................................................................................... 1

1.1 Statement of the Problem .................................................................................................. 1

1.1.1 Language processing requires reanalysis ................................................................. 1

1.1.2 Cognitive control is involved in reanalysis and conflict resolution........... 3

1.1.3 Bilingualism and cognitive control: Non-linguistic ‘advantages’........... 5

1.1.4 Bilingualism, cognitive control and language processing: Linguistic advantages? ................................................................. 8

1.1.5 Early bilinguals, sentence processing and cognitive control....................... 9

1.1.6 Emergent bilingualism, garden paths and cognitive control.................... 13

1.1.7 Conflict adaptation and garden path resolution.............................................. 17

1.2 Goals and Potential Contributions of the Current Studies.............................. 18

1.2.1 Experiment 1 ............................................................................................................. 20

1.2.2 Experiment 2 ......................................................................................................... 20

1.2.3 Rationale ............................................................................................................... 21

1.3 Definition of Terms ........................................................................................................ 22

Chapter 2: Review of the Literature .................................................................................. 25

2.1 Models of Sentence Processing....................................................................................... 25

2.1.1 Two-phase parsers................................................................................................ 26

2.1.2 Constraint-based models......................................................................................... 27

2.1.3 Sentence processing in Spanish ........................................................................... 31

2.2 Cognitive Control........................................................................................................... 35

2.3 Second Language Learning and Cognitive Control ................................................. 36
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.1 Cognitive capacity in second language acquisition</td>
<td>37</td>
</tr>
<tr>
<td>2.3.2 Working memory, phonological working memory and executive function in second language acquisition</td>
<td>38</td>
</tr>
<tr>
<td>2.3.3 Cognitive control in second language acquisition</td>
<td>45</td>
</tr>
<tr>
<td>2.3.4 Second language learners, ambiguity resolution and cognitive control</td>
<td>47</td>
</tr>
<tr>
<td>2.4 Early Bilingual Sentence Processing and Cognitive Control</td>
<td>48</td>
</tr>
<tr>
<td>2.5 Conflict Detection and the Cognitive-control Engagement Paradigm</td>
<td>53</td>
</tr>
<tr>
<td>2.6 Rationale for the Studies</td>
<td>57</td>
</tr>
<tr>
<td>2.6.1 Rationale: Experiment 1</td>
<td>57</td>
</tr>
<tr>
<td>2.6.2 Rationale: Experiment 2</td>
<td>58</td>
</tr>
<tr>
<td>2.7 Research Questions</td>
<td>59</td>
</tr>
<tr>
<td>2.7.1 Research questions: Experiment 1</td>
<td>59</td>
</tr>
<tr>
<td>2.7.2 Research questions: Experiment 2</td>
<td>60</td>
</tr>
<tr>
<td>Chapter 3: Research Design</td>
<td>61</td>
</tr>
<tr>
<td>3.1 Overview</td>
<td>61</td>
</tr>
<tr>
<td>3.2 Operationalizations</td>
<td>61</td>
</tr>
<tr>
<td>3.3 Selection of Participants</td>
<td>62</td>
</tr>
<tr>
<td>3.4 Target Structures</td>
<td>63</td>
</tr>
<tr>
<td>3.4.1 Non-linguistic stimuli</td>
<td>63</td>
</tr>
<tr>
<td>3.4.2 Linguistic stimuli</td>
<td>64</td>
</tr>
<tr>
<td>3.5 Materials</td>
<td>65</td>
</tr>
</tbody>
</table>
5.6.1 Experiment 1 .................................................. 125
5.6.2 Experiment 2 .................................................. 126
5.6.3 Conclusion ...................................................... 130

Appendix A: Linguistic Stimuli ........................................... 131
Appendix B: Elicited Imitation Task ....................................... 134
Appendix C: Scoring of Elicited Imitation Task ......................... 136
Appendix D: Results for Reading Times at Regions 1 and 2 ............... 138
References ........................................................................ 139
LIST OF FIGURES

Figure 1. Early bilingual reaction time data by condition in Pilot Study 1 using ANOVAs and residual reading time................................................................. 81

Figure 2. Emergent bilingual reaction time data by condition in Pilot Study 2 using ANOVAs and residual reading time ......................................................... 84

Figure 3. Low-proficiency participants’ mean reaction times by condition in Pilot Study 2 using ANOVAs and residual reading time ........................................ 85

Figure 4. High-proficiency participants’ mean reaction times by condition in Pilot Study 2 using ANOVAs and residual reading time ........................................ 85

Figure 5. Reading times for early bilinguals by condition.............................................. 92

Figure 6. Early bilingual log-transformed mean reaction times at Region 3............... 94

Figure 7. Reading times for emergent bilinguals by condition ................................. 101

Figure 8. Predicted logRTs for emergent bilinguals in each condition at three proficiency scores.................................................................................................. 103

Figure 9. Predicted logRTs for emergent bilinguals across conditions.................... 104

Figure 10. Predicted means for emergent bilinguals in congruent and incongruent condition........................................................................................................ 104
LIST OF TABLES

Table 1. Participant background information: Experiment 1…………………………. 89

Table 2. Descriptive statistics of early bilingual raw reaction time means by condition. 91

Table 3. Output for reading times from early bilingual linear mixed-effects models….. 93

Table 4. Descriptive statistics of early bilingual cognitive measures…………………… 95

Table 5. Chi-squared statistics for two exploratory models compared to Model1……… 95

Table 6. Output from two exploratory linear mixed-effects models compared to
Model1………………………………………………………………………………………… 96

Table 7. Participant background information: Experiment 2…………………………. 98

Table 8. Descriptive statistics of emergent bilingual raw reaction time means by
condition………………………………………………………………………………………… 100

Table 9. Output for reading times from emergent bilingual linear mixed-effects
models …………………………………………………………………………………………… 102

Table 10. Comparison of Region 5 output in congruent and incongruent conditions… 106

Table 11. Comparison of Region 5 output in transitive and intransitive conditions….. 107

Table 12. Descriptive statistics of emergent bilingual cognitive measures……………. 108

Table 13. Chi-squared statistics for three exploratory models compared to Model1….. 108

Table 14. Output from three exploratory linear mixed-effects models compared to
Model1………………………………………………………………………………………… 109
CHAPTER 1: INTRODUCTION

1.1 Statement of the Problem

1.1.1 Sentence processing requires reanalysis

Processing language in our surroundings involves constant uptake and interpretation. As the language parser builds a structure for interpretation, it receives more input that it must incorporate into the structure sequentially. However, the hierarchical structure of language and the linear nature of this uptake means the parser will face complex linguistic structures that lead to incorrect interpretations that must be resolved (Frazier & Clifton, 1998; Frazier & Fodor, 1978; MacDonald, Pearlmutter, Seidenburg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995).

Different languages contain different structures of interest for the researchers who investigate this uptake and interpretation, but in English and Spanish, the focus of this dissertation, the structures typically used to investigate revision or recovery from ambiguity are predominantly reduced relative clauses (RRCs), also called subject-object ambiguities, and prepositional phrase (PP) or relative clause (RC) attachment ambiguities. The PP- and RC-attachment ambiguity is of interest because the preference differs between Spanish and English. For example, in example (1), taken from Dussias and Sagarra (2007), both (1a) and (1b) are ambiguous.

(1a) An armed robber shot the sister of the actor who was on the balcony.

(1b) Un ladrón armado le disparó a la hermana del actor que estaba en el balcón.

Although this is a global ambiguity, one that is not resolved within the sentence, native speakers have been shown to have different preferences: the preference in English is that the RC attach to the hierarchically lower noun phrase (NP), “the actor”, but in
Spanish the preference is attachment to the hierarchically higher NP, “la hermana”, such that the actor is interpreted to be on the balcony in English and the sister in Spanish (Cuetos & Mitchell, 1988; Dussias, 2003; Dussias & Sagarra, 2007; Felser, Roberts, Marinis, & Gross, 2003; Fernández, 2003; Papadopoulou & Clashes, 2003)

Contrasting global ambiguities, RRCs are temporary ambiguities, also known as garden-path sentences. Looking at these sentences in their entirety, they are not ambiguous, but as the parser processes them from start to finish, they contain structures or patterns that lead the parser to expect one analysis based on probabilities. However, as the parser continues, that analysis becomes incompatible with new linguistic data, so the parser must reject the preferred or original structure and reassess to incorporate the new information. For example, without the complementizer “that”, the examples in (2), which are taken from Novick, Trueswell and Thompson-Schill (2005), are all technically temporarily ambiguous because “accepted” and “figured” both allow direct object and embedded sentence analyses. However, parsing and lexical preferences result in a garden path in (2a), less so in (2b), and not in (2c-d) (Novick et al., 2005).

(2a) The man accepted the money could not be spent.
(2b) The man accepted the fire could not be extinguished.
(2c) The man figured the money could not be spent.
(2d) The man figured the fire could not be extinguished.

In the first example (2a), the subcategorization frame of accept and the semantic plausibility of money as the direct object of accept leads the parser to interpret the money as the direct object. However, as the parser moves forward, it encounters information that is incompatible with the structure it is currently building (“could not”), so must reanalyze
the money as the subject of the embedded clause and reject the direct object analysis. In example (2b), although the semantic plausibility of “accepted the fire” is weaker than that of “accepted the money”, the parser must maintain the direct object analysis viable until it reaches the disambiguating information (again, “could not”). In the examples (2c-d), the probability of figured accepting an embedded clause rather than direct object is much higher, so the preferred first-pass parse is the correct embedded sentence analysis, meaning no garden path results in these cases.

1.1.2 Cognitive control is involved in reanalysis and conflict resolution

Cognitive control (CC), also referred to as inhibitory control or attentional control, is a cognitive variable involved in the resolution of conflict in our input, responsible for “the performance of specific tasks through appropriate adjustments in perceptual selection, response biases, and the online maintenance of contextual information” (Botvinick, Braver, Barch, Carter & Cohen, 2001, p. 624). Cognitive control tasks include the Stroop, flanker and Simon tasks, in which participants respond to a stimulus that may present more information than necessary, which in some cases will conflict with the information pertinent to the response. For example, in the flanker task, participants must indicate the direction of an arrow at the center of a screen, in isolation (control condition) or flanked by four other arrows pointing in the same (facilitative condition) or opposite direction (conflictive condition). The difference between the reaction times (RTs) in these conditions is interpreted as a reflection of CC.

In addition to these non-linguistic tasks, cognitive control is also involved in resolving the incorrect initial interpretations that occur during language processing mentioned above (see Novick et al., 2005, for a review). The use of online measures like
functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) has allowed researchers to confirm that a shared brain structure is involved in the resolution of both linguistic and nonlinguistic conflict: data from psycholinguistic, neuropsychological and imaging studies converge to show that high-level linguistic processing and non-linguistic cognitive control processes co-localize within the left ventrolateral prefrontal cortex (specifically the left inferior frontal gyrus, LIFG, and anterior cingulate cortex, ACC) (January, Trueswell, Thompson-Schill, 2009; Novick, Kan, Trueswell & Thompson-Schill, 2009; Teubner-Rhodes et al., 2016; Thompson-Schill, Bedny, & Goldberg, 2005; see Fedorenko, 2014, and Luk, Green, Abutalebi & Grady, 2012 for reviews). The relationship between the LIFG, cognitive control and garden path resolution has also been established in behavioral studies.

In healthy adults, longer reading times have been observed with more LIFG activation during the processing of temporary ambiguities, such as dispreferred PP attachment and RRC ambiguities (Mason, Just, Keller & Carpenter, 2003), with later presentation of disambiguating information correlating with greater latency and more LIFG activation (Fiebach, Vos & Friederici, 2004). In other populations, failure to abandon incorrect interpretations (or delayed abandonment) has been associated with cognitive control deficiencies due to underdevelopment of or injury to the prefrontal cortex, that is, children (Trueswell, Sekerina, Hill & Logrip, 1999) and patients (Novick et al., 2009), respectively.

Returning to example (2) from above, Novick et al. (2005) review studies of healthy adults, patients and children to unite evidence to propose that LIFG, and by extension CC, are involved in the reanalysis required in (2a) because of the conflict between the verb’s
subcategorization preference for a direct object, the semantic plausibility of “money” as that direct object, and the syntactic incompatibility of that analysis upon reaching “could”. The example in (2b) also requires LIFG for reanalysis, although “fire” is less semantically plausible as the direct object of the verb “accept” and therefore, abandoning the direct object analysis is not as taxing in their model. In examples (2c-d), the subcategorization frame of “figure” prefers the embedded sentence, so there is no conflict between the parser’s running structure and the final structure, so the LIFG is not recruited.

In summary, the LIFG can be considered a major player in the correct selection of information during goal-oriented tasks, both linguistic and non-linguistic; in the case of sentence processing, it implements the abandonment of an incorrect initial analysis to facilitate correct comprehension (Novick, Trueswell & Thompson-Schill, 2005). However, in addition to the attention given to the LIFG in sentence processing, the structure and its psychological correlate, CC, also receive considerable attention in research on bilingualism (e.g. Hilchey & Klein, 2011; Struys, Mohades, Bosch & van den Noort, 2015; Teubner-Rhodes et al., 2016; Torres & Sanz, 2016).

1.1.3 Bilingualism and cognitive control: Non-linguistic ‘advantages’

Bilinguals do not “turn off” one language when they are using the other. Rather, both languages are active to some extent at any given time (see Kroll, Dussias, Bogulski, & Valdes Kroff, 2012, for a review), and this is true across linguistic modalities (Dijkstra, 2005; Kroll & de Groot, 1997; Marian & Spivey, 2003; Kroll, Bobb & Wodniecka, 2006). Therefore, the bilingual must suppress one language to use the other. Green (1998), in his seminal work proposing the Inhibitory Control (IC) model for bilingual language processing, highlights the similarities between speech and goal-oriented actions, such as
the similarities between selecting a lemma and grasping an object. His IC model proposes that any given language task requires a supervisory attentional system (SAS), which takes into account “perceptual and cognitive cues” to activate specific networks of sequences and control the activation and, importantly, suppression of lemmas, according to the specific goals of the message (p. 68). Thus, although “table”, “chair”, and “mesa” might be within overlapping networks, the SAS only allows the activation of “table” when that is the target lemma, under standard attentional circumstances. Since the publication of this model, the importance of inhibitory control in bilingual language processing has been tested empirically, with studies generally focusing on whether the inhibition in bilingualism results in greater development of domain-general (i.e., non-linguistic) inhibitory control (i.e., cognitive control), as compared to monolingual comparisons.

The empirical research in this field is abundant and beyond the scope of the present dissertation. However, to summarize pertinent findings, bilinguals have been observed to be faster than monolinguals on cognitive control tasks like the flanker task, attentional network task (ANT), Simon task or Stroop task (Bialystok, 2006; Bialystok, Craik, Klein & Viswanathan, 2004; Costa, Hernández, & Sebastián-Gallés, 2008; Costa, Hernández, Costa-Faidella & Sebastián-Gallés, 2009). However, revisiting these studies in a review, Hilchey and Klein (2011) show that across 21 studies in 13 publications, bilinguals have faster response times (RTs) than monolinguals on both congruent and incongruent trials (see Hilchey & Klein, 2011, p. 633, Fig. 3), suggesting that the advantage is one of conflict monitoring, not interference or conflict resolution itself. In the limited studies when global RTs of bilinguals are not faster than those of monolinguals, but only incongruent RTs, effects dissipate across the blocks of the study (e.g. Costa, Hernández & Sebastián-Gallés,
2008). Again, this suggests that bilinguals may be better at adapting to conflict, but not overcoming the conflict itself, per se. As Teubner-Rhodes et al. (2014) put it, after comparing their four groups (bilingual and monolingual groups in high-conflict or no-conflict conditions), “it appears that bilinguals outperform monolinguals on the task that involves conflict, but not selectively on the trials that involve conflict” (p. 220, emphasis in original publication).

Investigations into task-switching have also revealed bilingual advantages, though again, not as clear-cut as one uniform “bilingual advantage”. Task-switching studies use tasks like a card sorting task, in which participants are given different sorting cues throughout the experiment, and they must sort the stimulus card according to the current cue. Researchers compare response latency in trials with the same sorting cues (repetition trials) and “switch” trials to calculate a switch cost, and compare RTs in repetition trials during mixed blocks to latency in single-cue blocks to calculate the mixing cost. Even receiving a sorting cue identical to the previous cue has shown to have a cost: the first trial immediately following a repeated cue is slower than the second trial, called the repeat cost (Hernández, Martín, Barceló & Costa, 2013). Studies have revealed reduced switching costs for bilinguals (Bialystok & Martin, 2004; Garbin et al., 2010; Prior & MacWhinney, 2010; though compare to Paap & Greenberg, 2013), especially true for bilinguals who engage in frequent language switching (Prior & Gollan, 2011); reduced mixing costs for bilinguals, but no reduced switch cost (Soveri, Rodriguez-Fornells, & Laine, 2011); and a reduced restart cost and faster response latencies, but no reduced switch cost in any of three studies (Hernández et al., 2013). Therefore, while bilinguals are hypothesized to have
higher capacities interacting with conflictive stimuli and tasks, how those advantages are empirically observed has varied.

Most research focuses on non-linguistic cognitive advantages as a result of early bilingualism through orthogonally designed studies comparing monolinguals and early bilinguals. However, another question emerges about how cognitive abilities are applied during language processing. As Teubner-Rhodes et al. (2016) report, “Few studies have tested whether the bilingual advantage cascades into language processing. Provided that the source of bilinguals’ cognitive advantage is the systematic control of two languages, these benefits should be observed in the linguistic domain – however, much of the work in this area focuses on the effects of bilingualism in non-linguistic contexts” (p. 213). Alternatively, however, we may not see any benefit in the linguistic context, precisely because this advantage is nullified by the nature of bilingualism. In other words, if we suppose an advantage for bilinguals emerges from balancing (to a greater or lesser degree) two languages throughout their lives, it may be the case that that advantage is neutralized when the bilingual has to suppress one of the two languages, while a monolingual does not have that additional conflict of an active second language (L2) for cognitive control to resolve. This theoretical question thus requires us to explore to what extent our understanding of monolingual language processing is also true of bilingual processing.

1.1.4 Bilingualism, cognitive control and language processing: linguistic advantages?

Several studies have documented that cognitive control plays a role in bilingual lexical processing, evidenced by delays in the processing of cross-linguistic homographs (Martín, Macizo & Bajo, 2010) and sustained inhibition of the first language (L1) following production of the L2 (Misra, Guo, Bobb & Kroll, 2012), but shorter cognitive control
recovery periods following lexical processing compared to monolinguals (Blumenfeld & Marian, 2011). These results reflect a switch cost associated with language switching, but suggest that it may also train cognitive control to recover from difficult lexical processing.

Of course, language extends beyond words, and bilingualism beyond knowledge of two lexicons. Bilinguals also maintain two morphosyntactic systems. As Kroll & Bialystok (2013) highlight, questions concerning bilingual sentence processing have generally focused on the interaction between or coexistence of two grammars in a bilingual system or on L2 learners’ eventual attainment of a “native-like” L2 grammar (p. 509). The nature of these research foci has typically led researchers to look to comparisons between early bilinguals and monolinguals, or between L2 near-natives and monolingual natives. However, it is important to differentiate these different populations of bilinguals, given the differences between the profiles that have been documented (see Kroll & Bialystok, 2013, and Montrul, 2005, for example).

1.1.5 Early bilinguals, sentence processing and cognitive control

While the research referenced above considers whether a cognitive control advantage emerges in early bilinguals due to constantly maintaining two languages, it does not consider how this advantage plays out in early bilingual sentence processing. As addressed above, prior research shows that cognitive control has at least two roles in language processing: first, cognitive control works to suppress one language during the processing of the other (Kroll, Dussias, Bogulski, Valdes Kroff, 2012), and second, cognitive control intercedes to resolve conflict when the parser must incorporate incompatible linguistic material into a running structure (Novick et al., 2005). Limited
studies have considered how these two roles of cognitive control interact during sentence processing by early bilinguals.

In their 2010 publication, Moreno, Bialystok, Wodniecka and Alain compare sentence processing in bilinguals and in monolinguals. Participants in this study were tasked with an acceptability judgement task and grammaticality judgment task. In the former, participants reported whether sentences were grammatically or semantically correct. In the latter, participants only reported whether a sentence was grammatically correct, even if the sentence was semantically anomalous. Bilinguals were less accurate on the acceptability task than their monolingual peers; however, on the grammaticality task, the more demanding task according to these researchers and others (e.g. with bilingual children, Bialystok 1986, 1988; Bialystok and Majumder 1998; Cromdal 1999), bilinguals behaved similarly to monolinguals. However, this performance is likely at ceiling (93-96%, depending on the condition). Despite this, the event-related potentials (ERPs) measured through electroencephalography (EEG) revealed smaller P600 amplitudes to syntactic violations, and the monolinguals’ mean N400 amplitudes were the same for semantically valid and anomalous sentences. These findings suggest that monolinguals did not process the semantics of the sentences when looking for syntactic violations, while bilinguals did process the anomaly and still obtained the same accuracy. The authors interpret this as bilinguals processing the syntactic anomalies more efficiently and the semantics even when not necessary, all while maintaining the same accuracy as monolinguals (Moreno et al., 2010, p. 566).

Observing the conflict between at-ceiling performance on the grammaticality task in Moreno et al.’s (2010) study and the assumption that it is the more cognitively
demanding task, Paap and Liu (2014) conducted a replication of this study. The authors manipulated the design in select ways, such as using self-paced reading (SPR) instead of rapid serial visual presentation (RSVP) or EEG. The authors find a pattern similar to Moreno et al. (2010), with monolingual advantages in both tasks, but they did not find that bilinguals “close the gap” on the grammaticality task (Paap & Liu, 2014, p. 70). That is, they did not find the same at-ceiling performance. They also offer the slower RTs and lower accuracy on the acceptability task as an indication that this is in fact the more cognitively demanding task.

One important caveat of the Moreno et al. (2010) study and its replication in Paap and Liu (2014) is the extent that participants were attending to the meaning of the sentence throughout the task. As highlighted by the ERP data from Moreno et al. (2010), the monolinguals did not reveal N400s that you would expect with semantically anomalous sentences. Since the task following the sentences was simply an assessment of grammaticality or acceptability, participants could have limited their attention to the semantics in at least one of the two tasks, and the critical question that this dissertation poses is to what extent can bilinguals fulfill the demands of two tasks that require cognitive control simultaneously.

In the same publication, Paap and Liu (2014) also replicated the homograph-interference task originally conducted by Gernsbacher, Varner and Faust (1990), in which participants read sentences that ended in a homograph. After the homograph, they saw an isolated word that related to one of the two main meanings of that homograph and report whether this word related to the meaning of the homograph as used in the sentence. The findings of this study revealed that bilinguals suffer more homograph interference than
monolinguals, and this interference lessens for both groups in trials with an increased interstimulus interval (ISI). Thus, this replication also does not reveal a bilingual advantage, although given the design of the study, the conflict mediated by the participants is more lexical than syntactic. Another caveat of this study and the study it replicates relates to the profile of bilinguals included: the bilingual groups in Moreno et al. (2010) and in both Experiment 1 and 2 of Paap & Liu (2014) were heterogeneous in their profiles in terms of the languages they spoke, age of acquisition (AoA), and percentage of daily English use.

Teubner-Rhodes et al. (2014) conducted the first study to my knowledge that directly studied early bilingual cognitive advantages and garden-path recovery. In addition to a self-paced reading task in which participants read temporarily ambiguous subject-first or object-first cleft sentences, they also conducted high-conflict and no-conflict N-back studies. Their results align with a conflict monitoring perspective: bilinguals revealed more accurate sentence comprehension, but the higher accuracy was true of both the ambiguous and unambiguous trials (and congruent and incongruent non-linguistic N-backs), suggesting again that early bilinguals are better at monitoring for conflict and that their advantage “may only impact late-stage revision processes” (p. 223). Unlike the homograph interference task in Paap and Liu (2014), this study targets syntactic processing. In addition, this study contrasts previous studies in that the stimuli are temporarily ambiguous and do require reanalysis. While Moreno et al. (2010) argue that the grammaticality judgment task is cognitively demanding, there is no syntactic revision involved in the processing. In other words, the stimuli in this paradigm are either correct or anomalous, but never temporarily so.
In summary, in complex language processing, monolinguals may have a behavioral advantage identifying anomalous or incorrect sentences, but at the physiological level, early bilinguals process more efficiently. In addition, early bilinguals reveal an advantage in their ability to revise garden-path sentences during real-time processing. However, the studies conducted on cognitive control and language processing maintain the orthogonal design common in the field that compares bilinguals and monolinguals, two groups with distinct language profiles and experiences, but recent work by Hsu and Novick (2016), addressed below, has presented researchers with an alternative that allows within-subject comparisons of cognitive control in the more or less engaged conditions. First, I will address emergent bilingualism, that is, L2 acquisition, and how cognitive control and syntactic reanalysis play a role in this process.

1.1.6 Emergent bilingualism, garden paths and cognitive control

Among the emergent bilingual/adult language learner literature, parsing preferences of globally ambiguous sentences receives considerable attention (e.g. Dussias, 2003; Dussias & Sagarra, 2007; Felser et al., 2003; Fernández, 2003; Papadopoulou & Clashen, 2003). These studies have an orientation towards the product of L2 acquisition: do non-natives interpret complete sentences like natives. However, garden paths and reanalysis of temporary ambiguities, which revolve around questions of the processes of L2 comprehension, have also received some attention (Dussias & Piñar, 2010; Frenck-Menstre & Pynte, 1997; Hopp, 2006; Jackson & van Hell, 2010; Jegerski, 2012; Juffs, 1998, 2005; Juffs & Harrington, 1995; Pozzan & Trueswell, 2016; Rah & Adone, 2010). Given this dissertation’s focus on CC’s role in language use and reanalysis, this review will be limited to the latter set of studies, where some form of reanalysis is necessary.
Motivated by differences in adult and child language processing, Pozzan and Trueswell (2016) conducted a study on intermediate L2 learners of English. Children have been shown to have proportionally greater difficulty recovering from garden paths and using referential context to facilitate reanalysis (e.g. Trueswell et al., 1999), so Pozzan and Trueswell (2016) investigated whether these differences are learner phenomena or should be attributed to the underdeveloped cognitive control of children. The authors compared two groups of adult English speakers (L2 learners and native speakers) to determine what aspects found in child processing are shared by adult language learners, and what aspects are not replicated in adults and therefore should be attributed to immature CC. While both groups of adults used referential context, the learner group behavior followed patterns from previous studies of children when recovering from syntactic misanalysis. The authors suggest this finding indicates that adult L2 learners’ emerging language skills account for the difficulties they face in garden path recovery.

