NEUROCOGNITIVE CORRELATES OF SYNTACTIC PROCESSING IN CHILD AND ADULT SECOND LANGUAGE LEARNERS

A Thesis in
Psychology
by
Fatemeh Abdollahi

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Master of Science

May 2016
The thesis of Fatemeh Abdollahi was reviewed and approved* by the following:

Janet van Hell  
Professor of Psychology and Linguistics  
Thesis Advisor

Ping Li  
Professor of Psychology and Linguistics

Lynn Liben  
Distinguished Professor of Psychology

Paola Dussias  
Professor of Spanish, Linguistics and Psychology

Melvin Mark  
Professor of Psychology  
Head of the Department of Psychology

*Signatures are on file in the Graduate School
ABSTRACT

Second languages (L2) are taught in millions of classrooms worldwide, and to almost all age groups. In spite of great efforts made to incorporate L2 teaching in classrooms, there are still fundamental questions of how the L2 is processed throughout learning, and what factors may affect this. The Competition and Proficiency-based Models describe how adult second language (L2) learning occurs. The Competition Model emphasizes the role of transfer of knowledge from the first language (L1) to L2 where grammatical structures are similar between languages, and competition between L1 and L2 in cases where grammatical structures are dissimilar. The Proficiency-based Model argues that as proficiency in the L2 changes, neurocognitive patterns of L2 processing will change, and learners will show increased sensitivity to L2 morphosyntax with increased proficiency.

Neurocognitive and psycholinguistic studies on L2 learning have largely focused on highly proficient adult learners. Child L2 learners may differ fundamentally from adult L2 learners as children and adults are at different developmental stages, in terms of L1 development, as well as cognitively (e.g., smaller working memory span in children). In the present study, critical questions in second language acquisition were explored through electrophysiological and behavioral methods: 1) How is L2 learning influenced by the (dis)similarity between L1 and L2 structures, in children and adults? 2) How do individual differences in cognitive and affective variables associate with morphosyntactic processing, in children and in adults?

Two groups of native English speaking participants at intermediate levels of proficiency in L2 Spanish completed a grammaticality judgment task to assess learned knowledge of Spanish morphosyntax. Adult learners were 18 years and older (late learners) and child learners were ages 9-11 years old (early learners). L2 learners were presented with Spanish sentences that have
similar morphosyntactic structures in L1 and L2, have dissimilar morphosyntactic structures in L1 and L2, or have a unique morphosyntactic structure in the L2 that is not present in the L1. Building on the syntactic violation paradigm that is often used in ERP research, sentences were grammatically correct or incorrect. Participants read the sentences while electroencephalography (EEG) was recorded, and made a grammaticality judgment after having read the sentence. The manipulation of grammatically similar, dissimilar, or unique structures tested the predictions of the Competition Model; proficiency was held constant, and results were compared to ERP predictions across the L2-learning proficiency span (Steinhauer, White and Drury, 2009). In addition to the sentence reading experiment in the L2, sentence processing in the L1 was measured, and tasks were administered that measured working memory, executive functions, L1 proficiency, and attitude and motivation to learn an L2. Using correlations, I explored how variability in performance on these factors may be associated with sentence processing, in child and adult L2 learners.

Adult learners demonstrated sensitivity to grammaticality through a P600 in the L2, but in only the dissimilar condition, suggesting both the Competition Model and Proficiency-based Model may not best represent L2 learning. In the L1 they demonstrated sensitivity as an AN+P600 in all three conditions. Measures of response dominance showed large variability in processing of both L1 and L2, within the learner group. This suggests that grand mean waveforms of traditional ERP can be misleading for actual processing patterns. In correlations of individual difference factors and language processing measures, only motivation and second language proficiency correlated, with increased proficiency with increased motivation. Child learners demonstrated no sensitivity to grammaticality in the L1. In the L2 there were no significant sensitivities, but emerging P600 patterns were present in processing sentences of
structures similar between the L1-L2, and unique to the L2, though differential in topography from adult P600 processing. Within group variability was observed in response dominance, in both the L1 and L2. In correlations of individual difference factors and language processing measures, no factors correlated.