Importantly, however, Pozzan and Trueswell (2016) do not account for a second possible explanation for their “learner phenomena”: that L2 processing impinges on cognitive resources which makes learners perform like children during reanalysis. That is, similar results may be obtained, not because the linguistic structure is equally unfamiliar for both sets of learners but because low-proficiency L2 processing occupies cognitive control abilities in adults (Abutalebi, 2008), making comprehension run aground in cases when revision is needed, as is also the case in children for whom these same abilities have not yet fully developed. Given that the authors only include an intermediate learner group, they could not test how proficiency played a role in their results and call for future research to include multiple proficiency levels to clarify these results.
A related hypothesis was proposed by Clahsen and Felser (2006): that non-native (L2) processing is inherently different from native processing. Clahsen and Felser proposed that both children and adults process their L1 online and continuously, but adult L2 speakers do not take advantage of all of the tools at their disposal, relying on semantic and lexical cues, while underusing syntactic ones. They called their hypothesis the shallow structure hypothesis (SSH). Indefrey (2006) questions whether this is more reflective of a high- vs. low-working memory capacity dichotomy in both L1 and L2 processing, which will be addressed in more detail below in Section 2.3. Other researchers questioned the initial hypothesis, and new research in human language processing has been published in the interim, so Clahsen and Felser (2018) made minor updates and clarifications to their hypothesis, including incorporating an account of new findings on differences in the processing time-course. These differences include delayed effects of syntactic violations in L2 processing compared to native processing (e.g., Boxell & Felser, 2017; Felser & Cunnings, 2012; Morgan-Short, 2014; c.f., Clahsen & Felser, 2018). These delayed processing effects may help explain the learner phenomena demonstrated in Pozzan and Trueswell (2012).

Several researchers have included proficiency as a variable in their studies on L2 garden path resolution. Jackson and van Hell (2011), Rah and Adone (2010), and Hopp (2006) all found disambiguation ability to be a factor of proficiency. In the first two studies, native speakers and advanced L2 learners did not reveal processing asymmetries during the processing of *wh*-ambiguities or reduced relative clauses, respectively. However, behavioral differences between the ambiguous and unambiguous condition were obtained in the intermediate learner group in both studies. Rah and Adone (2010) conducted offline
comprehension measures as well and found that both the intermediate and advanced group comprehended the structure offline, despite the processing difficulty. In Hopp (2006), native, near-native L2, and advanced L2 groups revealed processing difficulties at the disambiguating region, but only the advanced group’s slow-down lasted to the end of the sentence, while the other two groups were able to incorporate the disambiguating information and return to their normal pace. The findings of these studies are not entirely surprising: with higher proficiency level comes more experience and input in the L2. However, the results do reveal that recovery from ambiguity is a function of linguistic competency or proficiency, even when the ambiguity exists in the L1, and it is not reflective of inherent constraints of adult language learning.

Given the relationship between proficiency and cognitive capacity, such as the increased recruitment of the LIFG in neuroimaging studies (Abutalebi, 2008), or the correlation between L2 morphosyntactic development and cognitive capacity at lower levels but not at higher proficiency (Serafini & Sanz, 2016), we have reason to ask how L2 learners’ recovery from garden paths at different language proficiency levels relates to CC.

While this research agenda has been started, there remain several gaps in the literature. Juffs (2005), for example, studied whether recovery from *wh*-extraction garden paths by L2 learners of English would reflect working memory capacity (WMC), L1 typology, or both. He did not find a correlation between reading span task or sentence span, but did find that participants recovered faster if their L1 allowed *wh*-extraction. Dussias and Piñar (2010) also studied *wh*-extraction in L2 English and WMC, as well as the role of semantic plausibility. They found that only the natives and the higher-WMC group exploit the plausibility of the sentence to facilitate reanalysis. However, WMC is a very broad
measure, while cognitive control has been directly associated with reanalysis during syntactic processing. In addition, neither of these studies includes proficiency as a variable of interest. To date, no study to my knowledge has considered the how both proficiency and cognitive control (or any measure of cognitive capacity) affect L2 garden-path recovery within the same study.

**1.1.7 Conflict adaptation and garden path resolution**

In these situations where cognitive control is recruited for two language-specific purposes, one broad role (non-active language suppression) and one focused role (garden path reanalysis), language processing may be affected differently from the observations made in psycholinguistic studies on monolingual processing. However, given the polemic of the bilingual advantage in CC, the researcher faces an issue of validity upon comparing bilinguals to monolinguals using an orthogonal between-groups design in this area of research. Likewise, given the individual variation of cognitive control across the population, studying the role of cognitive control at different levels of proficiency in adult L2 learners may cause validity issues. To counteract these issues, a within-subject design that investigates language processing internal to each participant, in relatively more and less engaged cognitive control conditions, could reveal how CC’s role varies across these groups during language processing and reanalysis. One possible solution has come out of psycholinguistic research recently, using conflict adaptation to achieve these more and less engaged cognitive control conditions.

The detection of conflict, linguistic or otherwise, has been shown to trigger sustained CC; for example, the Stroop effect can be lessened for an incongruent trial if it is immediately preceded by another incongruent trial, known as conflict adaptation
This pattern has also been observed in adults’ performance in cross-task paradigms, where linguistic conflict facilitates subsequent non-linguistic conflict resolution (Kan et al., 2013), or vice versa (Hsu & Novick, 2016). In this latter study, syntactic reanalysis was facilitated by conflict adaptation, when the linguistic trial was immediately preceded by a conflictive Stroop trial (Hsu & Novick, 2016). By controlling the interleaving of the two types of stimuli, the researchers were able to draw conclusions about reanalysis in the more engaged condition, relative to reanalysis in the less engaged condition (in other words, within-subject comparisons), rather than make comparisons or correlations according to individual differences in CC. This paradigm, then, can be extended beyond monolingual psycholinguistic research to answer the questions remaining on the role of cognitive control on garden path resolution in different profiles of bilinguals.

1.2 Goals and Potential Contributions of the Current Studies

Provided the decreased recruitment of CC-related neural structures (e.g. Abutalebi, 2008) and the decreased role of executive functions associated with language development at higher proficiencies (e.g. Serafini & Sanz, 2016), conflict adaptation can serve to elicit the more and less engaged conditions of cognitive control within L2 learners at different proficiency levels, canceling to some extent the effects of individual differences between participants that do not relate to proficiency. If cognitive control plays a differential role in online processing at different proficiencies, this cognitive-control engagement paradigm will obtain different results from different groups. If cognitive control is recruited less for suppression of the non-active language at higher proficiencies, a garden path preceded by an incongruent trial will likely reveal faster disambiguation, as there will be more resources
readily available for reanalysis. However, if cognitive control is highly engaged due to low L2 proficiency, a ceiling effect may be obtained wherein this activation of cognitive control would not facilitate disambiguation, resulting in similar reanalysis abilities in both preceding-trial conditions. An alternative possibility may involve cognitive fatigue, with L2 learners whose cognitive control resources are fully engaged obtaining slower reanalysis following incongruencies because the sustained engagement of cognitive control will have depleted their cognitive control resources. This would coincide with findings from children, a group with underdeveloped CC, and reanalysis of the “kindergarten path” (Huang, Gerard, Hsu, Kowalski & Novick, 2016; Huang, Hsu, Leonard, Kowalski, & Novick, in revision). In short, for adult L2 learners, if increased proficiency aligns with decreased dependence on CC, this availability of cognitive control resources should be observed as a greater RT benefit from cognitive control engagement, and this relative increase in RT would likely not be seen at lower proficiencies.

Meanwhile, early bilinguals, with life-long practice employing cognitive control to suppress the non-active language, have been shown to have a benefit in cognitive control and conflict monitoring. However, as Kroll and Bialystok (2013) state, “Individuals are not bilingual or not, and tasks are not measures of inhibition or not: These are all continua in which experience expresses itself through multiple facets” (p. 502). Researchers have historically conflated different profiles of bilingualism into one, conflict can require more or less effort to resolve, and the bell curves of human nature will result in plenty of monolinguals who have more cognitive control than many bilinguals. For this reason, it is useful to work with cognitive control in relative terms (that is, in terms of more- and less-engaged). In addition, it is unclear how the cognitive control benefit found for the bilingual
population affects language processing, as it may be occupied by the suppression of the non-active language that is theorized to engender the benefit to begin with. How reanalysis proceeds under the more engaged and less engaged cognitive control conditions will contribute to our understanding of how bilinguals recruit this cognitive control benefit during language processing.

To that end, the current dissertation includes three studies. These studies aim to investigate the role of cognitive control in syntactic disambiguation and reanalysis by emergent L2 bilinguals and early bilinguals, given the dual role of cognitive control in these populations (non-active language suppression and syntactic revision).

1.2.1 Experiment 1

The first experiment of this dissertation aims to contribute to the research on early bilinguals by investigating how balanced bilinguals respond to linguistic conflict. Participants complete a self-paced reading task with subject-object ambiguities in fronted adverbial clauses in a cross-task cognitive-control engagement design, with flanker stimuli triggering conflict adaptation or not, following Hsu & Novick (2016) and Huang et al., (2016). This design aims to investigate how conflict resolution differs under the two conditions. In practice, this can inform us more about how early bilinguals employ cognitive control during language processing.

1.2.2 Experiment 2

The participants of this experiment are L2 learners at increasing proficiency levels, having completed at least the equivalent of six college-level courses. They engage with the same linguistic stimuli as participants in Experiment 1 during a self-paced reading task with a cross-task cognitive-control engagement design. This leverages the prolonged cognitive
engagement following incongruencies to make within-subject comparisons of the resolution of linguistic conflict at more engaged vs. less engaged cognitive control conditions at each proficiency.

1.2.3 Rationale

This project has outlined several gaps in the L2 and early bilingual psycholinguistic research. First, although there is an abundance of research connecting garden path resolution and cognitive control in monolinguals (Mason et al., 2003; Novick et al., 2005; Novick et al., 2009; Trueswell et al., 1999), there is a dearth of studies that consider the role of cognitive control in sentence processing in different profiles of monolinguals. Given the role of cognitive control in the suppression of the non-active language in bilinguals, our understanding of sentence processing and garden path resolution is limited without explicitly considering whether bilinguals reveal the same patterns as monolinguals, who do not employ cognitive control to suppress an alternate language. The first experiment in this dissertation aims to contribute to our understanding of balanced/early bilingual sentence processing and how early bilinguals employ cognitive control to resolve linguistic conflict when it is already otherwise employed, suppressing the non-active language.

We also know that cognitive control has a differential role across proficiency levels, from neuroimaging (e.g., Abutalebi, 2008) and behavioral studies (e.g., Serafini & Sanz), likely related to the suppression of the non-active language. By conducting a within-subjects design using the cognitive-control engagement paradigm, we may also contribute to the understanding of conflict adaptation and this new paradigm, which may provide a new tool for L2 psycholinguistic research, while contributing to our understanding of how cognitive control is employed in the L2.
1.3 Definition of Terms

Boggle: A processing interruption or recalibration, which in sentence processing leads to reanalysis.

Cognitive control: A cognitive ability involved in the resolution of information conflict, associated with structures within the left ventrolateral prefrontal cortex (specifically the left inferior frontal gyrus and anterior cingulate cortex). Elsewhere in the literature, this variable may be referred to as ‘inhibitory control’ or ‘attentional control’.

Conflict adaptation (also: Gratton effect): A phenomenon in which the detection of conflict in the input leads to temporarily sustained cognitive control procedures which leads to a reduction in processing costs of subsequent conflict.

Early bilinguals: Operationalized in this study as a profile of bilinguals who grow up speaking two languages from before age 5, who report using both native languages regularly and continuously, learning both languages simultaneously. In this dissertation, the early bilinguals selected for participation live in a society that promotes the regular use of both languages and they were educated in both languages from an early age (i.e., the early bilinguals in this study are not heritage language speakers).

Emergent bilinguals: Second language learners who begin the acquisition of their second language after age 5 and achieve fluency after puberty.
Gratton effect: See “conflict adaptation” above.

Garden path sentence (also: temporary ambiguity): A type of sentence that is unambiguous when considered offline, but during online processing, these sentences contain a point of bifurcation at which the parser’s running structure cannot account for incoming linguistic input, and therefore forces a reanalysis. The parser creates this incorrect initial analysis based on frequency of linguistic structures, semantics, or other probabilities.

Linguistic conflict: This term is used to discuss competing impulses triggered by the linguistic input, such as a mismatch between the principle of Late Closure and a verb’s intransitive subcategorization matrix, a linguistic conflict presented in the dissertation’s stimuli. Garden-path sentences also present linguistic conflict.

Non-active language: The context-inappropriate language, used to avoid the term “inactive”, given that both languages in a bilingual are active to greater or lesser extent.

Parser: A component of human language that processes online linguistic input.

Proficiency: The level of global ability in a given language, operationalized in this dissertation as the score on the elicited imitation task (Ortega, 2000), a validated measure of global Spanish proficiency (Bowden, 2007).
Reanalysis: The process of abandoning an initial parse upon receiving information indicating that this first syntactic structure being built is incorrect and cannot account for the linguistic data.

Stroop effect: Processing cost associated with conflict in the input

Temporary ambiguity: See “garden path sentence” above.
CHAPTER 2: REVIEW OF THE LITERATURE

2.1 Models of Sentence Processing

When participating in a conversation, it’s commonly assumed that it is the content of the conversation that creates difficulty for an interlocutor: for example, the more unfamiliar the content, the more difficult. In other cases, we hear reference to ‘ten-dollar words’, suggesting that a participant in a conversation may use unfamiliar lexicon. The grammatical or syntactic structure of the speech often seems entirely unremarkable. However, sentence processing is a highly complex system, incorporating various levels of coordination. Although this coordination is very rapid and generally very smooth, there are cases in which a syntactic peculiarity is encountered that causes a processing interruption, or ‘boggle’. One particularly frequent example is the famous garden path sentence below:

(1) The horse raced past the barn fell.

This sentence is used by psycholinguists to present an example of when reanalysis is necessary in sentence processing. Given the linear nature of language, the parser must create the syntactic structure online as they are encountered, processing new items and incorporating them into the existing structure according to syntactic rules and restrictions. In (1), the parser would first interpret \textit{raced} as the main verb of the sentence. However, when the listener encounters \textit{fell}, she must abandon her incorrect analysis and re-interpret the sentence in order to acquire the correct interpretation within a grammatically licit structure, such that \textit{raced past the barn} is a reduced relative clause (RRC) (i.e., \textit{that was raced past the barn}) while \textit{fell} fills the role of main verb.

The example in (1) presents a case of reanalysis, a key component of research on sentence processing that has been used to build several models of parsing, generally
divided into universal (two-phase) parsers and Constraint-based (one-phase) parsers. Frazier and colleagues’ Garden Path model (and its later update, Construal Theory) is the predominant two-phase model (Frazier & Fodor, 1978; Frazier & Clifton, 1998). Constraint-based or experience-based models have received more recent focus (e.g. Cuetos & Mitchell, 1988; Novick, Thompson-Schill & Trueswell, 2008; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995; MacDonald et al., 1994).

2.1.1 Two-phase parsers

In the two-phase models put forward by Frazier, Fodor and colleagues, the parser’s first pass incorporates new material into the active syntactic structure using a highly restricted system of incorporation limited to the syntax, and importantly only building one structure, until an alternative structure becomes necessary, at which point reanalysis becomes necessary. This is driven by syntactic economy, specifically Frazier’s Minimal Attachment principle, which requires the first pass to create the structure that requires the fewest syntactic nodes while remaining grammatically-permissible, as well as Frazier’s (1978) Late Closure principle, which requires new material be incorporated into the current constituent whenever grammatically possible. Thus, only the lexical items’ subcategorization information (e.g., what types of complements does the item accept, what is the part of speech) is consulted when constructing the syntactic structure. The parser’s second-phase considers semantic and pragmatic information beyond the syntax.

Following evidence that identical syntactic structures with differing lexical items may cause differences in ease of recovery or reanalysis based on biases of the lexical items, these researchers updated their two-phase model with the Construal Theory (Frazier & Clifton, 1998). Construal Theory allows additional lexical information, such as bias or
frequency data, to be consulted during the second phase. Compare the following structures, each with temporarily ambiguous structures that require reanalysis (Ferreira & Henderson, 1998, their examples 6a-b):

(2a) The woman knew the nervous man would leave.

(2b) The woman saw the nervous man would leave.

Both main verbs’ subcategorization frames allow a noun phrase (NP) or a propositional complement (CP). According to Construal Theory, then, the nervous man would be incorporated into the structure as a NP in order to minimize the number of nodes used during the first phase. However, upon encountering would leave, the parser must reanalyze the nervous man as the subject of a CP and reconstruct the structure accordingly. Here, the CP-bias of know makes reanalysis simpler than the reanalysis of (2b), where saw has a NP-bias (Ferreira & Henderson, 1998).

2.1.2 Constraint-based models

Despite the update to incorporate lexical biases, the Garden Path/Construal model cannot account for preferences that go beyond lexical biases to cross-linguistic phenomena. Differences in Spanish- and English-speaker preferences related to RC attachment were among the first of these cross-linguistic preferences observed. Specifically, RC attachment preferences in multiple genitive constructions that do not inherently cause boggles (i.e., they are not ungrammatical but rather globally ambiguous without the appropriate context) cast the two-phase, syntactically-restricted parser into doubt. Consider the following examples from Cuetos and Mitchell (1998, their examples 1a-b):

(3a) El periodista entrevistó a la hija del coronel que tuvo el accidente.
The journalist interviewed the daughter of the colonel who had had the accident.

The preference in Spanish is for the high-attachment reading that the daughter had the accident, which contrasts with the English preference that the colonel had the accident (Cuetos & Mitchell, 1998; Dussias & Sagarra, 2007). More importantly for theoretical purposes, it also contrasts with the universal preference that Frazier’s Late Closure principle would force. In addition, a corpus study showed that exposure to the high-attachment is, in fact, much more frequent in Spanish (Cuetos, Mitchell & Corley, 1996), which suggests that past exposure and frequency account for cross-linguistic preferences. Most current research in sentence processing recognizes this need for the parser to consider several sources of information simultaneously: frequency, syntax, semantics, and discourse, for example, may all contribute to the resolution of ambiguity during one-pass parallel processing. As the parser processes in these models, feasible structures are ranked and re-ranked according to probability of being the correct structure (MacDonald et al., 1994; Novick et al., 2005; Novick et al., 2009; Tanenhaus et al., 1995; Trueswell, Tanenhaus, & Garnsey, 1994).

For example, while in (2a-b) above, the frequency of the CP after each verb differs and this information is theoretically used to facilitate disambiguation in the second-phase of processing, the semantics of the NP that immediately follows the verb can also help disambiguate such a sentence. By not incorporating this information during a first pass, the Minimal Attachment principle may lead to syntactically uneconomical parsing. Take the examples in (4a-d), from Novick et al. (2005).

(4a) The man accepted the money could not be spent.
(4b) The man figured the money could not be spent.

(4c) The man accepted the fire could not be extinguished.

(4d) The man figured the fire could not be extinguished.

In the first example (4a), the subcategorization frame of *accept* and the semantic plausibility of *money* as the direct object of *accept* leads the parser to interpret *the money* as the direct object. However, as the parser encounters information that is incompatible with the structure it is currently building, it must reanalyze *the money* as the subject of the embedded clause. In the example (4c), although the semantic plausibility of “*the fire*” as the complement of “*accept*” is lower than that of “*the money*” in (4a), frequency data associated with the subcategorization frame of *accept* would still lead the parser to interpret the NP that follows as a direct object; however, the recovery from the ambiguity in (4c) is less burdensome than that of (4a). In the examples (4b) and (4d) on the other hand, the probability of *figured* accepting an embedded clause rather than direct object is much higher, so frequency data trumps concerns of the number of syntactic nodes. In these examples, then, we can see how multiple domains can interact to lead to levels of garden paths and reanalysis difficulty.

To understand empirically how these different domains interact during parsing, researchers exploit online data collection methods like eye-tracking (e.g. Betancort, Carreiras, & Sturt, 2009; Carreiras & Clifton, 1999), self-paced reading (e.g. Carreiras, Salillas, & Barber, 2004; Teubner-Rhodes et al., 2014), ERPs (e.g. Tokowicz & MacWhinney, 2005) and fMRI scans (e.g. Mason, Just, Keller, & Carpenter, 2003), generally using temporarily ambiguous constructions like (2) and (4) instead of globally
ambiguous structures like those in (3). With these methodologies, researchers can collect data that reveals the resolution of the ambiguity in real time.

Evidence shows that the parser does use information beyond the syntax to process linguistic information correctly, providing evidence for constraint-based models. For example, Trueswell and colleagues (e.g. Pozzan & Trueswell, 2016; Trueswell et al., 1999), using the visual world paradigm, have consistently found that the parser uses referential context (information available outside of the linguistic system, in this case in the visual domain) by controlling the use of definite articles in the stimuli. When there are two competitors in the visual world, (e.g. two frogs presented on the screen as candidates for selection), the use of a definite article “the” triggers participants to expect the information that follows the definite article to specify which of the two competitors is being referred to. Otherwise, the use of a definite article is anomalous, so listeners use that as a clue to distinguish which frog is the referent, showing that parsers are using more than syntactic and lexical input to disambiguate in real-time processing.

While the specific constraint-based models are still underspecified, certain assumptions are characteristic of them all: an ambiguous strain results in multiple alternative structures through bottom-up processing, and constraining evidence from several domains is integrated on the first (and only) pass to resolve the ambiguity, or rather, to select the most felicitous option. However, the extent to which each domain plays a role varies from model to model (for example, Tanenhaus et al., 1995, vs. MacDonald et al., 1994). These models have been shown to account for phenomena associated with bilingual sentence processing, as well, an area of psycholinguistic research that remains highly understudied and will be addressed below.
2.1.3 Sentence processing in Spanish

Research on sentence processing, reanalysis and garden paths in Spanish has contributed greatly to our understanding of these phenomena more broadly. For example, RC-attachment ambiguities such as that presented in (3) can be globally ambiguous in both English and Spanish (we don’t know whether the colonel or his daughter had the accident). However, it’s been established that uninfluenced Spanish parsing prefers high-attachment (e.g. Carreiras & Clifton, 1999; Carreiras, Salillas & Barber, 2004; Cuetos & Mitchell, 1988; Cuetos et al., 1996; Dussias, 2003; Dussias & Sagarrá, 2007; Fernández, 1995, 1999; Mitchell & Cuetos, 1991). Despite this preference, we can construct stimuli that take advantage of grammatical gender in Spanish to force the attachment of the RC to one of the two NP candidates, and even vary the semantic plausibility or other controllable elements of the sentence to test how these different factors interact, such as the stimuli in (5), (5a-b come from Dussias and Sagarrá, 2007, originally from Carreiras and Clifton, 1999):

(5a) El policía arrestó a la hermana del criado que estaba enferma desde hacía tiempo.

(5b) El policía arrestó al hermano de la niñera que estaba enferma desde hacía tiempo.

(5c) El policía arrestó al hermano de la monja que estaba involucrado en el atraco bancario.

(5d) El policía arrestó al hermano de la monja que estaba involucrada en el atraco bancario.
In (5a) and (5b), the long-distance gender concordance designated by the -a feminine marker in *enferma* disambiguates the ambiguity, indicating high-attachment in (5a) and low-attachment in (5b). This ambiguity in genderless English would not be disambiguated by “sick”, on the other hand. Meanwhile, in (5c) and (5d), the semantic connections associated with *monja*, “nun”, may lead the parser to assume that the temporarily ambiguous RC refers to the brother as it processes *involucrad*. However, the grammatical gender can be manipulated to force the interpretation that the nun was involved in the nefarious event, thereby allowing researchers to investigate how semantic plausibility moderates the use of disambiguating information.

Flexible word order in Spanish also offers avenues for researchers to study ambiguity resolution. For example, Teubner-Rhodes et al. (2016) investigated how Spanish-Catalan bilinguals recovered from the subject-first/object-first cleft sentence dichotomy that creates garden paths (from Teubner-Rhodes et al., 2016, examples 1-2):

(6a) *Este es el general que vigilaba al espía desde la ventana.*

(6b) *Este es el general que vigilaba el espía desde la ventana.*

In (6a), the animate direct object marker *a* establishes that the embedded structure follows the subject-first cleft structure. Although *a* is not particularly salient, this is the preferred structure in Spanish, and the object-first structure in (6b) leads to increased reading times and processing difficulties (Betancort et al., 2009; del Río et al., 2011). In English, with its more stringent word order, this ambiguity is not manipulatable, because cleft structures are restricted to SVO order, with *wh*-extraction of one of the two NPs (although other ambiguities associated with *wh*-extraction have been used to study different

Finally, researchers have compared processing of Spanish and of English as a means of comparing how different language typologies affect processing. For example, English, an obligatory explicit-subject language, has an ambiguity that arises from transitive verbs in absolute constructions (e.g. “John reads every night”) compared to in oblique constructions (e.g. “John reads the news every night”). By moderating the location of transitive verbs in a sentence, such as in a fronted adverbial clause (e.g., 7a), and comparing resulting reading times to those of a comparable intransitive control (e.g., 7b), researchers have added to our understanding of ambiguity resolution. Fronted transitive verbs in English lead to garden paths because the post-verbal noun phrase is initially interpreted as the object of the transitive verb in the oblique construction, but it must be reanalyzed as the subject of the matrix clause once the parser arrives at the matrix verb.

(7a) When the sculptor finished the work was three meters tall.

(7b) When the sculptor came back the work was three meters tall.

(7c) Cuando el escultor acabó la obra tenía tres metros de altura.

(7d) Cuando el escultor volvió la obra tenía tres metros de altura.

In Spanish, these same constructions create a similar dichotomy (consider 7c-d, from Jegerski, 2012, her examples 9a-b). Spanish is a null-subject language, and so the conflict in these stimuli differs slightly. First, the intransitive verb in the fronted adverbial clause followed immediately by a noun phrase triggers a conflict between the Late Closure principle and the intransitive verb’s subcategorization matrix, which manifests as slower reading times for the intransitive condition (e.g., 7d) at Region 3. The transitive verb in the
fronted clause in (7c) can be interpreted to be in the oblique construction with the null-subject pro serving as the subject of the matrix clause (Jegerski, 2012, p. 723), which conforms to Frazier’s Late Closure principle. Therefore, in Spanish, no reanalysis is forced by the transitive conditions. However, the transitive sentences in Jegerski’s (2012) study did reveal processing difficulties. Comprehension accuracy for transitive sentences was significantly lower than in the intransitive condition, and there was a descriptive trend of faster reading times for the intransitive condition at Region 4 (faster native readers: 924 ms vs. 873 ms; slower native readers: 1292 ms vs. 1254). Given Spanish’s preference for the referents of pro-subjects to also be subjects (Jegerski, VanPatten, & Keating, 2011), the increased reading times at Region 4 in the transitive condition were attributed to forcing this non-preferred referent-assignment (e.g., assigning la obra, a direct object, as the referent of the pro-subject).

These findings may alternatively provide some evidence for a constraint-based model in line with Trueswell and colleagues, which would maintain multiple possible structures available until they become incompatible. That the parser must maintain two active possibilities upon reaching the end of the linguistic input, with no disambiguating information available in the Spanish syntax, may help account for the processing difficulties that Jegerski (2012) found.