Ultimately we conclude that 1) L2 learning is differentially influenced by (dis)similarity between L1-L2 in children (though only preliminary findings due to small sample), and adults. While children show emerging sensitivity based on (dis)similarity, adults did not show this effect. Again, due to large within group variation, we cannot conclude that these findings are not due to small child sample size. 2) Individual variation in working memory, cognitive control, L1 and L2 proficiency, and attitude/motivation to learn an L2 largely did not associate with language processing measures. Only motivation and L2 proficiency were found to have a positive association for adult L2 learners.
# TABLE OF CONTENTS

LIST OF FIGURES.......................................................................................................................... vii

LIST OF TABLES .............................................................................................................................. x

Chapter 1: INTRODUCTION ............................................................................................................. 1
  The role of age of acquisition in L2 learning ................................................................................ 2
  Theoretical models of L2 morphosyntactic processing .............................................................. 6
    Competition Model ..................................................................................................................... 7
    Proficiency-based Perspective .................................................................................................. 13
  The role of individual difference variables in L2 morphosyntactic processing ...................... 20
    Working memory ....................................................................................................................... 20
    Executive Functions and Cognitive Control ........................................................................... 23
    L1 Proficiency ......................................................................................................................... 24
    Attitude and Motivation ......................................................................................................... 25
  The Present Study ..................................................................................................................... 26

Chapter 2: METHODS ....................................................................................................................... 28
  Participants ................................................................................................................................. 28
  Materials .................................................................................................................................. 29
  Procedure .................................................................................................................................. 30
  Individual difference measures ................................................................................................. 34
    Cognitive tasks ......................................................................................................................... 35
    Proficiency measures ............................................................................................................. 38
    Attitude/Motivation to Learn an L2 ....................................................................................... 38
  Neurocognitive method: ERPs .................................................................................................. 39

Chapter 3: ANALYSIS & RESULTS ................................................................................................. 41
  Analysis ..................................................................................................................................... 41
  Results ...................................................................................................................................... 44
    Adult ...................................................................................................................................... 44
    Child ...................................................................................................................................... 72

Chapter 4: DISCUSSION .................................................................................................................... 93
  Adult Language Processing ....................................................................................................... 93
  Child Language Processing ....................................................................................................... 110

Chapter 5: CONCLUSION & FUTURE DIRECTIONS ..................................................................... 114

REFERENCES ................................................................................................................................. 116

Appendix ......................................................................................................................................... 125
  Appendix A: Adult Learner RDI in L2 Processing ................................................................. 125
  Appendix B: Adult Learner RDI in L1 Processing ................................................................. 136
  Appendix C: Child Learner RDI in L2 Processing ................................................................. 141
  Appendix D: Child Learner RDI in L1 Processing ................................................................. 146
LIST OF FIGURES

Figure 1: Timeline of stimulus presentation during primary task. Participants pressed button to indicate response after question mark.................................................................44

Figure 2: Accuracy by condition in L2 Spanish to similar (Tense), unique (Gender) and dissimilar (Number) conditions, for adult L2 learners. Standard error is shown through error bars. Accuracy is expressed in percentage. Red bars are accuracy to ungrammaticality, green indicate accuracy to grammatical sentences.................................55

Figure 3: Accuracy by condition in L1 English to Tense, Reflexive and Subject-Verb conditions, for adult L2 learners. Standard error is shown through error bars. Accuracy is expressed in percentage. Red bars are accuracy to ungrammaticality, green indicate accuracy to grammatical sentences.................................................................56

Figure 4: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of L2 Spanish grammatical (green) and ungrammatical (dashed red) sentences. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10μV........................................58

Figure 5: Grand mean waveforms contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences across the three L2 Spanish conditions: (5A) Tense (similar), (5B) Determiner Number (dissimilar), and (5C) Determiner Gender (unique) conditions........................................................................................................59-61

Figures 6, 7 and 8: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 6) Tense (similar), 7) Determiner Gender (unique), and 8) Determiner Number (dissimilar) conditions. The solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered those with a positive RDI.........................................................65-66

Figures 9, 10 and 11: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 9) Tense (similar), 10) Determiner Gender (unique), and 11) Determiner Number (dissimilar) conditions........................................................................................................68-69

Figure 12: Grand mean waveforms contrasting amplitude of processing of L1 English grammatical (green) and ungrammatical (dashed red) sentences........................................71