Véliz, Riffó, Aguilar and Sáez (2011), who do find a garden path effect in both young adults and seniors who read these structures, argue that they do involve reanalysis because of the semantic reassignment involved. For example, in the example in (6c), while la obra receives the thematic role of goal in the first pass, further input (“tenía…”) forces either (i) reanalysis of the status of the transitive verb in the fronted adverbial clause (whether it is
in an oblique or absolute construction, in other words whether to interpret *la obra* as a theme or an experiencer) or (ii) the disambiguation of the referent of the null subject *pro*. As Dussias and Piñar (2010) point out, this process requires a change in theta-role and abstract Case from accusative to nominative in the syntax (p. 448-449; also see Juffs & Harrington, 1996; for a modified view, see Juffs, 2005; Jackson & Dussias, 2009). The two processing difficulties, for the intransitive condition at Region 3 and for the transitive condition at Region 4 (with spillover to Region 5), will be the focus of the current dissertation.

### 2.2 Cognitive Control

Cognitive control (CC), the ability to resolve information conflict, is required to correctly interpret all of these garden paths and syntactic ambiguities. I assume the term ‘cognitive control’, as opposed to other terms such as ‘inhibitory control’, following other researchers who reason that conflict can be resolved through inhibition of irrelevant information or promotion of relevant information, or a combination of both (Botvinick et al., 2001; Teubner-Rhodes et al., 2016). The use of online measures has allowed researchers to isolate the region of the brain involved in CC: linguistic ambiguity resolution and non-linguistic conflict resolution processes co-localize within the LIFG and ACC (Thompson-Schill, Bedny, & Goldberg, 2005; Novick et al., 2009; January, Trueswell, Thompson-Schill, 2009, Teubner-Rhodes et al., 2016). This region, then, can be considered the center for the correct selection of information during goal-specific tasks, and in the case of sentence processing, it allows the abandonment of an incorrect initial analysis to facilitate correct comprehension (Novick et al., 2005).
2.3 Second Language Learning and Cognitive Control

In addition to CC’s role in ambiguity resolution, cognitive control and other measures of cognitive capacity have been studied for their role in language processing by L2 learners, in psycholinguistics (Bialystok, Craik & Luk, 2008; Kroll & Sunderman, 2005, etc.), neuroimaging and cognitive psychology (Abutalebi, 2008; Abutalebi & Green, 2008; Indefry, 2006, etc.), and L2 acquisition (SLA) and applied linguistics (Linck et al., 2013; Serafini & Sanz, 2016, etc.). The extent of attention that cognitive control has received varies across these fields, with more attention in cognitive psychology and psycholinguistics than in SLA, which has prioritized other cognitive capacity variables. Important findings from these fields are addressed below.

Collecting neuroimaging data from several fMRI and PET studies of neural structures during bilingual processing, Abutalebi (2008) observed increased activity of the LIFG and other prefrontal structures critical for cognitive control among participants processing a non-native, non-proficient language. The suggestion of this finding is that, as the speaker attains a sufficient level of L2 proficiency, this extra activity fades (Abutalebi & Green, 2008). Bear in mind that the role of cognitive control is to promote relevant information (or suppress irrelevant information, see Botvinick et al., 2001), which in the case of bilingual processing can be considered competing forms and linguistic structures, such as conflict between languages. As repeated activation strengthens these networks over time and at increasing proficiency, the need for cognitive control to intervene decreases (Fedorenko & Thompson-Schill, 2014). Hernández and Meschyan (2006) found increased activity in the anterior cingulate gyrus and dorsolateral prefrontal cortex during L2 picture naming when compared to L1 picture naming, suggesting that lexical access in a less fluent
language involves more attentional effort, and Grant, Fang and Li (2015) conducted a longitudinal study and found similar results to these prior studies. In Grant, Fang and Li’s study, the researchers measured brain activity using fMRI during lexical decision-making in intermediate-advanced learners twice over the course of an academic year. They found that, after a year of classroom exposure, the neural activity observed in the first round of imaging decreased in regions associated with cognitive control such as the ACC. Like Abutalebi’s (2008) findings, this suggests that increased exposure to the L2 leads to decreased recruitment of CC-associated brain regions. Researchers disagree whether this decreased activity is a convergence to the activity involved in L1 activity or if the decreased activity is better understood within the notion of neural organizational efficiency, or in other words, whether L2 processing takes on patterns of L1 processing at higher proficiencies or whether it simply becomes more efficient (Indefrey, 2006). However, for our purposes, the observation that there is decreased hemodynamic activity in these regions is sufficient to understand that their influence lessens with increased L2 proficiency.

2.3.1 Cognitive capacity in second language acquisition

The conclusions from neuroimaging studies reflect findings from behavioral studies on language learning as well, but most studies in SLA have used different proficiency levels, different linguistic structures and processes, and importantly, different operationalizations of cognitive capacity, with the most focus on more holistic constructs like executive function (EF), working memory (WM), or phonological working memory (PWM). EF is a family of cognitive processes that include inhibition (both response [1]

[1] To unite disparate terminology and generalize similar findings, I use PWM to refer to any operationalization that references the phonological loop of Baddeley or uses a nonword repetition task or digit span task, whether termed phonological memory, phonological short-term memory or phonological working memory.
inhibition and interference control), working memory, and cognitive flexibility (Diamond, 2013). WM is likely the most widely used conceptualization of cognitive capacity in SLA literature, and it is generally agreed that WM is a limited-capacity storage and processing unit (Baddeley, 2007; Conway, Jarrold, Kane, Miyake, & Towse, 2007; Miyake & Friedman, 2012; Williams, 2012). Far and above the most widely referenced model in the L2 literature, Baddeley’s (1986, 2000, 2003) model has three domain-specific components dedicated to visual processing (the visuospatial sketchpad), phonological processing (the phonological loop, or PWM), and bidirectional communication with long-term memory (the episodic buffer). PWM is responsible for the processing and storage of phonological information. The studies that consider these cognitive capacity variables will be addressed first, followed by a review of the limited literature on cognitive control in SLA.

2.3.2 Working memory, phonological working memory and executive function in second language acquisition

To understand how the role of cognitive capacity in the L2 is understood by scholars, we can first consider a recent study on both PWM and EF. In a longitudinal study investigating the relationship between cognitive variables and the automaticity of morphosyntactic structures for L2 Spanish learners, Serafini and Sanz (2016) found robust correlations for learners at beginning and intermediate proficiency levels. Serafini and Sanz used the operation span and digit span tasks to measure EF and PWM, respectively. Linguistic knowledge was measured using an untimed grammaticality judgment task (GJT) testing knowledge of ten linguistic structures and an elicited oral imitation task used as a measure of L2 Spanish knowledge (Bowles, 2011). Both these tests were administered at three points: at the beginning and end of one academic semester, and again for a reduced
set of the original participants after four weeks without instruction. While a correlation between PWM and the online oral imitation task for beginner and intermediate learners was identified, no such relationship was found for advanced learners. PWM and the offline (i.e., untimed) GJT also correlated for these groups, especially for intermediate learners. Higher EF also correlated to greater performance for these groups, though the effect sizes were smaller. EF also predicted variance in development on the oral imitation task for beginners, suggesting that at lower proficiency levels, learners “relied heavily on controlled attentional processing (Engle, Kane, & Tuholski, 1999) to monitor their speech production while completing the OIT” (Serafini & Sanz, 2016, p. 639). Unlike beginner and intermediate learners, advanced learners did not obtain any robust relationships between these cognitive variables and their morphological knowledge of Spanish, suggesting a decreased reliance on EF and PWM at increasing proficiency levels. It is important to note that the same linguistic structures were assessed for all three groups, so while advanced learners did not reveal any reliance on WM, this may be a reflection of the early-acquired structures under study. Regardless, the findings can be related to decreased general effort or increased automaticity at higher L2 proficiencies.

Serafini and Sanz are not the first researchers to have noted the waning importance of WM at increasing proficiencies. For example, they note: “studies tend to more consistently report positive effects for lower proficiency learners whereas findings for advanced learners with more exposure to and practice in the target language are mixed” (Serafini & Sanz, 2016, p. 614). While this observation is true, another reading of the

2 Excellent reviews of the role of WMC and L2 development and processing have been published (see Michael & Gollan, 2005; Linck et al., 2013; and Kroll & Linck, 2007), but this dissertation will present a limited overview, given its focus on CC.
literature suggests (1) that the null results that have been found for lower proficiencies are reflections of learners’ inability to leverage WM for certain tasks that are still too complex for these low proficiencies, and more importantly, (2) that WM functions correlate to increasingly complex phenomena at increasingly higher proficiencies.

In terms of the first suggestion regarding null findings at lower proficiencies, we must consider the linguistic target under study and how that might interact with proficiency. Participants at low proficiency levels may not be under such low-level processing constraints that neither low-capacity nor high-capacity individuals can leverage WM to reveal differences. For example, Kormos and Sáfár (2008) find that while WM accounts for 30 percent of the variance among beginners in performance on the standardized Cambridge First Certificate Exam, PWM does not correlate to any subcomponents of the exam for beginners, but does correlate to oral fluency and vocabulary gains in their pre-intermediate group. In another set of studies, sensitivity to online gender and number agreement violations only correlated to reading span for the intermediate (seventh- or eight-semester) students and not the beginning (third-semester) students (Sagarra, 2007; Sagarra & Herschensohn, 2010). In these studies with null results for early-stage learners but not higher-level learners, the components of WM in the lower group may be dedicated to other more pressing tasks of L2 use (including L1 suppression). As these processes become more automatized, less effortful, or more convergent with L1 processes, WM functions can be allotted to these more advanced language processes.

In what follows, references to WM reflect the use of Daneman & Carpenter’s (1980) reading span or Water & Caplan’s (1996) sentence span (reading or listening). These studies included these tasks in at least the L1 and in some cases the L2, but in that case, without leading to disparate findings.
This idea also feeds into the second proposed interpretation of the literature, mentioned above. The mixed findings reported in the literature for advanced learners may be a reflection of a phenomenon whereby the role of individual differences in WM in distinguishing L2 development and processing is a function of both proficiency and the complexity of the L2 structure under study. In other words, as the other side of the same coin as above, finding correlations between WM functions and a certain structure at beginner and intermediate but not at advanced or post-advanced levels likely suggests the automatization or convergence of that structure for the late-stage learners, for whom WM is freed up to facilitate processing of more advanced structures. For example, for true beginners, WM and PWM has been associated with vocabulary learning and making associations between vocabulary and inflectional morphology like gender (Kempe & Brookes, 2008; Kempe et al., 2010; Williams & Lovatt, 2003); for intermediate learners, WM is correlated to more noticing of feedback (Mackey, Philp, Egi, Fujii & Tatsumi, 2002), and PWM is correlated to oral fluency development (O’Brien, Segalowitz, Freed, & Collentine, 2007); for learners with at least four semesters completed, WM has been correlated to online sensitivity to agreement violations (Coughlin & Tremblay, 2013; Sagarra, 2007; Sagarra & Herschensohn, 2010), and to reading comprehension (Harrington & Sawyer, 1996). However, for post-advanced learners with at least three years of formal L2 study, cognitive capacity has not been associated with sensitivity to subject-verb number agreement or noun-adjective gender agreement (Foote, 2011), nor to noun-adjective number or gender agreement, word order, or lexical recognition (Grey, Cox, Serafini & Sanz, 2015), but WM has been correlated to native-like use of semantic plausibility (Dussias & Piñar, 2010) and to syntactic comprehension and native-like cue
use (Miyake & Friedman, 1998). Consider the differences between the last two groups4. WM has been shown to play a role in agreement sensitivity for advanced learners, but this role dissolves as learners progress, freeing up WM resources. This in turn allows high-capacity learners to use more complex linguistic cues to facilitate comprehension and processing.

This is epitomized by Havik, Roberts, Van Hout, Schreuder and Haverkort’s (2009) observation of German students with advanced knowledge of Dutch processing subject-object RC ambiguities of longer or shorter length. Havik et al. found that the native speakers have a processing advantage for subject-first RCs over object-first RCs. In the native speakers with low WMC, this advantage was due to an interaction of the RC’s type and length. They had particular difficulty processing long object-first RCs. This same pattern was found in the high WMC learner group, but not for the low WMC learner group, for whom the authors posit that the task was too difficult. To see differences among the advanced group of learners that relate to WMC, Havik et al. included a task that was challenging even for native speakers, and as a result, they were able to see that WM does play a differential role even for late-stage language learners. WM is a monolithic cognitive construct conceptualized as a multi-modal aspect of short-term memory that emerges during cognitive tasks (Baddeley, 1986, 2000; Baddeley & Hitch, 1974; Cowan, 2008), and may even include long-term memory as a contributing module (Ericsson & Kintsch, 1995; Miller, 1956). However, as addressed previously, cognitive control is a specific

4 These groups in this section are a reflection of a synthesis of the literature and not of the terms used in the original studies. This was to allow some form of cross-study comparisons. For example, consider Sagarra & Herschensohn’s (2010) in-text terminology (beginner vs. intermediate) and their participants’ experience (three semesters vs. seven or eight semesters). The intermediate learners in Sagarra & Herschensohn (2010), after seven or eight semesters of language study, are dubiously comparable to O’Brien et al.’s (2006) “intermediate” learners, who only completed two semesters, even though the authors’ terminology denotes them all as “intermediate.”
capacity with a specific linguistic correlate: conflict resolution. Therefore, by comparing ambiguity resolution to the “default” cognitive variable in the L2 literature (i.e., WMC), the study introduces the possibility for other components of WM introducing confounding results.

In other words, given that WM is a monolithic variable, not every operationalization taps that cognitive resource in the same way, and therefore, not every linguistic structure will correlate to WM operationalizations in practice, nor are they expected to in theory. For example, consider the findings of Harrington and Sawyer (1992). In this study, the researchers compared three L2 English reading comprehension measures (the reading and grammar sections of the TOEFL and a cloze passage) to three cognitive capacity measures in the L2 (digit span, reading span and word span). 5 Predictably, the reading span correlated with both TOEFL sections (TOEFL Grammar and reading span, $r(30) = .57, p < .001$; TOEFL Reading and reading span, $r(30) = .54, p < .001$). The correlation between reading span and the cloze passage did not reach significance at the established alpha, but followed the same trend (cloze passage and reading span, $r(30) = .33, p < .06$), while the other span scores did not approach significance. These findings from Harrington and Sawyer (1992) are a clear example of why it is so important to select cognitive variables and operationalizations of those variables carefully during psycholinguistic research: had Harrington and Sawyer selected a different subset of cognitive capacity tests, they may have come to very different conclusions. They motivated the inclusion of the three tasks precisely because the reading span task taps both higher

---

5 The researchers also collected data for these same cognitive measures in the L1. The L1 span scores all correlated with their L2 equivalent ($p < .05$). However, L1 span scores were not correlated to the L2 comprehension measures.
order processing and storage functions of WM, while digit and word spans have not been shown to relate to such higher order skills, and the researchers wanted to show how L2 reading involves these higher order processes of WM.

Another approach to disentangling distinct aspects of cognitive capacity has been utilized by Kempe, Brooks and colleagues (Brooks, Kempe & Sionov, 2006; Kempe & Brooks, 2008; Kempe, Brooks & Kharkhurin, 2010). In these studies on novice learning of Russian in the laboratory setting, the authors consider general intelligence using Cattell and Cattell’s (1973) Culture Fair intelligence test. However, in the latter two studies, they also incorporate measures for WMC with Daneman and Carpenter’s (1980) reading span task. The selected operationalizations both include aspects of EF (Cowan, 2000; Kane & Engle, 2002), so by conducting a multiple regression analysis, they partialled out the shared variance to determine which findings related to EF and which to other subcomponents of general intelligence or working memory. They posit that EF is involved in noticing inflectional changes and segmenting lexical and inflectional morphology, but that WMC’s storage component is related to the memorization involved in learning the new words, stems, and affixes (Kempe & Brooks, 2008; Kempe, Brooks & Kharkhurin, 2010). This design allows the researchers to look at the multi-modal variables they chose while still being able to draw specific conclusions about how the different aspects of language learning interact with these different modules, and this can help us create more specific models for different stages of language learning.

Like Kempe and colleagues, Serafini and Sanz (2016) included multiple cognitive capacities to measure their relation to the acquisition of the same linguistic targets across increasing proficiencies to model how these capacities differentially effect language gains.
Because they include ten linguistic structures, however, they motivate the need for future studies to isolate fewer specific targets to increase subsequent studies’ power. Syntactic ambiguity resolution, directly associated with CC, is one target with a clear cognitive correlate that can answer that call. However, the behavioral studies that investigate cognitive control in the L2 are much fewer than those of the cognitive variables addressed above. These studies of cognitive control are addressed in the section that follows.

2.3.3 Cognitive control in second language acquisition

As has been highlighted, while most research in SLA uses WM or PWM, and to a lesser extent, EF, the current dissertation investigates the role of cognitive control in L2 learners. Cognitive control has received limited attention, but for example, it has been found to predict improved fluency, specifically decreased reformulations or false starts and reduced silent pauses in terms of both frequency and duration (Korko & Williams, 2016); and improved phonological segment perception (Darcy, Mora, & Daidone, 2016). This dissertation’s focus on CC, however, addresses two specific and established correlates: first, how cognitive control is related to ambiguity resolution in the L2, and second, how decreased recruitment of CC-related brain regions at increasing proficiencies is represented behaviorally. Our current understanding of how these three constructs interact, based on limited SLA, are addressed below.

Linck and Weiss (2011) investigated the role of both WMC and cognitive control on the acquisition of a L2 over the course of a semester, and found that WMC predicted gains during one semester in early-stage learners, while cognitive control was not a predictor. This finding is often used to discuss proficiency and cognitive capacity, although there are some issues with this interpretation, since beginning German and intermediate
Spanish participants were combined when no group differences were found in the statistical model’s predictor or criterion measures in the preliminary analyses, even though these participants completed different language measures (German: fifteen fill-in-the blank items from University of Wisconsin placement exam; Spanish: grammar and vocabulary section of the Diplomas de Español como Lengua Extranjera, DELE). That being said, WMC correlated with greater L2 proficiency at both the beginning and end of the semester, and was also found to be a predictor of change in proficiency from beginning to end. No significant results were obtained for CC. This is not entirely surprising, given the nature of the cognitive variables and of the language measures used. WM is a measure of storage and processing, which are both likely to play a role in classroom learning. While we would expect cognitive control to reveal itself as a predictor because L1 suppression is more laborious at early-stage learning, the language measures are offline and so likely do not require the participants to recruit cognitive control as much as WM or even long-term memory.

In addition to seeing less recruitment of the brain regions associated with cognitive control (i.e., the ACC, the LIFG), behavioral cognitive advantages have also been observed. In other words, bilingualism has been shown to benefit cognitive capacities in groups beyond the early native bilingual addressed earlier. For example, Poarch and van Hell (2012) found that bilingual and trilingual children revealed greater cognitive control than monolingual children in both the Simon and ANT tasks, as predicted and fitting within the earlier review. However, an additional group, children learning a non-native L2, revealed a descriptive advantage in the Simon task over the monolinguals and a descriptive disadvantage compared to the bilingual and trilingual groups. However, the early bilingual
and trilingual groups revealed an advantage in conflict resolution in the ANT task. This would indicate that bilingual advantages are continuous, not acquired through bilinguals’ “nativeness” or lack thereof, but rather through the practices associated with it, and redeemable to greater or lesser degrees. While Poarch and van Hell (2012) study children and this dissertation focuses on adult L2 learners, other research supports this same finding that cognitive control advantages fall on a continuum. Linck, Hoshino, and Kroll (2008) found evidence for a late bilingual advantage in CC. Comparing the Stroop effect as measured by the Simon task (Simon & Rudell, 1967) for adult L2 learners at different stages of proficiency, the authors found that greater exposure to the L2 revealed greater conflict resolution abilities on the Simon task, suggesting that bilingual advantages can be tapped into by non-native learners, even those who begin their bilingual experience after puberty. In addition, this has been supported by ERP data from Sullivan, Janus, Moreno, Astheimer and Bialystok (2014), where six months of introductory Spanish coursework revealed neural evidence of improved performance on a go-nogo task and a sentence judgment task (i.e., larger P3 amplitudes and smaller P600 amplitudes, respectively), compared to a control group. Despite this neural evidence, the groups did not differ behaviorally. This change indicates improved cognitive control following early-stage L2 learning in the short-term.

2.3.4 Second language learners, ambiguity resolution and cognitive control

A recent psycholinguistics study also merits attention in this conversation, although cognitive control is not directly included in the research design. Pozzan and Trueswell (2016), studying intermediate learners of L2 Italian, find that these adult subjects reflect child behavior during online ambiguity resolution, even when the structure under question
exists in their native language (English). Specifically, they observed that L2 adults used referential information in the input like native adults, but revealed difficulties abandoning incorrect initial parses similar to native children. In children, this so-called ‘kindergarten-path’ has been associated with immature cognitive control as a result of protracted maturation of prefrontal cortical structures (Choi & Trueswell, 2010; Novick et al., 2005; Trueswell, Sekerina, Hill, & Logrip, 1999; Woodard, Pozzan, & Trueswell, 2016). Pozzan and Trueswell (2016) posit that the similarities between children in their L1 and adults in an L2 reflect that garden-path resolution may compete with other restraints on CC, namely immaturity in the case of children or L2 suppression in the case of non-proficient adult L2 learners. That is, similar results may be obtained not because the structure is difficult for both sets of language learners, but because low-proficiency L2 processing occupies cognitive control resources in adults (Abutalebi, 2008), making comprehension run aground in cases when revision is needed. However, because they do not include cognitive control measures or advancing proficiency levels, they propose this question as a direction of future research. The current dissertation attempts to answer this call through the use of a paradigm described below in Section 2.5.

2.4 Early Bilingual Sentence Processing and Cognitive Control

Having addressed the role of cognitive control in late L2 learners, it is also necessary to discuss the role of cognitive control in another population of bilinguals: early bilinguals who grow up speaking two languages (and therefore consistently suppressing one language), thereby exerting their cognitive control at critical stages of development. Psychological research has revealed that for many populations, the increased use of a certain behavior results in observably improved skills on similar tasks or in observable
differences in associated brain structures. For example, architects outperform non-architects on tasks evaluating visuo-spatial ability (Salthouse & Mitchell, 1990), and video game playing has been correlated to heightened modification of perceptual-motor ability (Abutalebi et al., 2012). Just as video-gamers and architects serve as easily identifiable population distinctions, so too is the bilingual, although the profiles of bilingualism are more diverse and complex than those of architects or gamers (see Kroll & Bialystok, 2013, and Montrul, 2005, for example). Psychologists and psycholinguists have shown that bilinguals differ from their monolingual counterparts in many different facets, both physiologically and behaviorally (see Bialystok, 2008), but for the purposes of the current dissertation, we can look specifically to bilingual and monolingual processing of conflict.

One reality of bilinguals that distinguishes them from their monolingual peers is the consistent need to suppress one language according to the context. The bilingual does not separate the mental lexicons and grammars, but rather both are active during language use in any form (Dijkstra, 2005; Kroll & de Groot, 1997; Marian & Spivey, 2003; Kroll, Bobb & Wodniecka, 2006; also see Kroll, Dussias, Bogulski, Valdes Kroff, 2012, for a review). This perpetual activation was the subject of Green’s (1998) Inhibitory Control (IC) model of bilingual language processing, and has been attested in behavioral studies (e.g. Kroll & Bialystok, 2013; Kroll & Sunderman, 2005); imaging studies (e.g. Abutalebi et al, 2012); and patient data (e.g. Abutalebi, Miozzo & Cappa, 2000).

Within Green’s (1998) IC model, the competitor is suppressed by inhibitory control (i.e., cognitive control). Given two competing forms, the salient but contextually-infelicitous form is suppressed by cognitive control while the desired form is promoted. This model and much of the early research that followed investigated whether this constant
activation of cognitive control led to improved performance of bilinguals on nonlinguistic tasks, based on the idea that consistent use of cognitive control would lead to improved cognitive control (see Bialystok, 2008, for a review). This has become known as the ‘bilingual advantage’. Bilingual advantages have subsequently been attested through neuroimaging and behavioral studies. In neuroimaging, larger gray matter volume has been observed in bilinguals’ brain regions serving executive functioning (Olulade et al., 2015). Likewise, the ACC and the LIFG, both neural structures involved in conflict detection and resolution, have been observed to be activated more by monolinguals than by bilinguals when performing high-conflict tasks, which is an indication of more effortful processing of conflict (Abutalebi et al., 2012, Abutalebi, 2008). This tendency in the results has been confirmed in a meta-analysis conducted by Adesope, Lavin, Thompson and Ungerleider (2010), comparing cognitive control of monolingual and bilingual populations. As Teubner-Rhodes et al. (2016) suggest, “bilingualism apparently acts as a form of cognitive control training, bestowing measurable advantages in conflict monitoring – the ability to detect unpredictable conflict and flexibly adjust recruitment of cognitive control resources” (p. 227).

This same phenomenon that serves as cognitive control training for bilinguals also presents difficulties or ‘disadvantages’ in linguistic development or language use. These include decreased vocabulary size in one language (e.g. comparing English vocabulary size of monolingual English speakers and bilingual speakers who speak English and another language, Bialystok, Luk, Peets & Yang, 2009; Bialystok & Luk, 2011); slower lexical access in picture-naming tasks (Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007) and in lexical decision tasks (van Hell & Dijkstra, 2002); and decreased verbal fluency in
semantic (or categorical) and phonemic (or letter) fluency tasks (e.g. Gollan, Montoya & Werner, 2002; Rosselli et al., 2000). Nonetheless, when bilinguals are compared to monolinguals of equivalent vocabulary size, bilinguals have performed better than monolinguals on verbal fluency tasks (Bialystok et al., 2008). These differences, among many others observed in the research of Bialystok, Gollan, and other colleagues, are attributed by the researchers to the perpetual activation of both languages, thus causing competition and conflict between language forms from the target and non-target languages.

It is important to consider that these disadvantages relate to one role of CC: the suppression of one language during the use of the other. For example, lexical access during the semantic fluency task is argued to be slower because the speaker is searching through a greater raw number of words that overlap semantically in order to select those that are appropriate under the circumstances of the task (i.e., are semantically relevant and in the relevant language). However, as has been addressed earlier in this dissertation, there is another role of cognitive control that is important in language processing: ambiguity resolution. Given what we know about how monolinguals employ cognitive control during ambiguity resolution, it is important to ask how the early bilingual population resolves this type of conflict, given that cognitive control is involved both in language suppression and in garden-path recovery. Limited studies have taken on this question.

Previous studies suggest that bilinguals do suffer from more interference during language processing than their monolingual peers (Paap & Liu, 2014), but they are efficient sentence processors who are particularly skilled at monitoring for conflict in the input (Moreno et al., 2010; Teubner-Rhodes et al., 2016). While Paap and Liu (2014) found that bilinguals suffer more interference on the homograph-interference task (see Chapter 1),
this task requires cognitive control only after the sentence processing is completed, in order to suppress interfering homograph meanings. However, in studies that require cognitive control during sentence processing, bilinguals have revealed advantages over monolinguals. For example, although Moreno et al. (2010) found that monolinguals were more accurate during an acceptability judgment task (AJT), which includes both semantic and syntactic errors, the two groups performed similarly on a GJT, which requires the participants to suppress their response to semantic anomalies and only respond to the grammar of a stimulus. In addition, ERP analyses in this study revealed less neural activation from the bilinguals in response to syntactic violations, even when performing at the same accuracy, suggesting less effortful processing (Moreno et al., 2010, p. 566). An even more direct finding came from Teubner-Rhodes et al. (2016), who conducted a study on linguistic and non-linguistic conflict resolution, using garden-path sentences and N-back tasks. They found that bilinguals had more accurate sentence comprehension in both ambiguous and unambiguous sentences, and they outperformed monolinguals on the high-conflict N-back but not on the no-conflict N-back. These findings suggest that bilinguals’ advantage relates to monitoring for conflict during these demanding or conflictive tasks (and perhaps not to conflict resolution itself).

While the previous literature has provided us with some understanding into differences between bilingual and monolingual processing, the cognitive-control engagement paradigm (Hsu & Novick, 2016; Huang et al., 2016) can allow us even more insight. Whether early bilinguals cope easily with the added conflict involved in sentential ambiguities thanks to their lifelong cognitive control training or whether they react differently from monolinguals to additional linguistic conflict because their cognitive
control is otherwise engaged is a question that remains for empirical research. Given our limited understanding of ambiguity resolution in bilinguals, we can use the cognitive-control engagement paradigm, addressed below, to ask how bilinguals respond to ambiguity under different cognitive control states.

2.5 Conflict Detection and the Cognitive-control Engagement Paradigm

Conflict in our input has an associated processing cost. In psychology and psycholinguistics, this cost is often known as the Stroop effect, named after the Stroop task and its creator (Stroop, 1935). In the original of its many incarnations, the Stroop task requires participants to respond to the color of the ink in which a color-word, such as “green” is written, rather than reading the color-word itself, thereby suppressing the instinct to read the word (Stroop, 1935). The Stroop effect is one of the most important findings in modern experimental psychology. Since it was first published in English over 80 years ago, the Stroop task has been the subject of hundreds of studies and has been used in many distinct branches of cognitive science (MacLeod, 1991). Manipulations of the Stroop task all hinge on presenting conflicting information to examine the role of interference in different cognitive processes. The consensus of these Stroop-like interference tasks is that an incongruent trial is more difficult than a congruent trial, revealed by slower reaction times (RTs) and higher error rates (see MacLeod, 1991 for a review).

The Eriksen flanker task was developed with similar questions in mind, isolating cognitive control procedures while mitigating linguistic interference by eliminating words from the paradigm (Fan, McCandliss, Sommer, Raz & Posner, 2002). This task presents a center target arrow flanked by either congruently or incongruently oriented arrows. Participants must resist responding to the four flanking arrows and respond strictly based
on the center arrow’s direction, and the processing cost is the difference between the incongruent and congruent conditions. Due to the simple design utilizing non-linguistic conflict, the task can be used with populations with known executive function problems, such as children, patients and monkeys, and can be used to make conclusions about domain-general cognitive functions (Fan et al., 2002).

Much of the literature that utilizes these tasks considers the role of the anterior cingulate cortex (ACC) and cognitive control in the associated processing cost. The shared resources involved in ACC activation, CC, and conflict resolution have been well documented in the literature (Thompson-Schill et al., 2005; Novick et al., 2009; January et al., 2009, Teubner-Rhodes et al., 2016). And while it has been understood for many years that the conflict involved in these tasks engages CC, researchers have also understood that the Stroop or flanker effects extend beyond a trial-by-trial basis. The detection of conflict during a given trial has been shown to initiate sustained cognitive control procedures, specifically conflict resolution, that continue into the subsequent trial. For example, Logan and Zbrodoff (1979) observed that manipulations to the probability of conflict stimuli modulate the Stroop effect, where higher probabilities of conflictive trials result in proportionally faster responses on incongruent trials. Likewise, Gratton, Coles and Donchin (1992) showed that the preceding trial’s congruency contributes to the speed of a RT on the subsequent trial, specifically that an incongruent critical trial immediately preceded by another incongruent trial (condition i-I) leads to faster RTs than when incongruent critical trials are immediately preceded by congruent stimuli (condition c-I). Theoretically, these results have been attributed to economy: by sustaining CC, the cost of subsequent conflict resolution is attenuated. This is known as the Gratton effect or conflict
adaptation effect (Blais, Stefanidi & Brewer, 2014; Freitas et al., 2007; Gratton et al., 1992; Kerns et al., 2004).\textsuperscript{6}

As mentioned above, non-linguistic conflict resolution has been shown to co-localize within the ACC and to share cognitive control resources with linguistic conflict resolution (Thompson-Schill et al., 2005; Novick et al., 2009; January, Trueswell, Thompson-Schill, 2009, Teubner-Rhodes et al., 2016). Domain-general cognitive control is suggested to be responsible for both behavioral adjustments to resolve conflict during information processing (Botvinick et al., 2001) and reanalysis following incorrect linguistic interpretation (Novick et al., 2005). Despite this well-established understanding, it is only very recently that researchers used the principles of sustained cognitive control engagement to develop a cross-task cognitive-control engagement paradigm. Hsu and Novick (2016) developed such a paradigm that allows within-subject comparisons of conflict resolution in the engaged and non-engaged cognitive control state. Specifically, the researchers conducted a 2x2 experiment in the visual world eye-tracking paradigm on healthy adult sentence processing by pseudo-randomly interleaving Stroop trials and act-out instructions. Researchers collected both behavioral and eye movement data as

\textsuperscript{6} It should be noted that this finding was questioned by Mayr, Awh and Laurey (2003), who observe the same phenomenon as a case of stimulus-specific priming (the “repetition-priming” account). In a conceptual replication of Gratton et al.'s (1992) study, they found that faster reactions were associated with pairs of trials in which the latter’s target was a repetition of the former’s target. However, subsequent research confirmed the earlier “conflict monitoring” account, specifically because sequential dependency effects do present themselves regardless of repetition of stimuli (Ullsperger, Blynsma & Botvinick, 2005). For example, in the second experiment of Ullsperger et al. (2005), the Gratton effect or conflict adaptation effect was still observed when participants were presented with a numerical, underlined single-digit target between 1-9 with 4 competing flankers, also a single digit. This paradigm eliminated the risk of priming because the researchers were able to exclude trial-to-trial repetition. While the researchers found slower reaction times than in the standard two-option directional flanker task, they attribute this to a speed-accuracy tradeoff associated with nine options that did not diminish the effect of conflict engagement. In other words, preceding trial incongruency in this experiment still resulted in a conflict adaptation effect, where RTs in the i-I condition were 26% faster than in the c-I condition.
participants followed the instructions, which were presented as temporarily ambiguous (RRCs) or unambiguous sentences. It is important to bear in mind that critical comparisons in this study were pairings of Stroop and sentence conditions: incongruent-ambiguous (i-A) and congruent-ambiguous (c-A). In both the action responses and the fixations to the correct goal, a significant sentence-type-by-Stroop-trial-type interaction was encountered, whereby participants made fewer performance errors or fewer fixations to the incorrect goal on i-A trials than on c-A trials. In other words, they observed that participants were less committed to incorrect initial parses of ambiguities when the preceding Stroop trial was incongruent, compared to those parses following congruent Stroop trials. This was true of both behavioral and eye-tracking data. As expected, Stroop incongruity did not modulate consideration of the correct goal in the unambiguous linguistic condition (i-U did not statistically differ from c-U), likely because there was no conflict to be resolved and online interpretation and action responses approached ceiling. These results show that cognitive control engagement helps reject incorrect parses earlier in order to correctly carry out instructions.

However, the cognitive-control engagement paradigm has revealed a different pattern of results in individuals with compromised cognitive control abilities: in children, conflict adaptation occurs in singular task designs, but not cross-task designs (Huang et al., 2016), and patients with damage to the prefrontal cortex have also shown similar cross-task difficulties (Novick, personal communication, Nov. 3, 2016). The authors suggest this may be due either to the depletion of limited cognitive control resources or to task difficulty fatigue, or a combination of these. How conflict adaptation can provide us insight into ambiguity processing and the role of cognitive control in other populations for whom
cognitive control is under conditions that distinguish it from the cognitive control of monolingual adults is the subject of the three studies of the present dissertation, each addressed below.

2.6 Rationale for the Studies

2.6.1 Rationale: Experiment 1

The cognitive-control engagement paradigm can give us important insight into early bilinguals’ cognitive control advantage. Findings investigating the transcendence of lifelong language suppression to domain-general cognitive abilities have generally supported a bilingual advantage in cognitive control (e.g. Bialystok, 2006, 2010; Bialystok et al., 2009; Bialystok et al., 2004; Costa et al., 2008; Costa et al., 2009; Martin-Rhee & Bialystok, 2008; Teubner-Rhodes et al., 2016; but also see Hilchey & Klein, 2011, and Paap & Greenberg, 2013). However, what bilinguals are able to do with that advantage in more domain-specific terms is still unclear. Based on previous studies, dynamic engagement of cognitive control moderates the processing cost associated with ambiguity resolution in monolingual subjects. This reduced cost is due to greater efficiency resolving conflict when cognitive control is already actively monitoring for conflict. However, how this same type of dynamic engagement of cognitive control would affect bilinguals is unclear. As bilinguals suppress one language during language processing, they rely on cognitive control (Kroll et al., 2012), and therefore cognitive control is already engaged more in bilinguals than in monolinguals by nature of their bilingualism; however, this level of engagement is also CC’s neutral state for these individuals. Whether further engagement through a non-linguistic conflictive stimulus can lead to subsequent improved ambiguity resolution is a question that has not yet been asked empirically, but will help us understand
bilingual advantages and language processing.

2.6.1 Rationale: Experiment 2

Given the decreased recruitment of CC-related neural structures during non-native (L2) language processing (e.g. Abutalebi, 2008), the cognitive-control engagement paradigm can serve as a tool to force specific CC-engagement states, providing within-subject data reflecting how L2 syntactic ambiguity recovery differs in different states of CC-engagement. If cognitive control plays a differential role in online processing at different proficiencies, the cognitive-control engagement paradigm may reveal different responses from different groups: when cognitive control is not engaged to capacity (e.g., at higher proficiencies), linguistic conflict preceded by an incongruent flanker trial should reveal faster disambiguation than the same linguistic conflict preceded by a congruent flanker. However, if cognitive control is fully engaged by L1 suppression at lower proficiency levels, simulating the compromised cognitive control of children or patients, these lower proficiency learners may reveal a different pattern. One possibility is that the incongruency of the preceding non-linguistic trial will not be able to engage cognitive control beyond this maximum capacity, resulting in comparable reanalysis RTs in both congruent and incongruent conditions. Another possible pattern is that L2 learners with all cognitive control resources engaged will obtain slower reanalysis following incongruent flankers, if they too experience the ‘depletion’ or fatigue that is hypothesized to cause these findings in children and patients (Huang et al., 2016). Of course, given the novel nature of this research, there is also the possibility that cognitive-control engagement will result in faster resolution of linguistic conflict for all L2 learners, regardless of proficiency, if their resources are not maximally engaged. In other words, because the linguistic stimuli are
relatively complex in nature, the available range of proficiency levels from which to recruit is narrowed at the lower end, and so we may expect participants to have enough resources available to leverage cognitive-control engagement to resolve subsequent conflict. This dearth of prior studies upon which to base predictions makes it more advisable to ask questions about proficiency only in an exploratory nature, and to use findings from the present dissertation to inform future research into L2 sentence processing and cognitive control.

2.7 Research Questions

These experiments contribute to disentangling CC’s complex relationship with early bilingual language processing and with L2 proficiency and language processing. Such research is timely, considering recent research in SLA, psycholinguistics and neuroimaging that suggests that cognitive control plays a differential role across L2 development and considering the gap in literature on early bilingual processing and the role of CC. In addition to offering insight into avenues of future research through the two exploratory research questions that follow, the studies will aim specifically to answer the confirmatory research questions that follow:

2.7.1 Research questions: Experiment 1

1. Do early bilinguals reveal the same linguistic conflict sensitivities to fronted adverbial clauses in Spanish as native monolinguals and near-native L2 speakers? (confirmatory)

2. Does cognitive-control engagement improve linguistic conflict resolution in Spanish for these early bilinguals? (confirmatory)

3. Do cognitive individual differences of these early bilinguals contribute to understanding the resolution of linguistic conflict in the cognitive-control engagement paradigm?
2.7.2 Research questions: Experiment 2

4. Do emergent bilinguals reveal the same linguistic conflict sensitivities to fronted adverbial clauses in Spanish as native monolinguals and near-native L2 speakers? (confirmatory)

5. Does cognitive-control engagement improve linguistic conflict resolution in Spanish for these emergent bilinguals? (confirmatory)

6. Does proficiency modulate the resolution of these constructions or the effect of cognitive-control engagement? (confirmatory)

7. Do cognitive individual differences of the emergent bilinguals contribute to understanding the resolution of this linguistic conflict in the cognitive-control engagement paradigm? (exploratory)
CHAPTER 3: RESEARCH DESIGN

3.1 Overview

This chapter begins with an overview of how the main constructs in the design and conceptualization of the present dissertation were operationalized, namely early and emergent bilingualism, processing and comprehension, incongruency and transitivity, and proficiency. The rest of the chapter is dedicated to the specific aspects of the research design for Experiments 1 and 2. Many of the elements of the design were informed by two small pilot studies, presented at the end of the chapter. However, before presenting the results of the pilot studies and the changes that it motivated, the final design and methodology of Experiments 1 and 2 are presented, including the selection of participants, materials, procedures, scoring and coding protocols, and statistical analyses.

3.2 Operationalizations

*Early bilingualism* is operationalized in this dissertation as self-reported simultaneous and continuous use of two languages during childhood linguistic development, beginning minimally before the age of 5.

*Emergent bilingualism* is operationalized as the use of a L2 outside of the home, learned in an academic setting, with fluency in this L2 developed post puberty over many years. Among the emergent bilinguals in this dissertation, there were four individuals who were exposed to a second language in the home before age 5, but this exposure was neither continuous throughout their linguistic development nor simultaneous with their native English development, and all four reported Spanish to be their second most dominant language after English.
Linguistic processing is operationalized as reading time latency, the time during which a given word or region was exposed by the PsychoPy presentation script before the participant continued to the next word.

Comprehension is operationalized as correct responses to the comprehension questions that followed all linguistic stimuli. Only correctly comprehended stimuli and participants achieving at least 60% mean comprehension scores were included in final analyses, and three consecutive incorrect responses to CQs resulted in a ten-second break and a reminder that timely completion of the study required participants to read the sentences both carefully and at as natural a pace as possible.

Incongruency in the chapters that follow refer specifically to a flanker stimulus with flanking letters that are different from the response trigger. In other words, while “HHHHHHHH” is a congruent flanker stimulus, “HHHSHHHH is incongruent.

Transitivity refers to whether the verb in the fronted adverbial clause, at Region 2, was presented in the transitive or intransitive conditions. However, it is important to remember that comparisons of items at Regions 3, 4, and 5 include the exact same linguistic input.

Proficiency in the L2 for emergent bilinguals is operationalized as their score on the Elicited Imitation Task, which has been validated as a measure of global proficiency. This operationalization allows continuous values to be assigned to participants instead of institutional levels or other broad measures that obscure individual differences.

3.3 Selection of Participants

Two populations were targeted in this dissertation in order to fill gaps in the literature: first, a homogenous group of early bilingual university students was recruited
from Barcelona, Spain; and second, a heterogenous group of L2 learners of Spanish raised in the United States and at advanced, though different, stages of L2 acquisition. Participants were excluded from Experiment 1, the early bilingual study, if they acquired Spanish or Catalan after the age of 5 or if they reported any extended periods of minimal use of one of the two languages. Participants were excluded from either study if they reported an uncorrected vision deficiency, a dyslexia diagnosis, or an attention disorder diagnosis. To control the stage of attentional development as much as possible, participants under the age of 18 and over the age of 30 were also excluded from participation.

3.4 Target Structures

Non-linguistic and linguistic conflict resolution were both involved in the experimental design of the current dissertation. These stimuli are addressed below.

3.4.1 Non-linguistic stimuli

The flanker task was used to elicit participant data on non-linguistic conflict resolution and to elicit the more- or less-engaged cognitive control states. Each stimulus included seven letters, presented in the center of the screen and participants responded according to the center letter. In the design of the original test developed by Eriksen and Eriksen (1974), the letters H and K were assigned to the right key response, and the letters S and C were assigned to the left key response. In the interleaved task in this dissertation, only S and H were used. In the independent flanker task in Experiment 2, all four letters were used. The key response for S (and C) was the “f” key and the key response for H (and K) was the “j” key. This increased conflict by disassociating the
letter on screen from the key response while allowing the participant to use the same keys for linguistic and nonlinguistic response selection.

This critical aspect of the task is that it asks participants to respond only to the central letter by pressing one of two keys, despite any flanking letters that conflict with the center arrow. For example, while the correct key response for “SSSHSSS” is “j”, the correct key response for “SSSSSSS” is “f”. In addition to these incongruent and congruent conditions, a third condition, a single letter, was balanced across the experiment. The single letter stimuli were never associated to a critical sentence. The critical comparison is between the congruent and incongruent stimuli. The number of “f” and “j” responses and the number of incongruent and congruent stimuli were balanced across the experiment. Each stimulus was preceded by a fixation cross on screen for 1000 milliseconds (ms). After the fixation cross, participants have up to 2000 ms to respond to the flanker stimulus. Failure to respond within that time limit was scored as an incorrect response.

3.4.2 Linguistic stimuli

The critical linguistic stimuli included fronted adverbial clauses, adapted from Jegerski (2012). These stimuli modulated transitivity of the verb in a preposed adverbial clause, as can be seen in (8) below (with each regional division indicated by a backslash).

(8) a. **Transitive:** Cuando el escultor \ acabó \ la obra \ tenía tres metros \ de altura.

b. **Intransitive:** Cuando el escultor \ volvió \ la obra \ tenía tres metros \ de altura.

c. Region: 1 \ 2 \ 3 \ 4 \ 5

   “When the sculptor \ {finished/came back} \ the piece \ was ten feet \ in height.”
The stimuli are listed in the Appendix A with any alterations from the originals indicated. Any changes to the stimuli were made to emphasize the intransitivity of the critical verb (e.g., empezar changed to llegar) or to include vocabulary more familiar to US classroom learners (e.g. concursante changed to participante). The regions of interest were the following: (1) the postverbal NP in Region 3, where Jegerski (2012) found slower reading times following the intransitive verb due to the mismatch between Late Closure parsing and intransitive verbs’ subcategorization matrices; (2) the matrix verb in Region 4, where Jegerski (2012) found descriptively slower reading times for the transitive condition, likely reflective of pro referent-assignment processes; and (3) Region 5, the spillover region.

3.5 Materials

3.5.1 Experimental task

The linguistic stimuli were presented via a self-paced reading task, using the moving window paradigm. Before each sentence was presented, a fixation cross lasting 1000 ms reset the participant’s gaze on the center of the screen to prepare the participant for the trial. After the 1000 fixation period, the sentence appeared on screen, each stimulus consisting of a single sentence on the screen with each letter of each word masked by a hyphen. The sentences were presented on screen in their entirety, and as participants read, one word would be uncovered at a time, according to their own pace, following the moving window paradigm. After the pilot study, the width of the frame within which PsychoPy presented the stimuli was widened so that all sentences appeared on one line of text, thus avoiding any latency effects associated with participants reverting their gaze to the left edge of the screen to continue reading. Participants
uncovered each word by clicking the space bar. As one word was uncovered, the previous word was re-covered by hyphens. The latency of reading times of each word was recorded by PsychoPy (Peirce et al., 2019). After each sentence, participants answered a true/false comprehension question (CQ), for which the key responses were “f” (true) and “j” (false), the same key responses used in the flanker stimuli.

The experimental task included four lists created using a Latin Square design. Jegerski’s original design included 20 sentences with an transitive and intransitive condition for each sentence, and the study compared RTs of these sentences in each condition. Therefore, the study necessitated two lists of twenty sentences, each list containing ten of the sentences in their transitive condition and the other ten in the intransitive condition. The design of the current dissertation departed from Jegerski’s (2012) original design because of the introduction of the variability of the preceding trial’s congruency. In this dissertation, each participant saw both versions of each sentence, transitive and intransitive, always in separate presentational blocks, and both versions of a sentence were always presented to a given participant following the same type of flanker stimulus (i.e., congruent or incongruent). The four lists balanced the order of presentational blocks (Lists 1 and 2) and the congruency of the flanker that preceded each item (Lists A and B), resulting in Lists A1, A2, B1, and B2.

The forty stimuli from Jegerski (2012) were presented among 80 distractor sentences. These sentences were of varying structures and aimed to draw the participant’s attention away from the subject-object ambiguity. These sentences are included in

---

7 The sentential distractor items included preposed adverbial phrases that included a transitive verb with an explicit complement, sentences with long-distance gender or number disagreement, and sequential prepositional phrases with high-low attachment ambiguity.
Appendix A. Each sentence was preceded by a pseudorandomized flanker stimulus. In addition, thirty unpaired distractor flankers were introduced and were presented randomly (15 during the first half and 15 during the second) to disrupt participants’ expected patterns, adapted from Hsu & Novick (2016).s

Frequency of the words in the linguistic stimuli and distractors was assessed using the Corpus del Español’s WordAndPhrase tool, which compiles 20 million words from texts from the twentieth century and 2 billion words from texts published online since 2014. Participants in Pilot Study 2 completed a vocabulary training, addressed below, for any words that were not in the 3000 most frequent words in this list or that did not appear in their coursework. This training, which will also be provided for participants in studies 2 and 3, is addressed below.

Comprehension Questions (CQs)

True-False CQs were created for each of the sentences, including distractors. These CQs were used to assess and ensure comprehension throughout the study in order to prevent participants from proceeding through the regions of the sentences without actually reading them. This is a common method to ensure sustained attention during self-paced reading studies (e.g. Jegerski, 2012; see Jegerski, 2014). The sentences were presented in the center of the screen, with “Verdadero” (“true”) and “Falso” (“false”) respectively located half-way to the lower left- and right-hand corners of the screen. Participants responded by pressing the “f” or “j” key, which are the keys where their index fingers would rest when using their thumbs to press the space bar during the self-

s The flanker distractor items were chosen such that each of the six conditions (right vs. left; congruent vs. incongruent vs. single) was balanced across all trials.
paced reading task, and which were also the keys used to indicate a leftward or rightward facing arrow. In other words, participants did not have to shift their hands between the three tasks.

3.5.2 Non-experimental materials

Vocabulary Training

As mentioned above, frequency of the words in the linguistic stimuli and distractors was assessed using the Corpus del Español’s WordAndPhrase tool. Words in Jegerski’s (2012) stimuli that were not among the 3000 most frequent words on the corpus’ list were included in a decontextualized vocabulary test given to Beginner 1 students to assess how well these students would recognize the words, assuming that students who completed Beginner 1 would have a more limited vocabulary than the students who participated in this study. These students were asked to translate each word and rate how well they recognized the word on a scale of 1-4. If words in this test met three of the following four criteria, they were included in the critical stimuli: (i) the word is covered in the vocabulary lists of Beginner 1 or 2 (i.e., Intermediate I students are expected to know this word); (ii) the word elicited an exact translation accuracy of 85% or higher; (iii) the word received a correct recognition rate among all raters of at least 3; and (iv) the word is a cognate with no more than two graphemic changes due to Spanish phonotactics and spelling norms (e.g. violín, escultor). The seventeen words that did not meet these criteria were included in the vocabulary training, along with an additional 25 words. These 25 words were selected pseudo-scientifically, based on the investigator’s intuition about vocabulary as an instructor of Spanish. However, they did not meet the criteria established above or they were among the 3000 most frequent words in the corpus.
Pilot Study 2 participants were all required to complete a vocabulary training on PsychoPy before completing the study to establish a baseline vocabulary. This training presented 42 vocabulary items used in the sentences, and 17 of which appeared in the critical stimuli. The 42 words were divided into two blocks. The words in each set were presented one at a time with a photo that represented the word and a translation in English. After 1000 ms, participants were invited to move to the next word if and when they were ready. After each set of 21 words was presented, participants had an opportunity to review the words again or take a quiz to continue. The quiz presented four of the photos from the set and asked participants to identify which photo represented the word shown on the screen. A minimum score of 85% was required on the quiz from each block in order to continue to the next phase of the study.

Language Background Questionnaire

All participants completed a language background questionnaire (LBQ), edited from the LEAP-Q (Marian, Blumenfeld & Kaushanskaya, 2007). Information regarding languages spoken was collected for all participants, but the participants in Experiment 1 completed detailed questions for Spanish, Catalan and English; participants in Experiment 2 completed detailed questions for English and Spanish. The questions were presented via computer.

Elicited Imitation Task

The Elicited Imitation Task (EIT) was used as the primary proficiency measure (see Appendix B). The EIT was developed as a means of measuring global proficiency for research purposes (Ortega, 2000; Ortega, Iwashita, Norris, & Rabie, 2002). It has been validated in Spanish as an equally reliable measure of global proficiency as the
Simulated Oral Proficiency Interview (SOPI), a global proficiency measure based on the American Council of the Teaching of Foreign Languages proficiency guidelines (Bowden, 2007). In this task, participants are recorded listening to and repeating a set of 30 pre-recorded sentences. These sentences become progressively more complex structurally. The recordings are then scored on a scale of 0-4 according to the accuracy of each repetition, and the total score for each participant falls between 0 and 120 (see Appendix C).

**Automated Operation Span Task**

Executive working memory (updating function) was measured using the Operation Span (OSpan) task (Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005), which was administered on the computer via Inquisit Web 5.0.10. The automated version of this task was used. In the automated OSpan, participants must remember series of letters, ranging from 3-7 per series, that are interleaved with mathematical equations. Each series begins with a math equation with two operations (one multiplication/division and one addition/subtraction) that the participant solves mentally, they click to the next screen, which proposes an answer and asks the participant to respond if the answer is correct or not. Following this, a letter that the participant must recall later appears on screen. At the end of a series, a 4x4 letter matrix is provided and the participant must select the letters of the series in the order in which they were presented. 15 series are presented in the task, 3 repetitions of each of the five series lengths.

Before beginning, participants complete practice rounds of the letter recall only, of the math problems only, and of the interleaved tasks. The automated OSpan uses the math practice block to calculate each participant’s mean latency in solving the math problems.
This value plus 2 SDs is then used as the math timeout for subsequent math problems.