Figure 13: Grand mean waveforms contrasting amplitude of processing across three L1 English conditions: (13A) Tense, (13B) Reflexive, and (13C) Subject-Verb conditions........................................................................................................72-74
Figures 14, 15 and 16: Scatterplots presenting effect magnitudes in the N400 and P600 windows for individuals across the three L1 English conditions: 14) Tense, 15) Reflexive, and 16) Subject-Verb conditions.................................................................77-78

Figure 17: Accuracy by condition in L2 Spanish, Similar (Tense), Unique (Gender) and Dissimilar (Number) conditions, for child L2 learners..................................................84

Figure 18: Accuracy by condition in L1 English Tense, Reflexive and Subject-Verb conditions, for child L2 learners.................................................................85

Figure 19: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of L2 Spanish grammatical (green) and ungrammatical (dashed red) sentences. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16μV...............................86

Figure 20: Grand mean waveform from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences across the three L2 Spanish conditions: (20A) Tense (similar), (20B) Determiner Number (dissimilar), and (20C) Determiner Gender (unique) conditions........................................87-89

Figures 21, 22 & 23: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 21) Tense (similar), 22) Determiner Gender (unique), & 23) Determiner Number (dissimilar) conditions..................................................................................................92-93

Figures 24, 25 & 26: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 24) Tense (similar), 25) Determiner Gender (unique), & 26) Determiner Number (dissimilar) conditions..................................................................................................94

Figure 27: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of L1 English grammatical (green) and ungrammatical (dashed red) sentences.................................................................96

Figure 28: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences across three L1 English conditions: (28A) Tense, (28B) Reflexive, and (28C) Subject-Verb conditions..................................................................................................97-99

Figures 29, 30 & 31: Scatterplots presenting effect magnitudes in the N400 and P600 windows for individuals across the three L1 English conditions: 29) Tense, 30) Reflexive, & 31) Subject-Verb conditions..................................................................................................102-103

Figure 32: Grand mean waveform from nine representative electrodes contrasting grammatical (green) and ungrammatical (dashed red) sentences, for those individuals
demonstrating negative RDI, across three L2 Spanish conditions: (32A) Tense (similar), (32B) Determiner Number (dissimilar), and (32C) Determiner Gender (unique) conditions.................................................................139-141

**Figure 33:** Grand mean waveform from nine representative electrodes contrasting grammatical (green) and ungrammatical (dashed red) sentences, for those individuals demonstrating positive RDI, across three L2 Spanish conditions: (33A) Tense (similar), (33B) Determiner Number (dissimilar), and (33C) Determiner Gender (unique) conditions.................................................................144-146

**Figures 34 & 35:** Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences, for those individuals demonstrating negative RDI (Figure 34) and positive RDI (Figure 35), across three L1 English conditions: Tense, Reflexive and Subject-Verb conditions........................................................................................................................................147-148

**Figure 36:** Grand mean waveform from nine representative electrodes contrasting grammatical (green) and ungrammatical (dashed red) sentences, for those child learners demonstrating negative RDI, across three L2 Spanish conditions: Tense (similar), Determiner Number (dissimilar), and Determiner Gender (unique) conditions.................................154

**Figure 37:** Grand mean waveform from nine representative electrodes contrasting grammatical and ungrammatical sentences, for those child learners demonstrating positive RDI, across three Spanish conditions: Tense (similar), Determiner Number (dissimilar), and Determiner Gender (unique) conditions..........................................................................................................................156

**Figures 38 & 39:** Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences, for those child learners demonstrating negative RDI (Figure 38) and positive RDI (Figure 39), across three L1 English conditions: Tense, Reflexive and Subject-Verb conditions........................................................................................................................................158-159
LIST OF TABLES

Table 1: Examples of structure types tested in Spanish. *indicates ungrammaticality........41

Table 2: Examples of structure types tested in English. *indicates ungrammaticality........41

Table 3: Mean accuracy (%) in correctly identifying violations in grammaticality, broken down by grammatical and ungrammatical sentences in the L2 and L1, for each condition, for adult L2 learners. SD=Standard Deviation.................................................................56

Table 4: F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for L2 Spanish language processing in the 300-500 ms and 600-800 ms windows. Degrees of freedom are in parentheses. *.05<p<.1; **p<.05; ***p<.01......................................................62

Table 5: F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for L1 English language processing in the 300-500 ms and 500-800 ms windows. Degrees of freedom are in parentheses. *.05<p<.1; **p<.05; ***p<.01......................................................75