The letters appear on screen for 800 ms. See Unsworth, Heitz, Schrock, & Engle, 2005, for the advantages of the automated OSPAN.

3.6 Procedure

3.6.1 Procedure: Experiment 1

Early bilinguals in Experiment 1 were recruited from a group of students at the Universitat Pompeu Fabra who participated in a Spanish conversation program with Georgetown’s summer abroad program in Barcelona. They were paid 15€ for completing the study. The entirety of the study was conducted in Spanish. Before beginning the study, the researcher explained the study generally in a quiet room, obtained informed consent and assigned participants to one of four groups. These groups determined the list and presentational order of blocks for each participant. Following this, participants completed three training phases to prepare for the experimental phase: non-interleaved flankers, non-interleaved self-paced reading sentences, and interleaved flanker-sentence reading pairs. Participants needed to achieve minimally 80% accuracy on each of these training phases in order to move on to the experimental phase. Once they completed the training phases, they had an opportunity to rest and ask any questions before beginning their two experimental blocks, each of which contained 60 flanker-sentence pairs, 15 unpaired flankers, and a brief intra-block break. Following the first block, participants had another opportunity to rest before completing the second block, which had the same structure. In addition, if the participant answered three consecutive flankers or three consecutive CQs incorrectly, they were forced to take a 10-second break, during which they were reminded that both speed and accuracy were important to finish the study in a
timely fashion. Upon completion of the interleaved self-paced reading task, participants completed the automated OSpan and then the LBQ.

3.6.2 Procedure: Experiment 2

Emergent bilingual language learners in Experiment 2 were recruited from Georgetown University’s 5-week summer abroad program in Barcelona Spain, which requires the equivalent of at least six courses of university courses; from upper-level Spanish classes at Georgetown University; or from related graduate programs, such as the Spanish doctoral programs and the Master’s in Latin American Studies. Participants enrolled in Spanish courses were rewarded two points on a mid-semester test for participating in the study, and graduate student participants who were not enrolled in a course were paid $30. Participants completed the study in a large computer lab with up to two other participants. These participants all wore headphones and were sat in opposite corners of this large lab facing away from each other to minimize interference. Before beginning the study, the researcher explained the study broadly, obtained informed consent and assigned participants to one of the four groups to determine list and block order. Participants first completed the vocabulary training, which required minimally 85% accuracy. They then completed three training phases to prepare for the experimental phase: non-interleaved flankers, non-interleaved sentences in self-paced reading, and interleaved flanker-sentence reading pairs. The explanation of the study, informed consent, instructions for the flanker, and both instructions and sentences in the non-interleaved self-paced reading training were conducted in English. The instructions to the interleaved training phase were also presented in English, but then participants were instructed that the actual stimuli and CQs would be presented in Spanish. This was done
to facilitate the transition to Spanish and initiate the so-called “monolingual mode” in Spanish. All computerized instructions and stimuli following this point were presented in Spanish. In order to move onto the experimental phase, participants needed to achieve 80% accuracy or higher on each training phase.

Once they completed the training phases, they had an opportunity to rest and ask any questions before beginning their two experimental blocks (identical to the blocks in Experiment 1). Like in Experiment 1, participants who answered incorrectly on three consecutive flankers or CQs were forced to take a short break, of at least 10 seconds. Upon completion of the experiment, participants completed the LBQ. After the LBQ, they were escorted from the laboratory to complete the EIT in another space to control the noise level in the laboratory.

Two weeks after participating, participants returned to the lab to complete a self-paced reading task without interleaving flanker stimuli, a flanker task without linguistic stimuli, and the Automated Operation Span. Unfortunately, word-by-word reading times for the self-paced reading task were not recorded due to an undetermined glitch in the PsychoPy script, so separate analyses planned for the independent self-paced reading task were not conducted. However, the flanker and Operation Span tasks were included in Experiment 2.

3.7 Coding and Scoring

3.7.1 Self-paced reading

In the code prepared for the current dissertation, PsychoPy latency data collection occurs at the word level, so regional-level latency was calculated post-hoc for each participant’s 40 critical stimuli by combining the latency of each word of the region.
In order to account for data that was likely affected by external factors, data was trimmed at the both the group and the subject level. At the group level, the maximum was set at 4000 ms (Jegerski, 2018) and the minimum to 100 ms (Luce, 1986). The group-level maximum was reduced following the pilot study. However, subject-level cutoffs superseded all group-level cutoffs. These subject-level cutoffs were established as two standard deviations from each participant’s mean reading time. In normally distributed data sets, 95% of the data falls within two SDs from the mean. Outlying reading-time values that fell beyond the group-level or subject-level cutoffs were trimmed to the value closer to the mean, and in all cases, the subject-level cutoff was closer to the mean. The percent of data replaced was within the normal limits proposed by Jegerski (2014). In early bilingual data, 4.21% of the early bilingual data (269 replacements, 25 at the subject-level minimum and 244 at the subject-level maximum) and 4.36% of the emergent bilingual data (323 replacements, all at the subject-level maximum) was replaced. For more information on self-paced reading trimming, see Jegerski (2014).

After removing data that did not meet the above criteria, the reading time data was log-transformed to account for positive skew typical of reaction time data (Luce, 1986).

3.7.2 Comprehension check

The inclusion of reading times for a given sentence was contingent on the sentence’s corresponding CQ being understood. All data with incorrect CQs were removed from analyses. If less than 60% of a participant’s data was included in the analyses, the participant was not included in final analyses. All early bilinguals in Experiment 1 and all but one emergent bilingual in Experiment 2 obtained the 60% cutoff. Comprehension scores are addressed in more detail in Chapter 4.
3.7.3 Operation span

Participants OSpan scores were automatically generated by the automated task. The conventional absolute score was used, which is the sum of all letters recalled in all perfectly recalled sets. For example, if a participant only recalled all letters in their order for the three sets of length = 3, their score would be 9, but if she remembered all letters in their order for the six sets with lengths of 3 or 4 letters, her score would be 21 (3x3+3x4+0x5+0x6+0x7).

3.7.4 Elicited imitation task scoring

Proficiency was operationalized as the score on the elicited imitation task (EIT). The EIT includes 30 sentences of increasing length and complexity, each of which is scored from 0 to 4, allowing for proficiency scores between 0 and 120 (Ortega, 2000; see Appendix C for scoring protocol). The same research assistant who scored the EITs for the pilot study was recruited to score a subset of the full study’s EITs. The researcher and research assistant separately coded 17.5% of the recordings (30 repetitions from 7 of the 40 participants), for a total of 210 comparisons. Cohen’s Kappa coefficient was calculated to assess interrater reliability. Following recommendations from Cohen (1960) and Lowry (2019), linear weighting was used, as scores were ordinal (increasing scores represent increasingly more successful production of the target), and the reported kappa is observed as a proportion of the maximum possible kappa given the marginal frequencies. This is recommended in cases when raters are not limited to a certain number of ratings per category. Concordance was high between the two raters, $\kappa_w = .9775$, SE = .0322, 95% CI [.6703, .7965], from the observed statistic of $\kappa = .7334$. This is considered excellent by Fleiss (1981) and almost perfect agreement by Landis and Koch (1977). After confirming
that the researcher was consistent with the protocol, the remainder of the data was coded by the researcher alone.

3.8 Statistical Analyses

3.8.1 Linear mixed-effects models for linguistic and flanker stimuli

Linear mixed-effects models (LMEMs) were conducted for the main analyses in this dissertation. In all LMEMs, the maximally converging random effect structure was used. Unless otherwise specified, for all analyses of the critical regions, this effect structure included random intercepts by subject and by item, and by-subject random slopes for transitivity. Such models are common in item response theory, where subject and item factors are fully crossed (Doran, Bates, Bliese, and Dowling 2007). The alpha level was set to .05 for all statistical analyses, unless post-hoc tests required corrections for multiple comparisons. All data was analyzed using R, version 1.1.463 (R Core Team, 2013), with the lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017), MuMin (Bartoń, 2019), emmeans (Lenth, Singmann, Love, Buerkner, & Herve, 2020) and psych (Revelle, 2018) packages. Some figures were created using ggplot2 (Wickham, 2016) and ggeffects (Lüdecke, 2018) packages.

Before analyzing the linguistic data, flanker responses from the interleaved task were also submitted to a LMEM, with reaction time as the dependent variable, incongruency as a fixed effect and a random effect structure with by-subject adjustments to the intercept and by-subject random slopes for incongruency. In the case of the emergent bilinguals, for whom both interleaved and independent flanker trials were recorded, a paired samples t-test compared their incongruent and congruent reaction times from the independent flanker task.
3.8.2 Effect sizes

To calculate the effect size of factors, (i.e., transitivity and incongruency), Cohen’s $d$ was calculated for repeated measures, using the formula:

$$d = \frac{|m_1 - m_2|}{\sqrt{(s_1^2 + s_2^2 - 2rs_1s_2)}}$$

where $m$ is the mean in a given condition, $s$ the standard deviation of a given condition, and $r$ is Pearson’s correlational coefficient which corrects for interdependent means due to the within-subject nature of the design. Avery and Marsden (2019) performed a meta-analysis of L2 SPR studies, which highlights the importance of interpreting these effect sizes within a different framework from other SLA effects (e.g., the framework established by Plonsky & Oswald, 2014). Their framework suggests a reliable sensitivity to morphosyntax among L2 learners of $d = 0.20$, a reliable sensitivity to anomaly detection of $d = 0.19$, and larger effect sizes ($d = .41$) in anomaly detection in conditions where the L1 and L2 were similar (such as this dissertation). For native speakers, they found larger effect sizes: $d = .28$ for sensitivity to morphosyntax and $d = .41$ for anomaly detection. Because there is historically less usage of Cohen’s $d$ effect size in L2 SPR studies, caution is still advised when interpreting these statistics.

To calculate the effect size of proficiency in Experiment 2, a continuous variable, Cohen’s $f^2$ was calculated, using the formula:

$$f^2 = \frac{(R_{2AB} - R_{2A})}{(1 - R_{2AB})}$$
where $R^2$ is the squared multiple correlation, representing the variance accounted for in a model. A is the set of all fixed effects in a model except the variable of interest, and B is the variable of interest (Cohen, 1988). In other words, Cohen’s $f^2$ calculates the magnitude of the effect of a given variable using the variance from the complete model and the variance from a model excluding that variable. Cohen (1988) recommends interpreting effect size as small when $f^2 = .02$, as medium when $f^2 = .15$, and as large when $f^2 = .35$, so these guidelines are used to interpret $f^2$ throughout this dissertation.

3.9 Pilot Studies and Resulting Changes

Two small pilot studies were conducted for Experiments 1 and 2 to obtain preliminary evidence that would support the viability and validity of the experimental design. Descriptive statistics and basic analyses were run following the pilot studies. These descriptive statistics are described below, followed by a summary of the changes made based on the pilot studies.

3.9.1 Pilot study materials and procedures

The materials for the pilot study were administered, following the procedures outlined above, with the following exceptions. First, while participants in the pilot study saw all twenty transitive sentences, they only saw the intransitive version of ten sentences, all preceded by incongruent flankers, in order to increase the percentage of distractors. Secondly, participants in the pilot study completed the arrow flanker task, while participants in Experiments 1 and 2 completed the flanker task using letters as described above. Thirdly, participants in Pilot Study 2 did not complete an independent flanker trial separate from the interleaved flanker trials. Finally, the Automated Operation Span task included above was introduced following the pilot studies.
The resulting data was coded, trimmed, and scored as outlined above, with the following major exception. Due to limited data in the pilot study and to determine the direction of reading time effects, residual RTs were used for comparisons (RRTs) instead of log-transformed RTs. RRTs are a form of normalized reading time data that shows if a given RT is faster or slower than the expected reading time, based on participants’ mean reading times on a wider array of stimuli of the same length. Therefore, positive RRTs presented in the figures that follow represent RTs slower than expected, while negative RRTs represent RTs faster than expected.

3.9.2 Pilot study 1

Participants

Participants in Pilot Study 1 were five early bilinguals of Spanish and Catalan (3 female, 2 male). Four of the participants (Participants 2-5) were simultaneous Spanish-Catalan bilinguals, who began learning Spanish and Catalan before age 2. The fifth participant (Participant 1) was a sequential bilingual, who began speaking Catalan at the age of 7 when her family moved from Colombia. Due to low numbers in the pilot study, the simultaneous and sequential bilingual groups were combined. The participants were recruited from the Universitat Pompeu Fabra in Barcelona, Spain. Participants had completed between 1 and 4 years of university study. The average age was 20.6 (SD = 1.34).

The participants were asked to self-report their abilities in Spanish and Catalan on a scale of 0-5, where 0 was defined as “poor” and 5 as “near-native/native”. They ranked their abilities in the following five categories: reading, writing, speaking, listening and vocabulary knowledge. The four simultaneous bilinguals self-reported each of these
abilities in Spanish and Catalan at the near-native/native level ($M = 5, SD = 0$); the sequential bilingual reported her Catalan at “near-native or native” for reading, and at “excellent” for the other four abilities. However, she also rated her Spanish listening and speaking abilities as “excellent”, and only ranked her reading, writing and vocabulary knowledge at the “near-native/native” level.

Participants were also experienced foreign language learners. The four simultaneous bilinguals reported learning English before age 13 ($M = 7.5, SD = 3.27$) but did not have extended experience abroad (1 month or less), and reported using their non-native languages less than 10% of the time. The sequential bilingual reported learning English between 13 and 17. Three participants (2, 3, 4) also began learning a fourth language between the ages of 5 and 13 (Japanese, German, and French, respectively), but considered English to be their most proficient foreign language.

Results

After removing outliers as described above, the data was then cleaned with regards to comprehension checks, with a threshold set at 60% accuracy on the CQs (Marijuan, 2015). All early bilinguals met this threshold. Early bilinguals responded correctly on 93.33% of the CQs (Range: 90.00% to 96.67%). In terms of transitivity, early bilinguals responded correctly to CQs for 92.00% of transitive sentences and for 96.00% of intransitive sentences. Sentence-flanker pairs with incorrect CQs that were removed from analyses accounted for 6.67% of early bilingual data (10 stimuli pairs distributed across the 5 participants), leaving a total of 140 flanker-sentence stimuli pairs.

Residual reaction times in the three conditions included in the pilot study were only considered descriptively using analyses of variance in SPSS (see Figure 1). Residual
reaction times were not used in the main analyses; rather, log-transformed reading times were used. Rather than analyses of variance, the main analyses used linear mixed-effects models. Therefore, the graphs presented below are merely reflections of the reading time effects observed in the pilot studies, which had few participants. No statistical importance should be interpreted from the figures 1-4. The only describable difference between the congruent-Transitive (c-T) and incongruent-Transitive (i-T) conditions were found at Region 3, where the i-T RTs were faster. As a reminder, these were the exact same linguistic stimuli. The only difference between the conditions was that half of the participants saw a given item after a congruent flanker and the other half in the incongruent condition. Only three conditions were included in the pilot study: c-T, i-T, and incongruent-Intransitive (i-I). The fourth permutation (c-I) was not included, but deemed necessary following the pilot study. The changes motivated by Pilot Study 1 are addressed below.

Figure 1. Early bilingual reaction time data by condition in Pilot Study 1 using ANOVAs and residual reading time
3.9.3 Pilot study 2

Participants

Participants in Pilot Study 2 were 18 classroom learners of Spanish (7 female, 11 male). They were recruited from one of three Spanish language courses at Georgetown University. Three participants were recruited from Intermediate I, twelve from Advanced I, and three from Oral Review, a course that follows Advanced II for these students. Because of low accrual numbers in two of these levels and the overlapping scores on the secondary proficiency measure (elicited imitation task, EIT, addressed below), which was likely due to self-selection bias, the EIT was used to separate participants into proficiency groups. Three early bilinguals (2 Korean-English and 1 English-Spanish bilinguals) were removed from the sample after reviewing their language background questionnaire (LBQ). The remaining fifteen participants all identified themselves as native speakers of English only, all born in the United States, and with a mean age of 18.6 years (SD = .74). Eight participants began learning Spanish between the ages of 10 and 13, and seven between 13 and 18. Spanish was the first classroom-learned language for 13 of these participants, while the other two learned Hebrew or French after age 13.

A research assistant scored the recordings of the EITs in the pilot study. Before scoring the participants from Pilot Study 2, the researcher and research assistant independently scored two EIT recordings collected for a previous study. Cohen’s kappa (κ) was run to determine the extent of agreement between the two coders. Concordance was moderate between the two raters, κw = .7024 (95% CI, .196 to .521), SE = .0622. Disagreements were discussed one by one until a consensus was reached. After this, the two raters individually scored recordings from the first two participants from Pilot Study
2. Cohen’s Kappa was calculated again and a much higher rate of agreement was found, $\kappa_w = .9472$ (95% CI, .8091 to .9463), SE = .035. This is considered excellent by Fleiss (1981) and almost perfect agreement by Landis and Koch (1977). Following the second round of independent scoring, the research assistant scored the remainder of participants’ recordings from Experiment 2.

Results

As in Pilot Study 1, outliers were removed and data was trimmed according to the criteria above. One emergent bilingual in the pilot study fell below the 60% comprehension threshold, leaving 14 emergent bilinguals in subsequent analyses. The remaining participants responded correctly to 73.81% of the questions (Range: 60.00% to 93.33%). In terms of transitivity, emergent bilinguals responded correctly to CQs for 71.79% of transitive sentences and for 77.86% of intransitive sentences. Data removed due to incorrect CQs accounted for the removal of 26.19% of emergent bilingual data (110 stimuli pairs across the 14 participants), leaving 310 flanker-sentence pairs for the emergent bilinguals. As will be discussed below, these comprehension rates were lower than the desired comprehension level, motivating a change in the recruited population.

As in Pilot Study 1, RRTs in the three conditions that were included in the pilot study were only assessed descriptively (see Figure 2). Unlike in Pilot Study 1, the i-T and c-T conditions in Pilot Study 2 differed more conspicuously at several regions, most of which was interpreted as noise from the small sample (see Figure 2). However, there was a more considerable difference at Region 4. This was taken as a warning that the processing effects observed in early bilinguals may manifest later in the sentence for L2 learners.
Figure 2. Emergent bilingual reaction time data by condition in Pilot Study 2 using ANOVAs and residual reading time

A post-hoc division of the group according to EIT score (using the median EIT score as the dividing point) revealed a difference that aligns with two of the predicted patterns of results. In the lower proficiency group, means were slightly slower in the c-T condition, but the difference between Condition 1 and 2 was generally consistent across all regions except Region 1 (see Figure 3).

Meanwhile, in the higher proficiency group, we also can observe similar means in both the c-T and i-T conditions, regardless of cognitive control engagement, except at Regions 4 and 5 (see Figure 4). These differences, signaling the possibility of a differential role of proficiency, motivated the full study to be carried out, with the critical changes addressed below.
Figure 3. Low-proficiency participants’ mean reaction times by condition in Pilot Study 2 using ANOVAs and residual reading time

Figure 4. High-proficiency participants’ mean reaction times by condition in Pilot Study 2 using ANOVAs and residual reading time
3.9.4 Changes made following the pilot studies

Listed below are the changes to the experimental design motivated by the pilot study.

(1) The pilot study used analyses of variance (ANOVAs) to compare conditions, but after the pilot study, it was determined that the analyses used should be linear mixed-effects models.

(2) The 6000 group-level cutoff originally established only removed .2% of the near-native data, so the group-level maximum was reduced to 4000 ms for the full study. Even so, the subject-level maximum, two SDs above the participants’ means, was a more productive cutoff, accounting for all replacements made.

(3) The participants in the full studies completed the Automated Operation Span as a measure of working memory capacity. The participants in Experiment 2 (but not Experiment 1) also completed an independent flanker task that did not include interleaved self-paced reading stimuli.

(4) The inclusion of four conditions: c-T, i-T, i-I, and importantly, the fourth condition to serve as the baseline level, c-I. This final condition was determined to be critical for the full study to make any real conclusions about cognitive-control engagement.

(5) The rejection of changes made in the pilot studies to the intransitive sentences. Changes made at Region 5 to avoid task effects obscured differences between the reading times, despite precautions about balancing for syllables, word length, and word frequency. The intransitive and transitive versions were balanced across the lists such that, for example, if a given transitive version was seen in Block 1 after
the intra-block break, its intransitive counterpart was seen in Block 2 after the intra-block break.

(6) Given the low comprehension scores obtained by participants in Pilot Study 2, proficiency was re-operationalized using EIT score as the main determiner of proficiency. Number of courses taken was only used as a criterion to recruit participants. This removed the upper limit of the proficiency range. Using a continuous variable rather than a factor also bolsters the LME analyses.
CHAPTER 4: RESULTS

4.1 Overview

In this chapter, the results of all confirmatory hypothesis testing from both experiments are presented, as well as the exploratory questions corresponding to Experiment 1 and 2. All results for Experiment 1 are presented first, followed by a brief summary of the results. Results of Experiment 2 are then presented, followed by another brief summary. Discussions of the results and limitations of the experiments are presented in Chapter 5, with future directions and a conclusion.

4.2 Experiment 1

The first goal of Experiment 1 was to assess whether early Catalan-Spanish bilinguals reveal the same reading time effects in a cross-task experiment that have been found for other native Spanish speakers in self-paced reading tasks (RQ1). The second goal was to assess whether cognitive control engagement in this cross-task paradigm facilitated these processing effects (RQ2). These two research questions, though separate, are intimately related by the research design. An additional goal of Experiment 1 was to explore whether cognitive individual differences accounted for any linguistic processing effects in addition to the experimental manipulation of cognitive control (RQ3). In what follows, I describe the participant sample for this experiment, present the descriptive and inferential statistics along with the results of the exploratory analyses, and offer a summary of the main findings.

4.2.1 Participants

Participants of Experiment 1 were 29 native speakers of Spanish and Catalan, born and raised in Barcelona, Spain (26 female). Two participants were also native
speakers of other languages. Participants’ additional language background information is summarized in Table 1.

Table 1. Participant background information: Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>M (SD)</th>
<th>Mdn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>21.34 (1.56)</td>
<td>21</td>
</tr>
<tr>
<td>Number of native languages</td>
<td>2.07 (.26)</td>
<td>2</td>
</tr>
<tr>
<td>Number of foreign languages</td>
<td>2.41 (.98)</td>
<td>2</td>
</tr>
<tr>
<td>Age of exposure to Spanish</td>
<td>.25 (.91)</td>
<td>0</td>
</tr>
<tr>
<td>Age of exposure to Catalan</td>
<td>.31 (.93)</td>
<td>0</td>
</tr>
<tr>
<td>Age of exposure to first foreign language</td>
<td>5.43 (1.79)</td>
<td>6</td>
</tr>
<tr>
<td>Self-rated Spanish proficiency</td>
<td>Speaking: 9.93 (.22)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Understanding: 9.83 (.76)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Reading: 9.83 (.76)</td>
<td>10</td>
</tr>
<tr>
<td>Self-rated Catalan proficiency</td>
<td>Speaking: 9.74 (.61)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Understanding: 9.88 (.46)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Reading: 9.88 (.46)</td>
<td>10</td>
</tr>
<tr>
<td>Self-rated English proficiency</td>
<td>Speaking: 7.95 (1.06)</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Understanding: 8.48 (.97)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Reading: 8.64 (.94)</td>
<td>9</td>
</tr>
</tbody>
</table>

Note. Participants’ reported native languages also included Bulgarian (n = 1) and Arabic (n = 1), but both of these participants reported minimal active exposure to these third native languages and reported dominance in Spanish and Catalan from birth. Participants’ reported foreign languages included English (n = 29), French (n = 20), German (n = 8), Italian (n = 6), Chinese (n = 3), Arabic (n = 3), Korean (n = 1), and Galician (n = 1).

Before data was analyzed, raw reading times were transformed as outlined in Chapter 3, including group- and subject-level maximums and minimums. As mentioned in Chapter 3, only the subject-level bounds applied, accounting for 269 replacements (4.21% of the data), 25 at the subject-level minimum and 244 at the subject-level maximum. The values were then log-transformed to account for the positive skew typical of reaction time data (Luce, 1986). As addressed in Chapter 3, data inclusion was also contingent on the participant having a comprehension score of at least 60%, which all native speakers obtained. Early bilinguals’ mean accuracy for all items, including filler sentences, was 92.8% (SD = 3.47%), and for the critical items, mean accuracy was 89.7% (SD = 6.20%) for the critical items. Data corresponding to incorrect comprehension questions were removed from analyses. This included over 40% of the data.
corresponding to the intransitive version of one of the items, so to maintain the experiment’s balance, the item was removed from further analyses. The resulting comprehension scores were 91.37% for the transitive condition and 89.09% for the intransitive condition.

Upon examination of the data, three participants were exposed to an incorrect counterbalancing scheme. As outlined in Chapter 3, the intended counterbalancing scheme exposed any given participant to each item in both their intransitive (i) and transitive (t) versions. For a given item S, the participant saw both versions, Si and St, in the same congruency condition (either incongruent for both versions or congruent for both versions). The three participants exposed to the incorrect counterbalancing scheme were exposed to half of the items twice, but in differing congruency conditions. That is, if they saw the intransitive version of Item 1 in Block 1 following a congruent flanker stimulus, they also saw the intransitive version of Item 1 in Block 2, but preceded by an incongruent flanker stimulus. These participants were therefore removed from subsequent analyses, resulting in 26 participants, 19 items, and a total of 905 observations in each of the regions of interest.

Before analyzing the linguistic data, an analysis was conducted of the native bilingual data to evaluate if the interleaved flanker responses yielded a congruency effect. A linear mixed effects model (LMEM) was conducted with reaction time as the dependent variable, incongruency as a fixed effect and a random effect structure with by-subject adjustments to the intercept and by-subject random slopes for incongruency. This revealed a congruency effect (CE) in the predicted direction (incongruent: $M = 696$ ms, $SD = 224$ ms; congruent: $M = 645$ ms, $SD = 211$ ms), with statistically slower responses for
incongruent stimuli, estimate = 53.26, SE = 13.85, t = 3.847, p < .001 [log-transformed response times: estimate = .027, SE = .007, t = 3.703, p < .001]. This confirms that the flanker stimuli elicited a congruency effect in the canonical direction, despite being interleaved among self-paced reading stimuli, and so the main analyses were run to determine how native bilinguals responded to the linguistic conflict in a task designed to trigger conflict adaptation.

4.2.2 Results: Research questions 1 and 2

Descriptive statistics are presented for each region in Table 2. The regions of interest are Regions 3, 4, and 5. Mean values are also presented in Figure 5.

<p>| Table 2. Descriptive statistics of early bilingual raw reaction time means by condition |</p>
<table>
<thead>
<tr>
<th>Region</th>
<th>Sentence Type</th>
<th>All flankers M (SD)</th>
<th>Congruent M (SD)</th>
<th>Incongruent M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>Transitive</td>
<td>887 ms (218 ms)</td>
<td>875 ms (231 ms)</td>
<td>899 ms (218 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td>Intransitive</td>
<td>894 ms (257 ms)</td>
<td>893 ms (278 ms)</td>
<td>895 ms (251 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td></td>
<td>891 ms (231)</td>
<td>884 ms (244 ms)</td>
<td>897 ms (226 ms)</td>
</tr>
<tr>
<td>Region 2</td>
<td>Transitive</td>
<td>436 ms (177 ms)</td>
<td>427 ms (175 ms)</td>
<td>445 ms (195 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td>Intransitive</td>
<td>409 ms (128 ms)</td>
<td>403 ms (132 ms)</td>
<td>416 ms (171 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td></td>
<td>423 ms (142 ms)</td>
<td>415 ms (140 ms)</td>
<td>430 ms (169 ms)</td>
</tr>
<tr>
<td>Region 3</td>
<td>Transitive</td>
<td>733 ms (188 ms)</td>
<td>729 ms (197 ms)</td>
<td>737 ms (191 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td>Intransitive</td>
<td>769 ms (248 ms)</td>
<td>790 ms (273 ms)</td>
<td>748 ms (237 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td></td>
<td>750 ms (215 ms)</td>
<td>758 ms (231 ms)</td>
<td>743 ms (207 ms)</td>
</tr>
<tr>
<td>Region 4</td>
<td>Transitive</td>
<td>594 ms (160 ms)</td>
<td>603 ms (181 ms)</td>
<td>584 ms (181 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td>Intransitive</td>
<td>595 ms (133 ms)</td>
<td>591 ms (119 ms)</td>
<td>598 ms (158 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td></td>
<td>594 ms (142 ms)</td>
<td>597 ms (143 ms)</td>
<td>591 ms (147 ms)</td>
</tr>
<tr>
<td>Region 5</td>
<td>Transitive</td>
<td>1020 ms (255 ms)</td>
<td>1032 ms (275 ms)</td>
<td>1006 ms (247 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td>Intransitive</td>
<td>1102 ms (295 ms)</td>
<td>1092 ms (271 ms)</td>
<td>1112 ms (331 ms)</td>
</tr>
<tr>
<td>All sentences</td>
<td></td>
<td>1060 ms (248 ms)</td>
<td>1062 ms (243 ms)</td>
<td>1058 ms (258 ms)</td>
</tr>
</tbody>
</table>
LMEMs were run for each of the critical regions, with transitivity and incongruency as fixed factors and a random effects structure that included by-participant and by-item random intercepts, as well as by-participant random slopes for the fixed effect of transitivity. The reference level for transitivity was set as the intransitive condition, and the reference level for incongruency was set as the congruent condition. The results of the LMEMs are included in Table 3. Inferential statistics are presented in Appendix D for Regions 1 and 2.
The models revealed that the early bilinguals revealed the expected sensitivity to the postverbal NP, namely slower reading times when the postverbal NP followed the intransitive verb, estimate = -.0209, SE = .0117, t = -2.449, p = .015, with an effect size of $d = .330$, consistent with Avery and Marsden’s (2019) observations of native SPR effect sizes resulting from sensitivity to morphosyntax. Incongruency did not significantly slow reading times, estimate = -.0203, SE = .0119, t = -1.709, p = .088, although the effect size, $d = .365$, suggests that this may be an avenue for further research with greater power. There was a significant interaction between transitivity and incongruency, estimate = .0343, SE = .0166, t = 2.064, p = .039, with an effect size of $d = .184$. This interaction is depicted in Figure 6. Post-hoc pairwise comparisons using R’s `emmeans` package (Lenth et al., 2020) revealed that the differences in reading times were only significant in the congruent condition, estimate = .0287, SE = .0117, t = 2.449, p = .015, but transitive and intransitive reading times did not statistically differ following incongruent flankers, estimate = -.0057, SE = .0119, t = - .476, p = .634. This supports the prediction that a conflictive preceding trial will facilitate the resolution of linguistic conflict. In this case, the linguistic conflict is between the principle of late closure, which leads the parser to incorporate new linguistic
items into the running parsing structure, and the subcategorization matrix of the intransitive verbs, which precludes intransitive verbs from taking complements (i.e., direct objects). The reading time effect caused by this conflict in the intransitive condition was neutralized when cognitive control was preemptively engaged by a preceding incongruent flanker.

The other predicted reading time effect, longer reading times in the transitive condition at the matrix verb (Region 4) or the spillover region (Region 5), was not statistical in the model.

4.2.3 Results: Research question 3

In order to explore whether participants’ cognitive individual differences accounted for the effects of transitivity or incongruency on Region 3 reading times, non-
confirmatory hypothesis testing LMEMs were conducted that included the cognitive ID data collected. Cognitive ID measures are included in Table 4. These variables were the Operation Span score and the magnitude of the congruency effect (CE). As a reminder to the reader, the CE is the difference between response times for all correct incongruent and all correct congruent flanker trials.

Table 4. Descriptive statistics of early bilingual cognitive measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean (SD)</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated Operation Span</td>
<td>29.48 (14.85)</td>
<td>31</td>
</tr>
<tr>
<td>Congruency Effect (from interleaved task)</td>
<td>34 ms (26 ms)</td>
<td>36 ms</td>
</tr>
</tbody>
</table>

Chi-squared tests were conducted to compare the LMEMs with covariates with the original model. The results of the chi-squared tests are presented in Table 5. Neither model differed significantly from the original model and neither of the variables was a significant predictor within these models, while the significance of the other predictors (i.e., transitivity, incongruency, and transitivity*incongruency) did not change. The results are presented in Table 6.

Table 5. Chi-squared statistics for two exploratory models compared to Model1

<table>
<thead>
<tr>
<th>Model</th>
<th>$X^2$</th>
<th>df</th>
<th>$p$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>.003</td>
</tr>
<tr>
<td>ModelOSPAN</td>
<td>.554</td>
<td>1</td>
<td>.457</td>
<td>.011</td>
</tr>
<tr>
<td>ModelCE</td>
<td>.208</td>
<td>1</td>
<td>.648</td>
<td>.007</td>
</tr>
</tbody>
</table>

Because the variables of interest (OSpan or CE) in these exploratory studies were not significant in the covariate models, no further steps were taken to determine if they interact with the variables of Model1.
Table 6. Output from two exploratory linear mixed-effects models compared to Model 1

<table>
<thead>
<tr>
<th>Region 3 Model</th>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Constant Intercept</td>
<td>2.860</td>
<td>.0285</td>
<td>100.423</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>-.0287</td>
<td>.0117</td>
<td>-2.449</td>
<td>.015*</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>-.0203</td>
<td>.0119</td>
<td>-1.709</td>
<td>.088</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>.0343</td>
<td>.0166</td>
<td>2.064</td>
<td>.039*</td>
</tr>
<tr>
<td>Model OSPAN</td>
<td>Constant Intercept</td>
<td>2.895</td>
<td>.0553</td>
<td>52.343</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>-.0287</td>
<td>.0117</td>
<td>-2.450</td>
<td>.015*</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>-.0203</td>
<td>.0119</td>
<td>-1.712</td>
<td>.087</td>
</tr>
<tr>
<td></td>
<td>OSpan Score</td>
<td>-.0011</td>
<td>.0016</td>
<td>-0.723</td>
<td>.477*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>.0343</td>
<td>.0166</td>
<td>2.065</td>
<td>.039*</td>
</tr>
<tr>
<td>Model CE</td>
<td>Constant Intercept</td>
<td>2.895</td>
<td>.0419</td>
<td>69.074</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>-.0287</td>
<td>.0117</td>
<td>-2.451</td>
<td>.014*</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>-.0203</td>
<td>.0119</td>
<td>-1.708</td>
<td>.088</td>
</tr>
<tr>
<td></td>
<td>Congruency Effect</td>
<td>-1.0230</td>
<td>.8996</td>
<td>-1.140</td>
<td>.265</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>.0343</td>
<td>.0166</td>
<td>2.065</td>
<td>.039*</td>
</tr>
</tbody>
</table>

4.2.4 Summary of findings: Experiment 1

In Experiment 1, separate LMEMs were conducted for each region of interest, with the fixed effects of transitivity and incongruency, and with random intercepts for subject and item and by-subject random slopes for transitivity. Main effects for transitivity and incongruency were found at Region 3. The effect for transitivity reflects slower reading times for intransitive sentences, in line with Jegerski’s (2012) findings. The effect for incongruency reflects slower reading times following incongruent flanker trials. However, an interaction between transitivity and incongruency revealed that the increase in reading time for intransitive sentences was neutralized when they occurred after incongruent trials. Finally, exploratory analyses of cognitive individual differences suggest that the findings in Experiment 1 reflect the cognitive condition elicited by the preceding flanker trial, and not differences across participants in working memory or cognitive control.
4.3 Experiment 2

The goals of Experiment 2 were similar to Experiment 1, but extending the questions to a different population of bilinguals: emergent bilinguals. The first goal was to test whether L2 Spanish bilinguals reveal the reading time effects in a cross-task experiment that have been found for other native and near-native Spanish speakers in self-paced reading tasks (RQ4). The second research question asked whether preceding incongruent non-linguistic stimuli facilitated the resolution of linguistic conflict (RQ5). Because of the decreasing role of cognitive variables at increasing proficiency levels in the L2, RQ6 asked whether proficiency played a role in the answers to RQ4 and RQ5. Like in Experiment 1, these research questions are intimately tied by the research design and so the results are presented together in the sections that follow. An additional goal of Experiment 2 was to explore whether cognitive individual differences explained the differences in reading times by sentence type better than the experimental manipulation of cognitive control engagement (RQ7). In what follows, I describe the participant sample for this experiment, present the descriptive and inferential statistics along with the results of the exploratory analyses, and offer a summary of the main findings.

4.3.1 Participants

Participants of Experiment 2 were 40 native speakers of English (29 female) enrolled in undergraduate or graduate courses at Georgetown University. Four participants were also native speakers of another language, but each of these participants reported dominance in English since before reaching school age. Participants’ additional language background information is summarized in Table 7.
Table 7. Participant background information: Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>M (SD)</th>
<th>Mdn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>21.25 (3.37)</td>
<td>20</td>
</tr>
<tr>
<td>Number of native languages</td>
<td>1.10 (.31)</td>
<td>1</td>
</tr>
<tr>
<td>Number of foreign languages</td>
<td>1.63 (.77)</td>
<td>1</td>
</tr>
<tr>
<td>Age of exposure to first foreign language</td>
<td>10.54 (3.06)</td>
<td>11</td>
</tr>
<tr>
<td>Age of exposure to Spanish</td>
<td>11.14 (3.30)</td>
<td>12</td>
</tr>
<tr>
<td>Age of self-expressed fluency in Spanish</td>
<td>18.74 (2.32)</td>
<td>19</td>
</tr>
<tr>
<td>Self-rated Spanish proficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>7.16 (1.27)</td>
<td>7</td>
</tr>
<tr>
<td>Understanding</td>
<td>7.99 (1.13)</td>
<td>8</td>
</tr>
<tr>
<td>Reading</td>
<td>8.05 (1.16)</td>
<td>8</td>
</tr>
<tr>
<td>Spanish EIT Score (out of 120)</td>
<td>94.7 (16.31)</td>
<td>98</td>
</tr>
</tbody>
</table>

Note: Participants’ reported native languages also included Cantonese (n = 1), Polish (n = 1), Punjabi (n = 1), and Tamil (n = 1). In addition to Spanish, participants reported knowledge of foreign languages including French (n = 8), Portuguese (n = 6), Catalan (n = 2), American Sign Language (n = 1), Arabic (n = 1), Chinese (n = 1), Galician (n = 1), Hindi (n = 1), Hungarian (n = 1), Italian (n = 1), Tamil (n = 1), and Turkish (n = 1).

Raw reading times were examined before any further analyses. The presentational software did not record reading time data for three participants. The reading times of the remaining 37 participants were trimmed and transformed as outlined in Chapter 3. As in the native speakers, group-level and subject-level maximums and minimums were established. The group-level maximum (4000 ms) and minimum (100 ms) did not apply in any case because all participants’ subject-level cutoffs were within these bounds. Subject-level cutoffs were two standard deviations above and below the participants’ mean reading time. These bounds applied to 4.36% of the nonnative data, resulting in 323 replacements, all at the subject-level maximum. The reading times were then log-transformed to account for positive skew common among reaction time data (Luce, 1986).

Overall comprehension, including fillers and critical items, was assessed to establish that participants understood the sentences and maintained their attention. Mean
accuracy for the nonnative speakers was 87.7% ($SD = 6.19\%$) for all items, although mean accuracy for the critical items was lower ($M = 79.1\%, SD = 11.96\%$). One participant fell below the minimum comprehension threshold of 60%, leaving 36 participants for final analyses. It is important to note that for these remaining 36 participants, discrete trials were removed in all cases of an incorrect CQ.

Before the linguistic data was addressed, the flanker response times were analyzed to confirm that the non-linguistic components of the present study elicited the effects that have been established by previous literature. First, the independent (non-interleaved) flanker task was analyzed. This task was completed during the second data collection session, separate from the linguistic stimuli. A paired-samples t-test established that participants’ incongruent mean scores ($M = 578\text{ ms}, SD = 111\text{ ms}$) were statistically slower than their congruent mean scores ($M = 543\text{ ms}, SD = 100\text{ ms}$), mean of differences = 35.4 ms, $t(35) = 8.473, p < .001$, 95% CI = [27, 44]. This confirms that the independent (non-interleaved) flanker stimuli elicited a congruency effect in the canonical direction wherein congruent trials are faster than incongruent trials.

An analysis was also conducted to assess the flanker responses within the interleaved experiment in order to confirm that the interleaved design did not inhibit the congruency effect from emerging. A LMEM was conducted with reaction time as the dependent variable, incongruency as a fixed effect and a random effect structure with by-subject adjustments to the intercept and by-subject random slopes for incongruency. This revealed a congruency effect (CE) in the predicted direction (incongruent: $M = 743\text{ ms}, SD = 225\text{ ms}$; congruent: $M = 693\text{ ms}, SD = 214\text{ ms}$), with statistically slower responses for incongruent stimuli, estimate = 50.40, $SE = 11.45$, $t = 4.402, p < .001$ [log-
transformed response times: estimate = .0263, $SE = .0081, t = 3.242, p = .003$. Thus, both interleaved flanker stimuli and the independent flanker task elicited a CE in the canonical direction. Results of a Pearson correlation between participants’ two congruency effects revealed that participants’ performance on flanker stimuli in both the independent and interleaved tasks was also correlated, $r(36)= .42, p = .010$, 95% CI [.11, .66]. With these confirmation tests of the flanker stimuli conducted, the main analyses of the linguistic data were run. These results are presented in the following section.

4.3.2 Results: Research questions 4, 5, and 6

Descriptive statistics for the reading times of emergent bilingual sample are presented in Table 8. The data is presented graphically for each condition in Figure 7.

Table 8. Descriptive statistics of emergent bilingual raw reaction time means by condition

<table>
<thead>
<tr>
<th>Region</th>
<th>Sentence Type</th>
<th>All flankers $M$ ($SD$)</th>
<th>Congruent $M$ ($SD$)</th>
<th>Incongruent $M$ ($SD$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>Transitive</td>
<td>1112 ms (314 ms)</td>
<td>1145 ms (332 ms)</td>
<td>1063 ms (327 ms)</td>
</tr>
<tr>
<td></td>
<td>Intransitive</td>
<td>1075 ms (310 ms)</td>
<td>1084 ms (322 ms)</td>
<td>1058 ms (319 ms)</td>
</tr>
<tr>
<td></td>
<td>All sentences</td>
<td>1094 ms (308 ms)</td>
<td>1114 ms (322 ms)</td>
<td>1060 ms (311 ms)</td>
</tr>
<tr>
<td>Region 2</td>
<td>Transitive</td>
<td>491 ms (208 ms)</td>
<td>491 ms (191 ms)</td>
<td>491 ms (321 ms)</td>
</tr>
<tr>
<td></td>
<td>Intransitive</td>
<td>487 ms (198 ms)</td>
<td>479 ms (184 ms)</td>
<td>500 ms (315 ms)</td>
</tr>
<tr>
<td></td>
<td>All sentences</td>
<td>489 ms (199 ms)</td>
<td>485 ms (177 ms)</td>
<td>495 ms (301 ms)</td>
</tr>
<tr>
<td>Region 3</td>
<td>Transitive</td>
<td>916 ms (258 ms)</td>
<td>917 ms (250 ms)</td>
<td>913 ms (330 ms)</td>
</tr>
<tr>
<td></td>
<td>Intransitive</td>
<td>912 ms (289 ms)</td>
<td>918 ms (285 ms)</td>
<td>900 ms (325 ms)</td>
</tr>
<tr>
<td></td>
<td>All sentences</td>
<td>914 ms (264 ms)</td>
<td>918 ms (258 ms)</td>
<td>907 ms (306 ms)</td>
</tr>
<tr>
<td>Region 4</td>
<td>Transitive</td>
<td>737 ms (239 ms)</td>
<td>785 ms (275 ms)</td>
<td>666 ms (216 ms)</td>
</tr>
<tr>
<td></td>
<td>Intransitive</td>
<td>809 ms (255 ms)</td>
<td>783 ms (273 ms)</td>
<td>858 ms (339 ms)</td>
</tr>
<tr>
<td></td>
<td>All sentences</td>
<td>771 ms (234 ms)</td>
<td>784 ms (248 ms)</td>
<td>751 ms (247 ms)</td>
</tr>
<tr>
<td>Region 5</td>
<td>Transitive</td>
<td>1327 ms (437 ms)</td>
<td>1359 ms (483 ms)</td>
<td>1280 ms (450 ms)</td>
</tr>
<tr>
<td></td>
<td>Intransitive</td>
<td>1395 ms (391 ms)</td>
<td>1433 ms (376 ms)</td>
<td>1322 ms (524 ms)</td>
</tr>
<tr>
<td></td>
<td>All sentences</td>
<td>1360 ms (410 ms)</td>
<td>1396 ms (427 ms)</td>
<td>1298 ms (454 ms)</td>
</tr>
</tbody>
</table>

9 It is important to remember that these are mean reading times across all proficiency levels, but that the linear mixed-effects models use proficiency as a continuous factor to predict how transitivitiy, incongruency and proficiency interact. In other words, the model may find trends in the data that the raw data in this table hides by only presenting means.
LMEMs were conducted for each of the critical regions (Regions 3, 4, and 5), with the interacting fixed effects of transitivity, incongruency of the preceding trial, and proficiency. As in Experiment 1, the reference level for the factors were established as follows: for transitivity, as the intransitive condition, and for incongruency, as the congruent condition. At each critical region, the model with the maximal random effect structure included by-participant and by-item random intercepts, as well as by-participant random slopes for the fixed effect of transitivity. The results of the LMEMs for the critical regions are included in Table 9. Results of the LMEMs for Regions 1 and 2 are included in Appendix D.

Figure 7. Reading times for emergent bilinguals by condition
Table 9. Output for reading times from emergent bilingual linear mixed-effects models

<table>
<thead>
<tr>
<th>Region</th>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 3</td>
<td>Constant Intercept</td>
<td>3.361</td>
<td>.1313</td>
<td>25.592</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>- .0809</td>
<td>.0673</td>
<td>-1.202</td>
<td>.234</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.0673</td>
<td>.0681</td>
<td>- .989</td>
<td>.323</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>- .0045</td>
<td>.0013</td>
<td>-3.333</td>
<td>.002**</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>.1270</td>
<td>.0919</td>
<td>1.383</td>
<td>.167</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>.0007</td>
<td>.0007</td>
<td>1.002</td>
<td>.321</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>.0006</td>
<td>.0007</td>
<td>.850</td>
<td>.396</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>-.0012</td>
<td>.0009</td>
<td>-1.248</td>
<td>.212</td>
</tr>
<tr>
<td>Region 4</td>
<td>Constant Intercept</td>
<td>3.221</td>
<td>.1390</td>
<td>23.178</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>.0601</td>
<td>.0758</td>
<td>.793</td>
<td>.431</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.0531</td>
<td>.0807</td>
<td>.658</td>
<td>.511</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>- .0042</td>
<td>.0014</td>
<td>-3.083</td>
<td>.004**</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.0974</td>
<td>.1088</td>
<td>-.895</td>
<td>.371</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0007</td>
<td>.0008</td>
<td>-.854</td>
<td>.396</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0002</td>
<td>.0008</td>
<td>-.285</td>
<td>.775</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>.0005</td>
<td>.0011</td>
<td>.458</td>
<td>.647</td>
</tr>
<tr>
<td>Region 5</td>
<td>Constant Intercept</td>
<td>3.457</td>
<td>.1229</td>
<td>28.124</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>.1807</td>
<td>.0658</td>
<td>2.746</td>
<td>.008**</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.1574</td>
<td>.0764</td>
<td>2.061</td>
<td>.040*</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0039</td>
<td>.0012</td>
<td>-3.131</td>
<td>.003**</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.2451</td>
<td>.1028</td>
<td>-2.385</td>
<td>.017*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0018</td>
<td>.0007</td>
<td>-2.707</td>
<td>.008**</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0016</td>
<td>.0008</td>
<td>-2.036</td>
<td>.042*</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>.0022</td>
<td>.0010</td>
<td>2.149</td>
<td>.032*</td>
</tr>
</tbody>
</table>

At both Region 3 and Region 4, there was only a main effect for proficiency, showing that increased proficiency predicted faster reading times at these regions, irrespective of transitivity or incongruency of the condition. The predicted slowdown effect associated with the intransitive condition at Region 3 was not statistical, nor was there a statistical slowdown in the transitive condition at Region 4.

At Region 5, the spillover region, the main effect for proficiency was also found, estimate = -.0039, \( SE = .0012, t = -3.131, p = .003 \), with a medium effect size, \( f^2 = .148 \).

There were also main effects for both transitivity and incongruency in the predicted directions. Transitivity increased reading times, estimate = .1807, \( SE = .0658, t = 2.746, p \)
= .008, with an effect size of $d = .562$, considerably higher than the reference framework proposed by Avery & Marsden (2019) for this type of comparison (i.e., $d = .34$ for conditions of L1-L2 similarity). This likely reflecting wrap-up processing effects associated with the fronted absolute construction of the transitive verb. Incongruency increased reading times compared to congruency, estimate $= .1574$, $SE = .0764$, $t = 2.061$, $p = .040$, with an effect size of $d = .264$. This is in line with the prediction that the performance adjustment associated with increased conflict monitoring causes the parser to slow down following conflict detection. In addition to the main effects, each two-way and three-way interaction between proficiency, transitivity, and incongruency was also significant. These interactions are addressed below. To visualize the difference between the conditions, the log-transformed reading times for three proficiency scores are presented in Figure 8 and Figure 9. The graphs depicting the model are represented in Figure 10.

**Figure 8. Predicted logRTs for emergent bilinguals in each condition at three proficiency scores**
Figure 9. Predicted logRTs for emergent bilinguals across conditions

Figure 10. Predicted means for emergent bilinguals in congruent and incongruent condition
The negative interaction between transitivity and proficiency reflects that the main effect of transitivity at Region 5 decreases with increasing proficiency, estimate = \(-.0018\), SE = .0007, \(t = -2.707\), \(p = .008\). Likewise, the slowdown associated with incongruency affects lower-proficiency learners more than for higher-proficiency learners, estimate = \(-.0016\), SE = .0008, \(t = -2.036\), \(p = .042\). Finally, despite the model’s positive coefficients for the main effects of transitivity and incongruency (i.e., slower reading times in these conditions), these two variables in tandem resulted in faster reading times as predicted, *ceteris paribus*, estimate = \(-.2451\), SE = .1028, \(t = -2.385\), \(p = .017\), with an effect size of \(d = .389\). Thus, although incongruency slows reading times at Region 5 when all other factors are held equal (i.e., when the model’s reference level is the intransitive condition), the increased conflict monitoring that incongruency triggers proves to be advantageous when processing the linguistic conflict in a fronted absolute transitive clause.

Whereas the interaction between incongruency and transitivity was negative, their three-way interaction with proficiency was positive. Post-hoc analyses were conducted using R’s *emmeans* package (Lenth et al, 2020) to isolate the source of this interaction, the results of which are presented in Tables 10 and 11. The alpha-level was reduced to .0125 to correct for multiple comparisons. In all comparisons, proficiency was a significant predictor, reflecting decreased reading times at increasing proficiencies. Within the congruent condition, there was a significant difference between the two sentence types, estimate = .1728, SE = .6067, \(t = 2.849\), \(p = .0045\), as well as the interaction between proficiency and transitivity, estimate = \(-.0018\), SE = .0006, \(t = -3.007\), \(p = .0027\). We can interpret this as follows: for the lowest proficiency levels, in
cases following congruent flanker stimuli, Region 5 following transitive sentences was processed slower than Region 5 following intransitive sentences. However, with each 1-unit increase in EIT score, the difference between the two sentence types decreases by .0018. Between EIT = 100 and EIT = 101, the two sentence types switch: intransitive processing becomes slower than transitive processing, which reflects the ordering of the sentence types in the incongruent condition at Region 5 in this experiment (It also reflects the Region 5 reading time data from the native bilinguals in Experiment 1). Meanwhile, in the incongruent condition, only proficiency predicted faster reading times, estimate = .0051, \( SE = .0015, t = -3.413, p = .0014 \). Transitive and intransitive sentences did not differ statistically, and there was no interaction between proficiency and transitivity.

### Table 10. Comparison of Region 5 output in congruent and incongruent conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Parameter</th>
<th>Estimate</th>
<th>( SE )</th>
<th>( t )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Congruent</strong></td>
<td>Transitivity</td>
<td>.1728</td>
<td>.0607</td>
<td>2.849</td>
<td>.0045**</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0038</td>
<td>.0013</td>
<td>-2.836</td>
<td>.0073**</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0018</td>
<td>.0006</td>
<td>-3.007</td>
<td>.0027**</td>
</tr>
<tr>
<td><strong>Incongruent</strong></td>
<td>Transitivity</td>
<td>-.0605</td>
<td>.0825</td>
<td>-.733</td>
<td>.4641</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td><strong>-.0051</strong></td>
<td><strong>.0015</strong></td>
<td><strong>-3.413</strong></td>
<td><strong>.0014</strong>**</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>.0003</td>
<td>.0008</td>
<td>.394</td>
<td>.6939</td>
</tr>
</tbody>
</table>

*Note: Alpha level reduced for multiple comparisons, \( \alpha = .0125 \)*

Only proficiency was a significant predictor in the pairwise comparisons of the role of incongruency in the transitive and intransitive conditions. The results are presented below in Table 11. Taken with the results in Table 10, the post-hoc comparisons provide two key findings: first, that at higher proficiency levels, learners progressively process the different sentence types more similarly; and secondly and more to the point, that incongruent flanker stimuli neutralize the reading time differences
between transitive and intransitive sentences, which only manifested in the congruent condition.