Table 6: Mean individual difference scores, standard deviations and score ranges for variables related to Sensitivity to Grammaticality in the L2 and in the L1, Age of L2 Acquisition, L1 Exposure, L2 Exposure, and Self-rated L2 Proficiency.............................................80

Table 7: Mean individual difference scores, standard deviations and score ranges for performance on the Rey, O-Span, TVIP, PPVT, ANT, and Simon Tasks.................................................81

Table 8: Correlations between L1 Proficiency, L2 Proficiency, Attitude, Motivation, Working memory, Cognitive Control, RDI in L2, and RMI in L2.........................................................82

Table 9: Mean accuracy in correctly identifying violations in grammaticality, broken down by grammatical and ungrammatical sentences in the L2 and L1. SD=Standard Deviation 84

Table 10: F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for child L2 Spanish language processing in the 300-500 ms and 600-800 ms windows. Degrees of freedom are in parentheses.................................................................90

Table 11: F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for child L1 English language processing in the 300-500 ms and 500-800 ms windows. Degrees of freedom are in parentheses. *.05<p<.1; **p<.05; ***p<.01.................................100

Table 12: Findings from 6 studies on L2 learning contrasting findings based on cross-linguistic similarity of tense marking structure, and proficiency.................................107

Table 13: Findings from 11 studies on L2 learning contrasting findings based on cross-linguistic similarity of number agreement, and proficiency.................................109
Table 14: Findings from 9 studies on L2 learning contrasting findings based on cross-linguistic similarity of gender agreement, and proficiency

Table 15: F-statistics from ANOVA on mean amplitude measures for participants of negative RDI in the three L2 Spanish conditions in the 300-500 ms and 600-800 ms windows. Degrees of freedom are reported in parentheses. Gram. = Grammaticality; Cond. = Condition; Elec. = Electrode; Ant = Anteriority; Hem. = Hemisphere. *.05<p<.1; ** p<.05; *** p<.01

Table 16: F-statistics from ANOVA on mean amplitude measures for participants of positive RDI in the three Spanish conditions in the 300-500 ms and 600-800 ms windows. Degrees of freedom are reported in parentheses. Gram. = Grammaticality; Cond. = Condition; Elec. = Electrode; Ant = Anteriority; Hem. = Hemisphere. *.05<p<.1; ** p<.05; *** p<.01

Table 17: F-statistics from ANOVA on mean amplitude measures for child participants of negative RDI in the three Spanish conditions in the 300-500 ms and 600-800 ms windows. Degrees of freedom are reported in parentheses. *.05<p<.1; ** p<.05; *** p<.01

Table 18: F-statistics from ANOVA on mean amplitude measures for child participants of positive RDI in the three Spanish conditions in the 300-500 ms and 600-800 ms windows. Degrees of freedom are reported in parentheses. *.05<p<.1; ** p<.05; *** p<.01
Chapter 1: INTRODUCTION

The ability to speak multiple languages is of increasing importance in social and professional contexts. While there has been an increase in the study of the neural and cognitive processes involved in second language (L2) learning, little research has focused on young classroom L2 learners (Pufahl & Rhodes, 2011; for more information on dual language learners, see Paradis, Genesee & Crago, 2011; Sandhofer & Uchikoshi, 2013), who constitute a large portion of L2 learners. Neurocognitive and psycholinguistic studies on L2 learning have largely focused on adult learners (for a review, see van Hell & Tokowicz, 2010), who may greatly differ from their child counterparts in rate of learning and cognitive processes utilized in L2 learning. These differences may result from adult and child L2 learners being at different developmental stages, not only in terms of first language (L1) development and the level of entrenchment of their L1 – that is, how automatic L1 retrieval is or how engrained L1 syntax is, but also cognitively (e.g., smaller working memory span in children) and neurologically (e.g., white and grey matter development). Moreover, within a given age group, individuals differ in L1 proficiency, cognitive resources, and motivation/attitude towards L2 learning, and these individual differences potentially affect the rate and success of L2 learning.