### Table 11. Comparison of Region 5 output in transitive and intransitive conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Incongruency</td>
<td>-.1066</td>
<td>.0077</td>
<td>-1.376</td>
<td>.1693</td>
</tr>
<tr>
<td>Transitive</td>
<td>Proficiency</td>
<td>-.0058</td>
<td>.0014</td>
<td>-4.256</td>
<td>.0001***</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>.0008</td>
<td>.0008</td>
<td>1.003</td>
<td>.3165</td>
</tr>
<tr>
<td>Intransitive</td>
<td>Incongruency</td>
<td>.1523</td>
<td>.0728</td>
<td>2.093</td>
<td>.0369</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0039</td>
<td>.0012</td>
<td>-3.189</td>
<td>.0029**</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0016</td>
<td>.0007</td>
<td>-2.122</td>
<td>.0343</td>
</tr>
</tbody>
</table>

Note: Alpha level reduced for multiple comparisons, $\alpha = .0125$

### 4.3.3 Results: Research question 7

In order to consider the role of cognitive individual differences, the non-natives’ results from the flanker tasks (CE and congruency sequence effect, CSE) and the results of the Operation Span task were included in a series of subsequent analyses that were exploratory in nature. The CSE is a measure of conflict adaptation within Eriksen flanker tasks. It incorporates conflict adaptation by separating trials according to the preceding condition, and is calculated according to the following formula, in which the subscripts represent the congruency of the preceding trial and the critical trial:

$$CSE = (RT_{ci} - RT_{cc}) - (RT_{ii} - RT_{ic})$$

Higher CSEs reflect a greater difference between $RT_{ci}$ and $RT_{ii}$. In other words, participants with high CSEs reveal more conflict adaptation than participants with low CSEs

The LMEMs with these variables were not used for confirmatory hypothesis testing. The descriptive results of the cognitive tasks are included in Table 12.
Participants’ scores were separately incorporated into the prior model, Model 1, as covarying fixed effects to determine if more variance was accounted for when one of these cognitive individual differences was included.

| Table 12. Descriptive statistics of emergent bilingual cognitive measures |
|--------------------------------|----------|---------|
| Cognitive Measure | M (SD)   | Median  |
| Automated Operation Span | 53.45 (15.09) | 52.5    |
| Congruency Effect (from independent task) | 36 ms (25 ms) | 34 ms   |
| Congruency Sequence Effect (from independent task) | -16 ms (99 ms) | -13 ms  |

The three models with cognitive covariates were compared to Model 1 using chi-squared tests, the results of which are presented below in Table 13, along with the percentage of variance in the data that each model accounted for. None of the three models was statistically different from Model 1. In addition, the results of the new models are reflective of the results of Model 1 while none of the newly introduced cognitive individual differences was found to be a significant predictor, as shown in Table 14.

| Table 13. Chi-squared statistics for three exploratory models compared to Model 1 |
|----------------------------|-------|----------------|--------|---------|
| Model                   | $X^2$ | df  | p       | $R^2$ |
| Model 1                 | --    | --  | --      | .1301  |
| Model 1:OSPAN           | .038  | 1   | .846    | .1296  |
| Model 1:CE              | 2.475 | 1   | .605    | .1309  |
| Model 1:CSE             | 1.1069 | 1 | .293    | .1358  |

These results suggest that cognitive condition accounts for the differences in reading time more than individual differences across participants. This will be addressed further in Chapter 5.
Table 14. Output from three exploratory linear mixed-effects models compared to Model

<table>
<thead>
<tr>
<th>Region 5 Model</th>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>p.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeli</td>
<td>Transitivity</td>
<td>.1807</td>
<td>.0658</td>
<td>2.746</td>
<td>.0075**</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.1574</td>
<td>.0764</td>
<td>2.061</td>
<td>.0395*</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0039</td>
<td>.0112</td>
<td>-3.131</td>
<td>.0034**</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.2451</td>
<td>.1028</td>
<td>-2.385</td>
<td>.0173*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0018</td>
<td>.0007</td>
<td>-2.707</td>
<td>.0084**</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0016</td>
<td>.0008</td>
<td>-2.036</td>
<td>.0420*</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>.0022</td>
<td>.0010</td>
<td>2.149</td>
<td>.0319*</td>
</tr>
<tr>
<td>Modelospan</td>
<td>Transitivity</td>
<td>.1812</td>
<td>.0658</td>
<td>2.753</td>
<td>.007**</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.1582</td>
<td>.0764</td>
<td>2.071</td>
<td>.039*</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0038</td>
<td>.0113</td>
<td>-3.032</td>
<td>.004**</td>
</tr>
<tr>
<td></td>
<td>OSpan</td>
<td>-.0006</td>
<td>.0112</td>
<td>-4.70</td>
<td>.642</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.2460</td>
<td>.1028</td>
<td>-2.393</td>
<td>.017*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0018</td>
<td>.0007</td>
<td>-2.713</td>
<td>.008**</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0016</td>
<td>.0008</td>
<td>-2.044</td>
<td>.041*</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>.0022</td>
<td>.0010</td>
<td>2.156</td>
<td>.031*</td>
</tr>
<tr>
<td>Models</td>
<td>Transitivity</td>
<td>.1807</td>
<td>.0658</td>
<td>2.747</td>
<td>.007**</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.1567</td>
<td>.0763</td>
<td>2.052</td>
<td>.040*</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0041</td>
<td>.0113</td>
<td>-3.225</td>
<td>.003**</td>
</tr>
<tr>
<td></td>
<td>Congruency Effect</td>
<td>.3303</td>
<td>.3700</td>
<td>.893</td>
<td>.379</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.2444</td>
<td>.1028</td>
<td>-2.379</td>
<td>.018*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0018</td>
<td>.0007</td>
<td>-2.706</td>
<td>.008**</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0016</td>
<td>.0008</td>
<td>-2.024</td>
<td>.043*</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>.0022</td>
<td>.0010</td>
<td>2.140</td>
<td>.033*</td>
</tr>
<tr>
<td>Modelc</td>
<td>Transitivity</td>
<td>.1791</td>
<td>.0659</td>
<td>2.719</td>
<td>.008**</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.1552</td>
<td>.0763</td>
<td>2.034</td>
<td>.042*</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0044</td>
<td>.0113</td>
<td>-3.302</td>
<td>.002**</td>
</tr>
<tr>
<td></td>
<td>Congruency Sequence Effect</td>
<td>.2073</td>
<td>.1970</td>
<td>1.052</td>
<td>.300</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.2424</td>
<td>.1027</td>
<td>-2.360</td>
<td>.018*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0018</td>
<td>.0007</td>
<td>-2.685</td>
<td>.009**</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0016</td>
<td>.0008</td>
<td>-2.011</td>
<td>.045*</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>.0022</td>
<td>.0010</td>
<td>2.127</td>
<td>.034*</td>
</tr>
</tbody>
</table>

4.3.4 Summary of findings: Experiment 2

In Experiment 2, separate LMEMs were conducted for each region of interest, with the fixed effects of transitivity, incongruency, and proficiency, and with random intercepts for subject and item and by-subject random slopes for transitivity. Main effects for transitivity, incongruency, and proficiency were found at Region 5, with slower reading times for transitive sentences, for sentences following incongruent flankers, and for participants with lower proficiency scores. There was an interaction between transitivity
and incongruency, whereby the main effect of transitivity and incongruency was reversed when these two conditions occurred in tandem. There were also interactions between transitivity and proficiency and between incongruency and proficiency, whereby lower proficiency participants revealed more of a processing effect for transitivity and for incongruency than higher proficiency participants. Finally, there was a three-way interaction between transitivity, incongruency, and proficiency. Post-hoc analyses using R’s emmeans package (Lenth et al, 2020) revealed that this three-way interaction was driven by the congruent condition, where transitive and intransitive conditions differed significantly, revealing that preceding incongruent flankers neutralized the difference between sentence types. However, the direction of the interaction also revealed that at increasing proficiencies, participants processed the two sentence types more similarly, even in the congruent condition. Finally, exploratory analyses suggest that the findings in Experiment 2 can be attributed to cognitive condition and not to individual differences in working memory or cognitive control.
5.1 Overview

In the chapter that follows, I discuss the results of Experiments 1 and 2 in relation to the hypotheses presented in Chapter 3, as well as how these findings fit within current work in bilingualism research, including psycholinguistics and second language acquisition. Finally, I discuss the limitations of the research and directions for future research.

5.2 Experiment 1

Experiment 1 was driven by the question of whether early bilinguals would reveal the same conflict resolution effects of cognitive-control engagement that monolinguals have shown (Hsu & Novick, 2016). This question is pertinent given that previous research shows their cognitive control is constantly activated to suppress their contextually inappropriate language (see Kroll, Dussias, Bogulski, & Valdes Kroff, 2012). Thus, using the stimuli from Jegerski (2012), Experiment 1 first asked whether early bilinguals revealed the same sensitivities to fronted adverbial clauses as monolinguals (RQ1), and if so, if cognitive-control engagement reduces the resulting reading time effects (RQ2). The role of different cognitive variables was explored in RQ3.

5.2.1 Research question 1

The transitivity dichotomy led to the prediction of a reading time effect at Region 3, specifically slower reading times in the intransitive condition due to the clash between the parser’s impulse to incorporate the postverbal NP as a direct object according to Late Closure and the intransitivity of the verb in the preceding region. Trends in Jegerski
(2012) suggested that the transitive condition may be read slower at Region 4, similarly to the garden-path effect found in non-pro-drop languages. Region 5 was studied as a potential area of spillover or wrap-up effects.

The results of Experiment 1 correspond to the findings of Jegerski (2012): at Region 3, there was a statistical slow-down for intransitive sentences, but there were no statistical reading time effects at the other regions. The similarity of the findings in Jegerski (2012) and Experiment 1 is unsurprising, since both monolinguals and bilinguals should experience Late Closure in the same way. RQ1’s main purpose was to confirm that the experimental design did not alter predicted reading time effects, in order to reliably ask RQ2. The results of RQ2 are addressed below.

5.2.2 Research question 2

The thrust of this experiment is in RQ2. The incongruency manipulation included in the cognitive-control engagement paradigm led to two predicted reading time effects. First, that incongruency would lead to slower reading times as a result of increased conflict monitoring. This would manifest as a main effect for incongruency. Although Region 3 reading times that followed incongruent trials ($M = 758$ ms, $SD = 231$ ms) were descriptively slower than congruent trials ($M = 743$ ms, $SD = 207$ ms), the model did not reveal a significant main effect as predicted, estimate $= -.0203$, $SE = .0119$, $t = -1.709$, $p = .088$, and the effect size was $d = .365$. However, the second and more pertinent prediction concerning incongruency for this experiment concerned its interaction with transitivity. If pre-trial incongruency were to lead to greater conflict monitoring, and therefore decrease cognitive control’s lag time, the effects for transitivity were predicted to be diminished following incongruent trials. The interaction was statistically significant,
estimate = .0343, \( SE = .0166, t = 2.064, p = .039 \). Post hoc tests revealed that the main effect of transitivity was driven by the congruent condition where intransitive condition RTs were statistically slower than transitive condition RTs, estimate = .0287, \( SE = .0117, t = 2.449, p = .015 \). Transitive and intransitive sentences with preceding incongruent flankers were statistically equivalent. In other words, pre-trial non-linguistic conflict facilitates the resolution of linguistic conflict in Region 3.

### 5.2.3 Research question 3

The third research question was exploratory, given the paucity of prior research investigating cross-task conflict adaptation in early bilinguals. It asked how cognitive variables that are involved in language inhibition in early bilinguals would interact with the experimental manipulation of cognitive control (i.e., incongruent vs. congruent) conditions. The latency on the flanker trials in the interleaved task provided a measure for participants’ congruency effect and participants also completed the automated OSpan, a non-linguistic working memory task. LMEMs that included these individual differences as covariates were compared to the original model, Model1, but they were not significant covariates. Thus, it was determined that the cognitive control condition was a better predictor of reading time differences than cognitive individual differences in Experiment 1.

### 5.3 Experiment 2

Non-highly proficient, non-native processing has been shown to recruit cognitive control to greater degrees than native or highly proficient processing (Abutalebi, 2008). How this increased recruitment of cognitive control for L2 processing affects other cognitive control processes, like conflict adaptation, was the broad idea behind
Experiment 2. To begin to ask this question empirically, the cognitive-control engagement paradigm was extended to the emergent bilingual population. Do emergent bilinguals reveal the same conflict resolution effects from cognitive-control engagement that Hsu & Novick (2016) found for monolinguals? Experiment 2 first asked whether emergent bilinguals would reveal the same sensitivities to fronted adverbial clauses that near-native L2 speakers revealed in Jegerski’s (2012) study (RQ4), and if so, if cognitive-control engagement assuages the resulting reading time effects (RQ5). Whether proficiency modulates the role of preceding conflict in linguistic conflict resolution was addressed by RQ6, because of the prior research that has shown the decreasing role of cognitive control resources at increasing proficiency levels (e.g., Abutalebi, 2008; Hopp, 2006; Rah & Adone, 2010). Finally, the role of different cognitive variables was explored in RQ7.

5.3.1 Research question 4

In order to study the role of proficiency in this paradigm, participants were recruited from a range of proficiency levels that ranged from undergraduate students who had completed a short-term stay abroad to near-native speakers completing their doctoral studies. Given this difference in L2 learner profiles when compared to Jegerski’s (2012) near-native sample, the potential for different effects was recognized for Experiment 2. Importantly, the sample of this dissertation had a greater daily reliance on English than the near-natives in Jegerski’s study, a group that had lived in Mexico for an average of 14 years. Jegerski found that transitive reading times at Region 4 were descriptively slower for her near-native participants. Given that Region 4, the matrix verb, is the region of forced disambiguation for these sentences in non-pro-drop languages like English, the
higher dominance of English in this group of L2 bilinguals was predicted to influence their non-native processing and reveal differences at Region 4. In addition, prior research, including the results of the pilot studies that were conducted, suggests that processing effects observed in L1 speakers reveal themselves later for L2 speakers, so Region 5 was also of great interest for reading time effects.

The main effect for transitivity found at Region 3 in Experiment 1 and for near-natives in Jegerski (2012), namely, slower reading times for intransitive sentences, was not found in Experiment 2. There was also no main effect for transitivity at Region 4. Descriptively, the reading times do reflect slower reading times of the intransitive sentences at Region 3 and the transitive sentences at Region 4, as predicted. At Region 5, a significant main effect was found, revealing statistically slower transitive sentences, estimate = .1807, $SE = .0658$, $t = 2.746$, $p = .008$. This likely reflects wrap-up effects caused by the subject-object ambiguity in the transitive condition, which will be explored further below.

5.3.2 Research question 5

As in Experiment 1, a main effect for incongruency was predicted, whereby incongruent flankers were predicted to slow reading times by increasing conflict monitoring. However, more to the point for this dissertation, a negative interaction between transitivity and incongruency was also predicted if there was a main effect for transitivity, as increased conflict monitoring was expected to facilitate recovery from the linguistic conflict, as found in Hsu & Novick (2016). In other words, the increased conflict monitoring caused by the preceding incongruent flanker was predicted to slow down reading times for non-conflictive sentences, but if conflict was encountered in the linguistic
input, the active conflict monitoring was predicted to facilitate the resolution of this linguistic conflict.

Both a main effect for incongruency and an interaction between congruency and transitivity were found at Region 5. The main effect for incongruency reflected slower reading times for the incongruent condition compared to the congruent condition in the reference level (i.e., i-I vs. c-I), estimate = .1574, SE = .0764, t = 2.061, p = .040. However, the interaction revealed that congruent-transitive stimulus-pairs were processed slower at Region 5 than the reference level (i-I), estimate = .1807, SE = .0658, t = 2.746, p = .008, and also slower than incongruent-transitive stimulus pairs, estimate = -.2451, SE = .1028, t = -2.385, p = .017. In other words, although incongruent flankers slowed reading times of intransitive sentences at Region 5, incongruency facilitated the processing of the transitive sentences at Region 5 for L2 speakers. This can be interpreted to signal that the incongruent flankers trigger conflict monitoring, which in turn results in slower processing of the input, unless the parser finds conflict to resolve.

5.3.3 Research question 6

Proficiency was predicted to have negative main effect on reading times, but a three-way interaction between transitivity, incongruency, and proficiency was also predicted. Although this paradigm is novel, the cognitive control resources more available to higher proficiency learners than lower proficiency learners led to a prediction that high proficiency learners would reap more benefit from the cognitive-control engagement paradigm (i.e., would resolve linguistic conflict faster after incongruent trials, which would result in any differences being more substantial for higher proficiency learners than for lower proficiency learners).
The negative effect of proficiency on reading times was obtained in all regions as a main effect. Unsurprisingly, proficiency’s predicted interactions with the other variables were not obtained where no main effects for these variables was found. However, at Region 5, an interaction with transitivity and incongruency was found. This interaction revealed that the magnitude of the negative effect of incongruency on transitivity was reduced with increasing proficiency, estimate = .0022, $SE = .0010$, $t = 2.149$, $p = .032$. This means that in response to the incongruent flanker trials, higher proficiency learners revealed smaller differences between the two sentence types than lower proficiency learners. Through pairwise comparisons to elucidate the source of the three-way interaction, it was revealed that significant differences in the results were driven by the congruent condition, where sentence type was a significant predictor for differences in reading time and where there was an interaction between transitivity and proficiency. No difference between sentence types and no interaction was found in the incongruent condition. In the pairwise comparisons, incongruency was not a significant predictor in either condition, though its effect was negative on transitive sentences and positive on intransitive conditions, as predicted. These results were interpreted to show that the incongruency of the preceding trial facilitated the resolution of linguistic conflict that manifested in trials with congruent preceding trials.

This direction for the three-way interaction was not predicted. The prediction was that at increasing proficiencies, cognitive-control engagement would have a greater effect to facilitate the processing of the transitive sentences at Region 5. Instead, the findings suggest that at increasing proficiencies, the differences between the two sentence types becomes narrower across conditions. Consider the trajectories and patterns from Figures 8
and 10 above (page 103 and 104). This unexpected direction of the interaction likely reflects that the wrap-up effects found at Region 5 in this experiment are largely due to the lower proficiency learners, who are driving the positive main effect for transitivity, and because of this, they are the participants who have the most to gain from cognitive-control engagement. This and other possibilities will be explored below in Section 5.6.

**5.3.4 Research question 7**

The final research question in Experiment 2, RQ7, did not engage in confirmatory hypothesis testing, but rather was exploratory in nature, asking whether cognitive individual differences revealed more about participants’ performance on the self-paced reading task, or if the manipulation of cognitive control through the interleaved flanker stimuli accounted for the differences. To answer this question, the three cognitive measures that were collected (OSpan, CE, and CSE) were included in separate LMEMs as covariates. The inclusion of the covariates did not alter the significance of the other predictors when compared to the results of Model1. However, the covariates themselves were not statistically significant and the models did not statistically differ from Model1. This was interpreted to tell us that it is cognitive condition (i.e., the status of the preceding trial) that accounts for the differences in reading time across the conditions of the experiment and not cognitive individual differences. This likely reflects that participants’ performance changes, comparably across the board, according to whether the preceding trial gives cognitive control a head-start to resolve the linguistic conflict that follows, not according to these cognitive individual differences.

**5.4 Conclusions**
The results from both Experiment 1 and Experiment 2 reveal that the cognitive-control engagement paradigm (Hsu & Novick, 2016) does transcend both from monolingual speakers of English to bilingual speakers of Spanish (both early and emergent), and from the visual world paradigm to self-paced reading. In this dissertation, where a significant main effect for reading time by sentence type was found, the effect was diminished or neutralized when conflict was presented in the immediately preceding flanker trial. These results are a testament to the importance of the context in which we process our input. Although the models above predict differences between sentence types following neutral congruent flankers, by presenting conflictive flankers before these same sentences, the differences that were found in the congruent condition do not manifest. In other words, conflict adaptation induced by non-linguistic stimuli facilitates the resolution of linguistic conflict.

The linguistic stimuli in this dissertation were first employed by Jegerski (2012), who compared Spanish monolinguals and near-native L2 Spanish learners. If we consider only the congruent conditions, the early bilinguals in Experiment 1 replicate the results from Jegerski’s (2012) experiment. In that study, a main effect for transitivity was found only at Region 3, where the postverbal NP was processed slower following the intransitive verbs than the transitive verbs. This effect is associated with a mismatch between Frazier’s Late Closure principle, which triggers the parser to incorporate the postverbal NP as a complement of the verb, and the intransitive verb’s subcategorization matrix, which does not allow for any complements.

From the basis of Jegerski’s (2012) publication, the current dissertation not only extends the question to a new population, namely early (and balanced) bilinguals, it also
introduces preceding flanker trials to manipulate cognitive state during sentence processing. The findings indicate that when the preceding trial is incongruent, the resulting conflict detection processes neutralize the difference between the two sentence types. Although Hsu & Novick (2016) studied native English speakers responding to a different linguistic structure via the visual world paradigm, they also found that this form of cognitive-control engagement, from non-linguistic to linguistic stimuli, facilitated or accelerated sentence reinterpretation processes involved in recovery from garden paths.

Experiment 1 was driven by the question of whether early bilinguals would reveal these same effects of cognitive-control engagement because of previous research that shows their cognitive control is constantly activated to suppress their contextual inappropriate language (see Kroll, Dussias, Bogulski, & Valdes Kroff, 2012). When cognitive control is already engaged for language suppression, can further engagement facilitate conflict resolution? The answer proposed here is yes. Does this reflect that cognitive-control engagement does not affect other background conflict resolution processes, or that slightly different processes are engaged in language suppression and conflict resolution, or that a bilingual’s lifelong training in language suppression provides enough of an advantage to engage in both tasks simultaneously? These are questions that remain for future research, and are addressed further below.

Jegerski (2012) established that near-native late L2 bilinguals experience the same sensitivity to these subject-object that native monolingual speakers experience. This dissertation, in Experiment 2, asked if emergent bilinguals of lower proficiency levels also reveal this effect. Whereas Jegerski’s participants had a mean age of 41.6 years and had lived in Mexico for an average of 14.1 years, the present sample’s mean age was 21.3
years, with much more limited experience living abroad ($M = 7.2$ months, $SD = 12$ months), and at most had spent $4$ years abroad. This sample was chosen precisely to elucidate how cognitive control and proficiency interact in sentence processing at non-native, non-highly proficient proficiencies, as studies suggest that cognitive control has a decreasing importance in processing (e.g., Abutalebi, 2008; Hopp, 2006; Rah & Adone, 2010) and learning (Serafini & Sanz, 2016) at increasing proficiencies. The results show that this sample on the whole does not reveal the conflict at Region 3, like monolinguals, early bilinguals, or near-native bilinguals. Rather, this sample reveals a wrap-up effect in Region 5 from the garden path of the transitive condition. This delay is not entirely unexpected, given the added demands of low proficiency L2 processing (e.g., cognitive capacity limitations, slower processing speed, slower lexical access; see Cunnings’ (2017) keynote article and its responses in *Bilingualism: Language and Cognition, 20*).

At this region, the observed pattern is similar to the findings from Experiment 1: the transitive and intransitive sentences are only processed differently in the congruent condition, while incongruent preceding trials neutralize this difference. This effect, however, runs in the opposite direction, where the transitive condition is processed slower because of the subject-object garden path created by the fronted transitive verb in the absolute condition.

The effect is also constrained by proficiency, with a greater reduction in the garden-path effect as a result of the conflict manipulation for low proficiency learners than for higher proficiency learners. For example, the model predicts a participant with an EIT score of 80 will yield (log-transformed) reading times in these conditions that differ by 134 ms ($c-T = 1520, i-T = 1387$), while a participant with an EIT score of 100
will yield reading times differing by 73 ms (c-T = 1169, i-T = 1096), and a participant at the highest possible EIT score of 120 will yield reading times that differ by 32 ms (c-T = 899, i-T = 867). This likely reflects floor effects, as the predicted means of the model for this highest possible EIT score of 120 are even lower than the early bilinguals’ mean reading times in Experiment 1 (c-T = 971, i-T = 940). This finding was not predicted, but it is not entirely surprising. The narrowing effect of the preceding incongruent trial was likely merely more visible when learners were reading at non-native-like speeds. In other words, the prediction that low proficiency learners would not be advantaged by preceding conflict because of maxed out cognitive control resources did not stand. Lower proficiency levels did resolve conflict faster at Region 5 when primed for conflict by the preceding conflict.

A limitation of the present study, to be addressed below, relates to the power of Experiment 2, particularly as it pertains to the upper extreme of the proficiency range, who may be experiencing conflict at earlier regions in the sentence that is not statistical because of how lower proficiency learners’ data adjust the curve. This question remains for future research, but from this sample, we can confirm that delayed garden-path effects that appear at Region 5 in lower proficiency learners (and in this sample as a whole) are lessened by preceding conflict.

It’s worth reminding the reader here that the lowest proficiency learners of Experiment 2 are not low proficiency learners in the general understanding of the term, but rather within the context of this dissertation. Following the pilot study, the complex sentences were determined to be too difficult for intermediate learners to reliably and consistently process, so the learners in this sample had at least completed the equivalent
of six semesters of college Spanish (up to and including Advanced 2) and were completing a 5-week stay abroad. The prediction that low proficiency learners would not react to the paradigm in the same way as more proficient participants because of their increased recruitment of cognitive control resources for L2 processing, which was informed by research on children, seniors, and impaired patients with cognitive control deficiencies, was disconfirmed. Nonetheless, the findings of this dissertation, including this disconfirmed prediction, contribute to several areas of investigation and open several avenues for future research, the goals of the next two sections of this chapter.

5.5 Implications

This dissertation is a first foray into bilingual processing and the cognitive-control engagement paradigm, but despite its novelty, the results carry several implications. These implications are addressed below.

Firstly, this use of the cognitive-control engagement paradigm constitutes its first use to my knowledge to explore bilingual processing, its first use to study linguistic conflict in Spanish, and its first use with self-paced reading instead of the visual world paradigm. Despite these changes, the pre-trial presentation of non-linguistic conflict still facilitated the resolution of subsequent linguistic conflict. Given that the effect of cross-task conflict adaptation found by Hsu & Novick (2016) is externally valid, or transferrable to different populations (early and emergent bilinguals) processing a different language (Spanish) in a different linguistic form (reading), the paradigm can be used to explore language processing and cognitive control for a wide variety of subject profiles, linguistic typologies and linguistic tasks, in addition to the follow-up questions posed in Section 5.6 for the populations in this dissertation. The dissertation offers
psycholinguists validation for Hsu & Novick’s novel tool that will facilitate the exploration of many questions about ambiguity processing and cognitive capacity.

Beyond the methodological implication, the dissertation contributes to literature on subject-object ambiguity resolution in Spanish, a pro-drop language. Specifically, the study shows that early bilinguals experience the same processing effect demonstrated by monolingual and near-native L2 Spanish speakers, caused by the conflict between the parser’s preference to include the Region 3 NP as a direct object in the fronted clause (complying with the Principle of Late Closure) and the intransitive verbs’ subcategorization matrix disallowing direct objects. That this conflict extends to early bilinguals is not surprising as they build the same subcategorization matrices as their monolingual and L2 peers. However, the results confirm that Jegerski’s (2012) findings are a general phenomenon for native and near-native parsers.

Experiment 2 revealed that lower-proficiency L2 processing of fronted subject-object ambiguities in Spanish does not obtain the same results as near-native L2 processing. This is not entirely surprising, but both the absence of the previously observed processing effect at Region 3 and the presence of an effect at Region 5 open several avenues for future research, which are addressed in Section 5.6.

5.6 Limitations and Future Research

Whereas both experiments have shown that pre-trial cognitive-control engagement can facilitate the resolution of linguistic conflict, they do not elucidate how these populations might differ from monolinguals or from each other in terms of cognitive control recruitment, precisely because both groups responded to the paradigm similarly to monolinguals. Questions concerning this cross-task conflict adaptation
remain. For example, does cognitive-control engagement not affect background conflict resolution processes, like language suppression? Novick and colleagues’ research involving patients and children (e.g. Huang et al., 2016; Huang et al., in revision), two populations with cognitive control deficiencies, found that in cross-task designs, preceding conflict depleted cognitive control resources, so it seems unlikely that the manipulation does not affect any cognitive control processes, but perhaps language suppression in bilinguals is inherently different from other types of conflict resolution. More research on multilingual populations is needed to elucidate how the manipulation affects language suppression specifically. Some specific limitations and areas for further research that emerge from this idea are addressed below.

5.6.1 Experiment 1

The early bilinguals responded to cross-task cognitive control engagement much like the monolinguals in Hsu & Novick’s (2016) first study using the paradigm. Of course, as addressed above, setting aside any controversy in the literature, the bilingual advantage in cognitive control, like the oft-cited taxi driver advantage in spatial processing (Maguire et al., 2000) and the video gamer advantage in visual processing (Green & Bavelier, 2003), exists because of extended and frequent training. The maintenance of two language systems forces the development of cognitive control mechanisms that can accommodate them. The findings from Experiment 1 fit within this conceptualization. Healthy adult early bilinguals reflect healthy adult monolinguals more than children or patients. Their advantage in cognitive control allows them to maintain two language systems while engaging with their environment like their monolingual peers. Therefore, it follows that they process linguistic conflict like their monolingual
peers, even while suppressing an additional language. The results similarly suggest that linguistic conflict resolution is not affected by nonactive language suppression (e.g., suppression of Catalan when using Spanish) in this population any more than monolinguals are affected by suppressing non-target linguistic tokens (e.g., suppression of words that are semantically or phonologically related to the target). Despite the added task of language suppression, early bilinguals are also advantaged by the early engagement of cognitive control that results from pre-trial conflict. Future studies may explore to what extent this is true. For example, does a higher degree of cognitive control engagement reveal differences between monolinguals and bilinguals? Modulating the type of conflict in the non-linguistic trials may help answer this question. Researchers may also explore whether an increase in the resources dedicated to language suppression changes the effect of conflict control engagement. For example, introducing false cognates, dissimilar linguistic structures, or language switching may alter to what extent language suppression recruits cognitive control, which may in turn alter the effects of pre-trial conflict.

5.6.2 Experiment 2

Experiment 2 is the first study to bring the cognitive-control engagement paradigm to non-native language processing, and as a result, it includes both more limitations and more areas for future research to explore. The first several areas addressed below are methodological in nature, while the last areas are theoretical.

The results of Experiment 2 show that non-near-native linguistic conflict resolution reveal a different processing effect than native and near-native processing of the same stimuli. The nature of the observed effect will be explored more below.
However, the *absence* of a processing effect at Region 3 for L2 participants may reflect a case of shallow processing (e.g., in line with the Shallow Structure Hypothesis), where the L2 learners process semantic and lexical information at first-pass but not syntactic information. In other words, they may not reveal the same slowdown at Region 3 because their shallow parse does not create syntactic conflict between the intransitive verbs’ subcategorization matrices and the linearly (but not structurally) adjacent NP. They are processing the sentences semantically and lexically, which is demonstrated by their high comprehension scores. However, a shallow (syntactically impoverished) parse is only one of several possibilities to explain the absence of the reading time effect at Region 3, and therefore, more research is needed to explore this possibility.

Although previous research with these stimuli has not observed a reading time effect at Region 5, to the best of my knowledge, this is the first study that uses fronted adverbial clauses to study emergent (i.e., non-near-native) bilingual L2 processing. The selection of the range of proficiency levels included in Experiment 2 was in some sense a sample of convenience: all participants in the second experiment were current students at Georgetown University, whether enrolled in an undergraduate, master’s, or doctoral program. The levels were chosen to offer a range of proficiency at a lower level than Jegerski’s (2012) study. Despite the differences, the trends in this sample’s data has offered insight for future research. For example, there is a breadth of proficiency levels between the sample in Experiment 2 and the sample in Jegerski’s (2012) study, as well as a range within the current sample that can be compared cross-sectionally. For example, in a cross-sectional study with more power, might we find that reading times effects shift at increasing proficiency levels, revealing the expected effect at Region 3 for the highest
levels in the study? If so, does cognitive control engagement’s role in facilitating recovery also shift? The current dissertation cannot answer these questions, but presents them as interesting and important questions for future research. For example, the sample in Experiment 2 is too heterogeneous and not large enough to split the group and draw any conclusions. However, given Jegerski’s (2012) findings and the direction of the interaction in Experiment 2, it is worth asking at what stage learners move from processing delays in the spillover Region 5 to more native-like processing at Region 3. To ask this question, the researcher could ask how reading times differ across regions and conditions in a cross-sectional study with more clearly defined and homogenous proficiency groups than this dissertation can provide.

In addition to asking how subject-object ambiguity resolution develops between the stages of proficiency included here (undergraduate and graduate students) and in Jegerski’s (2012) non-native sample (near-native expats), future research can explore if resource depletion occurs when truly low proficiency levels engage in the cognitive-control engagement paradigm (i.e., is the predicted effect found for learners who have not automatized L2 processing to the same extent?). This dissertation’s lowest proficiency participants had completed at least two advanced Spanish courses and were completing a two weeks into a 5-week summer abroad. This population was chosen because of the complexity of the sentences presented in the experiment, but other types of linguistic conflict, like gender and number agreement violations, can be used in future research with a lower proficiency level.

In terms of theoretical questions remaining for future research, while the processing effect at Region 5 is clearly a manifestation of non-native processing, this
dissertation cannot clarify the nature of this processing effect. For example, a garden-path
effect that is shifted one region because of L2 processing difficulties is different
theoretically from a wrap-up effect at the sentence-final region due to a more complex
sentence, even though either would take place at Region 5 for the sentences studied in
this dissertation. Future research could explore this empirically by modulating the length
of the sentence remaining after the disambiguating region. If the effect found here is a
“shifted” garden-path, it would manifest at the region immediately following the matrix
verb. However, an alternative explanation is that the parser slows as the end of a complex
sentence approaches in order to process it globally (i.e., once all the items are revealed,
similar to the proposals in the Shallow Processing Hypothesis). Although this may also
be a delayed garden-path, it entails different implications for L2 processing: is L2
syntactic conflict processed online like in the L1, or does the parser only return to process
complex syntactic relations on an as-needed basis?

Another theoretical question is whether this wrap-up effect at Region 5 is the
result of added complexity due to the nonpreferred referent assignment of the null subject
*pro* (a complexity in Spanish but not English), or if it is non-target-like processing
transferred from the L1 English. In other words, does the low-proficiency L2 parse of
these stimuli permit the postverbal NP to be the direct object of the fronted transitive verb
like the native Spanish parse (a parse that complies with Late Closure but forces the
parser to violate another parsing preference and assign the fronted direct object as the
referent for the subject *pro*)? Or on the other hand, does it include a forced reanalysis at
the matrix verb in Region 4, in line with comparable English stimuli, but reflective of
non-target-like processing? This difference is more difficult to tease apart with these
stimuli, though perhaps highly specific comprehension questions can elicit offline data to begin to answer the question, and the use of a structure that is temporarily ambiguous in Spanish but not in English would also provide the researcher new insight into the L2 processes.

5.6.3 Conclusion

The present dissertation provides a first step in a promising area for future research by showing that the cognitive-control engagement paradigm transfers to populations beyond the monolingual English speakers studied in Hsu & Novick (2016). Early bilinguals and emergent bilinguals also adapt to conflict in cross-task studies, despite the different roles of cognitive control in these populations. While many questions concerning language suppression and conflict adaptation remain, the paradigm (and the confirmation of its transferability) provides a tool for researchers interested in cognitive capacity and language processing, and the results offer insight into two populations that have been understudied in research on these variables.
APPENDIX A: LINGUISTIC STIMULI

Ambigous/Transitive:

1. Después de que comieron el pollo se enfrió de una vez [inmediatamente]
   “After they ate the chicken got cold right away.”

2. Justo cuando aparcaron el carro viejo volvió a fallar
   “Just as they parked the old car died again.”

3. Antes de que cantaran esa canción no era muy conocida
   “Before they sang that song was not very well known.”

4. Cuando mi madre lavaba la ropa quedaba limpiecita [muy limpia]
   “When my mother used to wash the clothes came out nice and clean.”

5. Después de que limpiaron la casa brillaba por todas partes
   “After they cleaned the house was shiny clean all over.”

6. Cuando la artista pintó el cuadro cayó de la pared
   “When the artist painted the painting fell off the wall.”

7. Cinco días después de que bebieron el vino se volvió vinagre.
   “After they drank the wine turned into vinegar.”

   “When the contestant [participant] guessed the answer appeared on the screen.”

9. Después de que corrí el maratón no me parecía tan imposible.
   “After I ran the marathon didn’t seem so impossible to me.”

10. Mientras el maestro tocaba el violín resonaba por todo el salón.
    “While the maestro played the violin resonated throughout the hall.”

11. Cuando la novia descendió la escalera le pareció muy larga [Cuando la chica lava
    los platos se quedan bien sucios]
    “When the girl washes the plates they remain quite dirty.

12. Siempre que mi hermana maneja el carro suena raro.
    “Whenever my sister drives the car makes strange sounds.”

13. Mientras el padre leía la biblia antigua se partió en dos.
    “While the father was reading the ancient bible split in two.”

14. Mientras el novio pagaba el diamante se cayó al piso.
    “While the groom was paying a call the diamond fell to the floor.”
15. Cuando los jardineros terminaron el jardín quedó lleno de flores. 
   “When the gardeners were finished the garden was full of flowers.”

16. Cuando el escultor acabó la obra tenía tres metros de altura. 
   “When the sculptor finished the piece was three meters tall.”

17. Después de que la arquitecta empezó los planos se volvieron más sencillos. 
   “After the architect started the plans became simpler.”

18. Este año cuando gritaron el grito sonó hasta por las montañas. 
   “This year when they shouted el grito was even heard in the mountains.”

19. Cuando Dora compró la comida se le quedó en el mercado. 
   “When Dora bought the food was left behind in the market.”

20. Siempre que la niña salta la cuerda le da por la cabeza. 
   “Whenever the girl jumps the rope hits her in the head.”

Unambiguous/Intransitive (Compare to sentence #__):

1. Después de que hablaron el helado se derritió rápido (#1) 
   “After they talked the ice cream melted quickly.”

2. Justo cuando llegaban el carro viejo se chocó contra un poste de luz (#2) 
   “Just as they arrived the old car crashed against a lightpost.”

3. Antes de que celebraran esa canción no era muy popular (#3) 
   “Before they celebrated that song was not very popular.”

4. Cuando mi madre ayudaba la ropa se quedaban manchada (#4) 
   “When my mother helped the clothes ended up stained.”

5. Después de que salieron la casa se quedó muy limpia (#5) 
   “After they went out the house was left very clean.”

6. Cuando la artista gritó el cuadro cayó y se rompió (#6) 
   “When the artist yelled the painting fell and broke.”

7. Después de que celebraron la leche se volvió agria (#7) 
   “Five days after they celebrated the milk turned sour.”

8. Cuando la participante se rindió la respuesta le pareció obvia (#8) 
   “When the participant gave up the answer seemed obvious.”

9. Después de que llegué el maratón empezó inmediatamente (#9) 
   “After I arrived the marathon started immediately.”
10. Mientras el maestro descansaba el violín resonaba por todo el salón (#10)
   “While the maestro rested the violin resonated throughout the hall.”

11. Cuando la chica cena los platos terminan muy sucios
   “When the girl eats the plates they remain quite dirty.

12. Siempre que mi hermana sale el carro se queda sin gasolina (#12)
   “Whenever my sister goes out the car is left without gas.”

13. Mientras el padre rezaba la biblia antigua se le cayó al piso (#13)
   “While the father prayed the ancient bible fell to the floor.”

14. Mientras el novio salía el diamante se le cayó al piso (#14)
   “While the groom was leaving the diamond fell to the floor.”

15. Cuando los jardineros se fueron el jardín era el más bonito del barrio (#15)
   “When the gardeners left the garden was the prettiest in the neighborhood.”

16. Cuando el escultor volvió la obra estaba instalada en el parque (#16)
   “When the sculptor returned the piece was installed in the park.”

17. Después de que la arquitecta renunció los planos cambiaron mucho
   “After the architect quit the plans changed a lot.”

18. Este año cuando se reunieron el grito sonó por todo el vecindario (#18)
   “This year when they got together el grito was heard throughout the neighborhood.”

19. Cuando Dora salió la comida se le quedó en el horno quemándose (#19)
   “When Dora left the food was left in the oven burning.”

20. Siempre que la niña juega la pelota le da por la cabeza (#20)
   “Whenever the girl plays the ball hits her in the head.”
APPENDIX B: ELICITED IMITATION TASK

REPERTITION TASK
ELICITED IMITATION TASK for L2 SPANISH


*Recording script:*
You are going to hear several sentences in English. After each sentence, there will be a short pause, followed by a tone sound {TONE}. Your task is to try to repeat exactly what you hear. You will be given sufficient time after the tone to repeat the sentence. Repeat as much as you can. Remember, DON'T START REPEATING THE SENTENCE UNTIL YOU HEAR THE TONE SOUND {TONE}. Now let's begin.

We drove to the park
I'll call her tomorrow night
You can buy meat at the butcher shop
My brother just bought a brand new computer
Sometimes they take their dog for a walk in the park
We're going to play volleyball at the gym that I told you about

That was the last English sentence

Now, you are going to hear a number of sentences in Spanish. Once again, after each sentence, there will be a short pause, followed by a tone sound {TONE}. Your task is to try to repeat exactly what you hear in Spanish. You will be given sufficient time after the tone to repeat the sentence. Repeat as much as you can. Remember, DON'T START REPEATING THE SENTENCE UNTIL YOU HEAR THE TONE SOUND {TONE}. Now let's begin.

1. Quiero cortarme el pelo (7 syllables)
2. El libro está en la mesa (7 syllables)
3. El carro lo tiene Pedro (8 syllables)
4. Él se ducha cada mañana (9 syllables)
5. ¿Qué dice usted que va a hacer hoy? (9 syllables)
6. Dudo que sepa manejar muy bien (10 syllables)
7. Las calles de esta ciudad son muy anchas (11 syllables)
8. Puede que llueva mañana todo el día (12 syllables)
9. Las casas son muy bonitas pero caras (12 syllables)
10. Me gustan las películas que acaban bien (12 syllables)
11. Después de cenar me fui a dormir tranquilo (13 syllables)
12. El chico con el que yo salgo es español (13 syllables)
13. Quiero una casa en la que vivan mis animales (14 syllables)
14. A vosotros os fascinan las fiestas grandiosas (14 syllables)
15. Ella ha terminado de pintar su apartamento (14 syllables)
16. El niño al que se le murió el gato está triste (14 syllables)
17. Ella sólo bebe cerveza y no come nada (15 syllables)
18. Me gustaría que el precio de las casas bajara (15 syllables)
19. Cruza a la izquierda y después sigue todo recto (15 syllables)
20. Me gustaría que empezara a hacer más calor pronto (15 syllables)
21. Una amiga mía cuida a los niños de mi vecino (16 syllables)
22. El gato que era negro fue perseguido por el perro (16 syllables)
23. Antes de poder salir él tiene que limpiar su cuarto (16 syllables)
24. La cantidad de personas que fuman ha disminuido (17 syllables)
25. Después de llegar a casa del trabajo tomé la cena (17 syllables)
26. El ladrón al que cogió la policía era famoso (17 syllables)
27. Le pedí a un amigo que me ayudara con la tarea (16 syllables)
28. El examen no fue tan difícil como me habían dicho (17 syllables)
29. ¿Serías tan amable de darme el libro que está en la mesa? (17 syllables)
30. Hay mucha gente que no toma nada para desayuno (17 syllables)
APPENDIX C: SCORING OF ELICITED IMITATION TASK


SCORING PROTOCOL FOR ELICITED IMITATION TASK

Note: The scoring protocol was adapted from the scoring system developed by Ortega, Iwashita, Rabie, and Norris (in preparation).

Table J.1. EIT score 0 descriptor

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Nothing (Silence)</td>
<td>- Manana (10-item 4)</td>
</tr>
<tr>
<td>• Garbled (unintelligible, usually transcribed as XXX)</td>
<td>- El examen que [gibberish] (09-item 28)</td>
</tr>
<tr>
<td>• Minimal repetition, then item abandoned:</td>
<td>- Despues mue- XX tranquilo (01-item 11)</td>
</tr>
<tr>
<td>- Only 1 word repeated</td>
<td>- Tu que sepa a- m- muy bien (12-item 6)</td>
</tr>
<tr>
<td>- Only 1 content word plus function word(s)</td>
<td>- Me gustaria las se se se el XXX (16-item 18)</td>
</tr>
<tr>
<td>- Only function word(s) repeated</td>
<td></td>
</tr>
<tr>
<td>- Only 1 or 2 content words out of order plus extraneous words that weren’t in the original stimulus</td>
<td></td>
</tr>
</tbody>
</table>

Table J.2. EIT score 1 descriptor

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• When only about half of idea units are represented in the string but a lot of important information in the original stimulus is left out; sometimes the resulting meaning is unrelated (or opposed) to stimulus</td>
<td>- Antes de poder seguir (3 sec.) perdio su cuarto (02-item 23)</td>
</tr>
<tr>
<td>• Or when string doesn’t in itself constitute a self-standing sentence with some (related or not to stimulus) meaning (This may happen when only 2 of 3 content words are repeated and no grammatical relation between them is attempted)</td>
<td>- Dudo que sepa ma- tambien (04-item 6)</td>
</tr>
<tr>
<td></td>
<td>- Seria en que el libro esta en la mesa (11-item 29)</td>
</tr>
<tr>
<td></td>
<td>- El gato que eran pelo negro dan something el perro (14-item 22)</td>
</tr>
<tr>
<td></td>
<td>- El ladron que XX la policia famoso (11-item 26)</td>
</tr>
<tr>
<td></td>
<td>- Despues de cenar fue en- tranquilo (03-item 11)</td>
</tr>
<tr>
<td></td>
<td>- Le pendu una amiga que XXX la tarea (01-item 27)</td>
</tr>
<tr>
<td></td>
<td>- La cantidad de personas fumar alguno, alguno (03-item 24)</td>
</tr>
</tbody>
</table>

Table J.3. EIT score 2 descriptor

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• When content of string preserves at least more than half of the idea units in the original stimulus; string is meaningful, and the meaning is close or related</td>
<td>- Despues de cenar me fui a X tranquilo (11-item 12)</td>
</tr>
<tr>
<td></td>
<td>- Ella sola cerveza y no come nada (05-item 17)</td>
</tr>
</tbody>
</table>
to original, but it departs from it in some slight changes in content, which makes content inexact, incomplete, or ambiguous

<table>
<thead>
<tr>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Quieres una casa que viven los alemanes animales (07-item 13)</td>
</tr>
<tr>
<td>- El chico con lo que es algo es español (08-item 12)</td>
</tr>
<tr>
<td>- El chico con yo salgo es muy bien (15-item 12)</td>
</tr>
<tr>
<td>- Después a trabajo tome la cena (16-item 25)</td>
</tr>
</tbody>
</table>

Table J.4. EIT score 3 descriptor

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Original, complete meaning is preserved as in the stimulus. Strings which are ungrammatical can get a 3 score, as long as exact meaning is preserved. Some synonymous substitutions are acceptable:</td>
<td>- Me gustaría el precio de las casas baraja (2 sec.)</td>
</tr>
<tr>
<td>- Anything with or without ‘muy’ (very) should be considered synonymous.</td>
<td>- El niño que se murió cayo esta triste (02-item 16)</td>
</tr>
<tr>
<td>- Substitutions of ‘y’/‘pero’ (and &amp; but) are acceptable.</td>
<td>- [gibberish] se ducha cada mañana (18-item 4)</td>
</tr>
<tr>
<td>• Changes in grammar that don’t affect meaning should be scored as 3.</td>
<td>- Quiero cortar mi pelo (05-item 1)</td>
</tr>
<tr>
<td>(Ambiguous changes in grammar that could be interpreted as meaning changes from a NS perspective should be scored as 2. That is, as a general principle in case of doubt about whether meaning has changed or not, score 2.)</td>
<td>- Las calles de esta ciudad son anchas (13-item 7)</td>
</tr>
<tr>
<td></td>
<td>- El chico que yo salgo es español (06-item 11)</td>
</tr>
<tr>
<td></td>
<td>- El chico con el salgo es español (05-item 11)</td>
</tr>
<tr>
<td></td>
<td>- El examen no fue tan difícil como me han dicho (12-item 28)</td>
</tr>
<tr>
<td></td>
<td>- Las casa son muy bonitas pero caras (07-item 9)</td>
</tr>
<tr>
<td></td>
<td>- Quiero una casa en que viven mis animales (12-item 13)</td>
</tr>
<tr>
<td></td>
<td>- Dudo que sabe a manejar muy bien (11-item 6)</td>
</tr>
<tr>
<td></td>
<td>- Ella he terminado X pintar sus apartamento (11-item 15)</td>
</tr>
</tbody>
</table>

Table J.5. EIT score 4 descriptor

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Exact repetition: String matches stimulus exactly. Both form and meaning are correct without exception or doubt.</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D: RESULTS FOR READING TIMES AT REGIONS 1 AND 2

In Regions 1 and 2, the maximal random effect structure included random intercepts for subject and item, but not random slopes for transitivity, which is supported by the design given that the effect of transitivity would not be detectable until Region 3.

Table D.1.
Output for reading times from early bilingual linear mixed-effects models

<table>
<thead>
<tr>
<th>Region</th>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>Constant Intercept</td>
<td>2.934</td>
<td>.0280</td>
<td>104.817</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>-.0169</td>
<td>.0103</td>
<td>-1.633</td>
<td>.103</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>-.0107</td>
<td>.0105</td>
<td>-1.020</td>
<td>.308</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>.0220</td>
<td>.0147</td>
<td>1.497</td>
<td>.135</td>
</tr>
<tr>
<td>Region 2</td>
<td>Constant Intercept</td>
<td>2.553</td>
<td>.0347</td>
<td>73.778</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>.0035</td>
<td>.0164</td>
<td>.214</td>
<td>.831</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>-.0008</td>
<td>.0166</td>
<td>-.048</td>
<td>.962</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>.0059</td>
<td>.0233</td>
<td>.251</td>
<td>.802</td>
</tr>
</tbody>
</table>

Table D.2.
Output for reading times from emergent bilingual linear mixed-effects models

<table>
<thead>
<tr>
<th>Region</th>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>Constant Intercept</td>
<td>3.227</td>
<td>.1323</td>
<td>24.778</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>.0094</td>
<td>.0559</td>
<td>.169</td>
<td>.8661</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.0119</td>
<td>.0529</td>
<td>.226</td>
<td>.8216</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0028</td>
<td>.0013</td>
<td>-2.097</td>
<td>.0423*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.0207</td>
<td>.0710</td>
<td>-.291</td>
<td>.7708</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0002</td>
<td>.0006</td>
<td>-.268</td>
<td>.7890</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0001</td>
<td>.0005</td>
<td>-.107</td>
<td>.9149</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>.0001</td>
<td>.0007</td>
<td>.208</td>
<td>.8350</td>
</tr>
<tr>
<td>Region 2</td>
<td>Constant Intercept</td>
<td>2.983</td>
<td>.1552</td>
<td>19.221</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td></td>
<td>Transitivity</td>
<td>.0197</td>
<td>.0804</td>
<td>.246</td>
<td>.8060</td>
</tr>
<tr>
<td></td>
<td>Incongruency</td>
<td>.0884</td>
<td>.0948</td>
<td>.932</td>
<td>.3514</td>
</tr>
<tr>
<td></td>
<td>Proficiency</td>
<td>-.0039</td>
<td>.0016</td>
<td>-2.420</td>
<td>.0198*</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Incongruency</td>
<td>-.0230</td>
<td>.1271</td>
<td>-.181</td>
<td>.8564</td>
</tr>
<tr>
<td></td>
<td>Transitivity*Proficiency</td>
<td>-.0002</td>
<td>.0008</td>
<td>-.295</td>
<td>.7679</td>
</tr>
<tr>
<td></td>
<td>Incongruency*Proficiency</td>
<td>-.0006</td>
<td>.0009</td>
<td>-.622</td>
<td>.5341</td>
</tr>
<tr>
<td></td>
<td>Transitivity<em>Incongruency</em>Proficiency</td>
<td>-.0001</td>
<td>.0013</td>
<td>-.063</td>
<td>.9499</td>
</tr>
</tbody>
</table>
REFERENCES


Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in 
second and native languages. *Quarterly Journal of Experimental Psychology: 
Section A, 50*(1), 119–148.


Bridging language and attention: Brain basis of the impact of bilingualism on 

general comprehension skill. *Journal of Experimental Psychology: Learning, 

Gollan, T. H., Starr, J., & Ferreira, V. S. (2015). More than use it or lose it: The number-
of-speakers effect on heritage language proficiency. *Psychonomic Bulletin & 

Strategic Control of Activation of Responses. *Journal of Experimental Psychology: 
General, 121*(4), 480–506.


Mapping form to meaning in two languages. In J. F. Kroll (Ed.), *Handbook of
in one mind: What bilinguals tell us about language processing and its
consequences for cognition. In: B. Ross (Ed.), *The Psychology of Learning and
and proficient bilinguals. In I. Kecskes & L. Albertazzi (Eds.), *Cognitive aspects of
bilingualism* (pp. 237–269). New York: Springer
and Bilinguals: The Development of Lexical and Conceptual Representations. In
C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language
Conflict Adaptation Similarities in Children and Adults. *Developmental
Neuropsychology, 27*, 343-357.


Paap, K., & Liu, Y. (2014). Conflict resolution in sentence processing is the same for bilinguals and monolinguals: The role of confirmation bias in testing for bilingual advantages. *Journal of Neurolinguistics, 27* (1), 50–74.