The present study examined morphosyntactic processing of adult and child L2 learners in L2 and L1, investigating within group variation, as well as between group differences. In an Event-Related Potential (ERP) study, fourth/fifth grade and adult L2 Spanish learners with similar amounts of L2 Spanish instruction read L2 Spanish sentences that contained morphosyntactic structures that varied in similarity to structures in their L1 English. L1 and L2 proficiency, executive functions, working memory, and
attitude and motivation towards learning an L2 were also measured. This experimental
design allowed us to study the effects of learner-related factors, (i.e., age of acquisition
(AoA)) while controlling for the amount of language exposure, executive control, L1 and
L2 proficiency, and attitude/motivation on L2 sentence processing, in sentences that
contain L2 grammatical structures that vary in terms of similarity to grammatical
structures in the L1. Measures of performance (both ERP and behavioral) can inform
debates on L2 morphosyntactic processing in younger and older L2 learners at an
intermediate level of L2 proficiency. To gain insight into their L1 morphosyntactic
processing, and the extent to which electrophysiological signatures of morphosyntactic
processing in L2 are comparable to those in L1, children and adults also read sentences in
L1 while ERPs were recorded. In the Introduction I will discuss the theoretical
background motivating this study, followed by the methodology employed.

The role of age of acquisition in L2 learning

In the discussion of whether children or adults are “better” L2 learners, the critical
period hypothesis has been a long-standing and influential hypothesis of L2 learning
trajectory. This hypothesis (Lenneberg, 1967; Long, 1990) claims that, due to constraints
in plasticity as a result of maturational developments, adult learners are unable to acquire
a second language at native-like levels of proficiency.

In a classical study, Johnson and Newport (1989) tested 46 adult native speakers
of Chinese and Korean on their knowledge of twelve English grammatical structures. The
speakers varied on the age at which they immigrated to the U.S. (AoA), either before
puberty (defined as ages 3-15 years), or after puberty (defined as ages 17-39 years).
Participants ranged in length of exposure to English in the U.S. from 3 to 26 years. They
were evaluated through a grammaticality judgment task and their performance was compared to that of native English speaker controls. Participants who had entered the U.S. before age 7 trended like native speakers, but performance of the participants who had entered the U.S. after the age of 7 decreased as AoA increased. Variables such as years of formal foreign language instruction and years of exposure to the L2 did not correlate with performance. Measurement of attitude toward the L2 and United States did correlate positively with performance, with improved grammatical performance as positive affect increased. Johnson and Newport interpreted these findings as evidence for the critical period hypothesis, citing a decline in L2 performance after puberty.

This and similar studies have found that as AoA has increased, language proficiency often decreased (e.g., Birdsong & Molis, 2001; Weber-Fox & Neville, 1996). However, in these studies there has often been a confound of AoA with L2 proficiency or the amount of L2 exposure, which potentially presents an inaccurate picture of the effects of AoA on the rate and success of L2 learning. In Johnson and Newport (1989), the study most frequently cited to support the critical period, exposure to the L2, as well as other confounding factors, were not systematically controlled. The range of number of years of L2 exposure of the participants was large (3-26 years), and participants with less exposure to the L2 were all in the late learner groups, suggesting that the observed effects that led to the conclusion of a critical period of L2 learning were driven by differences in L2 exposure rather than AoA. In fact, Bialystok and Hakuta's (1994) re-examination of data from Johnson and Newport (1989) found that it was only after an AoA of 20 years (far past the “critical period”) that a steep decline in accuracy of grammaticality judgments was observed. Likewise, Flege, Yeni-Komshian and Liu (1999) tested 240
native Korean speakers of varying ages of acquisition for proficiency and accentedness in their L2 (English), and found that the level of accentedness increased as AoA increased. However, once L2 learners were separated by factors other than AoA (e.g., years of education in the U.S., years living in the U.S.), differences in morphosyntactic performance that had been attributed to AoA were no longer significant.

While it is largely accepted that there is a critical period for L1 learning (e.g., cases such as Genie (Curtiss, 1977)), successful cases of late L2 language acquisition have been long ignored in the L2 literature (Marinova-Todd, Marshall & Snow, 2000). Significant numbers of late learners who perform like native speakers on various linguistic tasks have been found (Bongaerts, 1999; Cranshaw, 1997; Ioup, Boustagui, Tigi & Moselle, 1994; van Wuijtswinkel, 1994). Phonology has been regarded as the only aspect of L2 learning that appears to have a true critical period, though this is also contested by cases of native-like accentedness in late learners (Bahrick, Hall, Goggin, Bahrick & Berger, 1994; MacKay, Flege & Imai, 2006). These cases of native speakers who have attained native-like competence in the L2 are important to this debate. While at present more researchers agree that there is an age-related decline in the rate and success of L2 learning, the existence of highly successful (very) late L2 learners indicates that there cannot be a set time (hard-wired in the brain) after which native-like competence in the L2 can no longer be achieved. This also suggests that, in order to examine the variability in the trajectory of L2 learning, more insight into the role of learner-related factors is needed.

Very few studies have attempted to control for effects of AoA and L2 exposure (Marinova-Todd et al., 2000; Muñoz, 2006; Snow & Hoefnagel-Höhle, 1978). One of
these studies is the Barcelona Age Factor project by Muñoz (2006). Data from 530
Spanish learners of English was examined on measures of L2 reading (detection of
textual cohesion in a narrative), writing (dictation), listening (dictation as well as
reception of information in an oral interview), speaking (production in an oral interview),
and general L2 proficiency. Learners were from four age groups (ages 8, 11, 14 and 18+
years old), with rate of exposure controlled (tested after 200 hours and then again after
416 hours of exposure). Muñoz found a graded effect for age, with 18+ year-old learners
significantly outperforming 14 year-old L2 learners, who outperformed the 11 year-old
L2 learners, who in turn outperformed the 8 year-old learners in all assessment measures
except for listening and speaking, on which there were no significant differences between
groups. These trends held across both testing points. These findings demonstrate that
comparable performance at listening and speaking can occur across a variety of AoAs.
More importantly, the older L2 learners demonstrated faster and more successful learning
than the younger learners. Similar advantages of older over younger learners in the
learning of syntax, morphology, and pronunciation have been observed by Snow and
Hoefnagel-Höhle (1978; see also Marinova-Todd et al., 2000).

The observed advantage of older over younger learners in Muñoz (2006) and
Snow and Hoefnagel-Höhle (1978) brings up an additional factor, and potential
confound, in studying the role of AoA in L2 processing: learner-related variability in
cognitive capacities and in level of L1 development. Measures of individual differences
in cognitive processing and L1 proficiency are therefore needed to see how much of a
role these play in AoA effects on L2 learning.
The studies described above, while influential in investigating the question of how AoA impacts L2 learning, show a need for research in which confounds of exposure and proficiency are controlled for in exploring how AoA impacts L2 learning (cf. Hernandez & Li, 2007). In the current project we explored the role of AoA (by studying two L2 learner groups: adults and children\(^1\)) on L2 learning, as measured through L2 sentence processing, while controlling for amount of exposure to the L2 (by testing both groups at intermediate levels of learning). We also investigated how variability in ERP responses during L2 processing manifest in each individual, as measured through two means: 1) calculations of Response Dominance Indices (RDI) and 2) Response Magnitude Indices (RMI). Moreover, we investigated how variation in learner-related factors affects L2 sentence processing, through including measures of working memory capacity as well as other executive functions, L1 proficiency, and attitude/motivation to learn an L2.

In the next section, I will discuss two theoretical perspectives on L2 morphosyntactic processing that formed the theoretical framework for my Master's thesis research, followed by a more detailed discussion of the potential role of executive functions, working memory, L1 proficiency, and attitude/motivation to learn an L2 on L2 processing.

**Theoretical models of L2 morphosyntactic processing**

In the past decades, several theoretical views on morphological processing in the L2 have been proposed. In the present thesis, I focus specifically on two perspectives that emphasize transfer between the L1 and L2 (i.e., the Competition Model) and the impact of L2 proficiency, in particular with respect to neurocognitive signatures (Steinhauer et

---

\(^1\) Data from a pilot group of 7 child participants is included in the current study. In testing during
al., 2009); for models emphasizing qualitative differences between L2 and L1 processing, and L2 and L1 speakers, see Clahsen and Felser (2006) and Ullman (2001; Declarative/procedural model).

**Competition Model**

The Competition Model (MacWhinney & Bates, 1989), a theory of L2 morphosyntactic acquisition and processing, argues that morphosyntactic learning in the L2 occurs with great influence from the morphosyntactic patterns of the L1. This model builds on the idea of transfer between L1 and L2. The model proposes that in language learning there are a variety of morphosyntactic cues that enable mapping of linguistic form with meaning (e.g., word order enables assignment of agent and patient differentially based on cues, and can manifest differentially across languages) that are in competition with one another while a sentence is being processed. These cues will have differential weights based cue validity in the language in question, calculated as a composite measure of the type, availability and reliability of the cue in question (Li & MacWhinney, 2013; MacWhinney, 2005). When L2 learning occurs, new cue validity must be learned, as the weights of cues from the L1 and L2 are never exactly the same. This can result in transfer from the L1 to the L2, or in competition between the L1 and L2, based on similarities in cue validity. Transfer of a structure from the L1 to the L2 occurs when the two structures are weighed similarly between the L1 and L2. Competition occurs when there are two cues that are in direct competition with one another (i.e., have different weights in the L1 versus the L2). A third pattern is formed by structures that are unique in the L2 and have no corresponding structure in the L1 (e.g., gender marking exists in Spanish but not in English). For unique structures there is no competition (nor transfer) between languages, as the structure is only present in the L2.
The Competition Model presents similarity between two languages as a factor that influences rate of language acquisition. Cues that are similar across the two languages transfer most easily, and are expected to be learned most quickly by an L2 learner. Unique structures are also learned relatively easily, as there is no competition between the two languages. Structures that are dissimilar between the two languages (i.e., competing cues) are acquired with great difficulty.

A key concept in the Competition Model is entrenchment of the L1, through which concepts and structures in the L1 become solidified and predictable through reliable and recurrent use. This entrenchment allows for more automatized L1 use, as concepts can quickly be retrieved and utilized. However, entrenchment can be problematic for L2 learning, as ways in which structures and concepts are used in the L2 will almost always differ to some extent from those entrenched in the L1. Thus, in L2 learning entrenchment from the L1 must be overcome in order to correctly learn the L2. This is why structures that are similarly represented between the L1 and L2 will be learned most quickly, as these can utilize prior entrenchment of structures in the L1. Those which are unique to the L2 should be acquired relatively easily as well, as they do not have competitors entrenched in the L1 – these can develop their own unique representations in the L2 system. However, those structures that are represented dissimilarly between the L1 and L2 encounter the most difficulty, because retrieval of how L1 structures are used is automatic from the entrenched system. It will require time to overcome the competition from the entrenched L1 structure and different L2 use and map new representations of the L2 use onto the structure.
The Competition Model is based on adult L2 learners, and does not make specific predictions for child L2 learners. It can be hypothesized, however, that the level to which L1 entrenchment aids or hinders L2 learning should not be the same for adults and children. As an L1 is developed, entrenchment grows deeper. Thus, in adults representations of structures would be expected to be very well-established, whereas in children who are still developing their L1, the concepts are not as entrenched, and transfer and competition between L1 and L2 structures will be less influential on L2 learning patterns. If this hypothesis is correct, child L2 learners would not be affected to the same degree as adult L2 learners by the similarity or dissimilarity between grammatical structures in the L1 and the L2.

The predictions of the Competition Model have been tested in many studies, most of which focused on adult L2 learners, and the impact of L1-L2 similarity and dissimilarity of grammatical structures (transfer vs. competition); to the best of my knowledge, the role of entrenchment in L1-L2 transfer and competition has not been studied (but are studied in this thesis) in child and adult L2 learners. In adult L2 learners, there is accumulating evidence that the similarity or dissimilarity between languages plays a role in the rate and success of L2 learning, although the research is not always conclusive (for reviews, see MacWhinney, 2012; Tolento & Tokowicz, 2011; van Hell & Tokowicz, 2010). Below, I discuss the studies that tested the Competition Model using the ERP technique, as this is the technique I used in my Master's thesis research.

Tokowicz and MacWhinney (2005) presented adult English-speaking students who were in their first, second, third or fourth semester of learning L2 Spanish with structures similar between the L1 and L2 (e.g., auxiliary omission “Su abuela
His grandmother *cooking/cooks very well”), dissimilar between the L1 and L2 (e.g., determiner number “*El/Los niños están jugando/*The-SING/the-PL boys are playing”), and unique to the L2 (e.g., determiner gender “Ellos fueron a *un/una fiesta/They went to *a-MASC/a-FEM party”), while their brain activity was recorded using the Event-Related Potential (ERP) technique. The learners were asked to make a grammaticality judgment after reading each sentence. L2 Spanish learners demonstrated P600s in response to violations for the L1-L2 similar structures and the L2-unique structures, but not for the dissimilar structures, even after four semesters of instruction. Behavioral accuracy was found to be near chance for all three structures, even while sensitivity to structures was shown through ERP measures (to the similar and unique constructions). These results show that there is L1 to L2 transfer as suggested by the Competition Model, in that similar L1-L2 structures as well as those unique to the L2 are learned by these L2 learners, whereas cues that contrast between the L1-L2 were not yet acquired at this level of proficiency. However, we do know from studies of highly proficient non-native L2 speakers (Ioup et al., 1994) that these dissimilar structures are ultimately learned at a later stage of L2 learning. The findings of this study suggest that higher proficiency levels are necessary for these structures to be acquired.

Using a longitudinal design, Osterhout, McLaughlin, Pitkänen, Frenck-Mestre and Molinaro (2006) examined adult English L2 learners of French at 1 month, 4 months, and 8 months of classroom instruction. When tested using both ERP and behavioral measures, they found that learners displayed sensitivity to the structure similar between French and English at even 1 month of L2 instruction, but still did not display sensitivity to the
structure dissimilar between languages after 8 months of L2 instruction. This again supports the Competition Model in that structures that are more similar between two languages will be learned the most quickly, and in a native-like manner.

Sabourin and Stowe (2008) used ERP and behavioral measures to test performance of advanced Romance language-Dutch and German-Dutch L2 learners on structures of verbal domain dependency and grammatical gender processing in the L2 (Dutch). These L2 learners had a minimum of 3 years immersion in the Netherlands, and an average of 9 years of Dutch instruction. In both groups verbal domain dependency was a similar structure, while grammatical gender processing was a dissimilar structure for Romance language-Dutch L2 learners. An attenuated P600 was reported for both groups with the similar structure, whereas Romance-Dutch learners did not show this sensitivity on the dissimilar structure (grammatical gender processing). This again suggests that based on the cross-linguistic differences between L1-L2, L2 learners will show differential sensitivities to the same structures, depending on the similarity to their L1.

Using ERP and behavioral measures, Foucart and Frenck-Mestre (2011) investigated the effects of L1-L2 similarity on the performance of advanced German-French L2 learners when processing violations of grammatical gender in the L2 (French). Where manipulations of gender were similarly expressed between the L1-L2, the advanced German-French L2 learners demonstrated the same P600 ERP pattern as native French control participants. However, no effect other than a small early negativity was found in the advanced learners when the structures were dissimilar between L1-L2, while native speakers showed the expected P600 response. Furthermore, in the advanced learners' responses to similar constructions, while all participants were sensitive to
determiner-noun gender mismatch violations in which the nouns' gender was the same in German and French (e.g., French: *la montre* (feminine); German: *die Uhr* (feminine) “the watch”), only half of the participants were sensitive to violations of a determiner-noun gender mismatch when the noun was not matched in gender across German and French (e.g., French: *la clef* (feminine); German: *der Schlüssel* (masculine) “the key”). Possible differences in proficiency are suggested as a cause, with those who are more proficient being more sensitive to determiner-noun mismatches that have a gender discrepancy between the L1-L2. These findings again show support for the Competition Model, due to sensitivities to structures similar between the L1-L2, and no sensitivities to those that were dissimilar.

Finally, in a study of proficient L1 Chinese-L2 English learners, Chen, Shu, Liu, Zhao and Li (2007) tested ERP and behavioral processing of a structure unique to the L2: Subject-Verb agreement. While this is an important syntactic element in English, Chinese does not have grammatical morphology for marking case, number, or gender. Both L2 learners and native L1 English participants were highly accurate in behavioral judgments to stimuli, unlike performance of L2 learners in some other studies who performed at chance on behavioral measures (Osterhout et al., 2006; Tokowicz & MacWhinney, 2005). However, in ERP processing L2 learners did not match native L1 participants – while native L1 participants demonstrated expected LAN/P600 processing to ungrammatical sentences, proficient L2 learners demonstrated an N400/N600 pattern. While structures similar between the L1-L2 were not tested to fully assess applicability to the Competition Model, this is evidence that L2 learners do not always trend in ERP processing as predicted by the model.
Proficiency-based Perspective

A second perspective in the literature on L2 sentence processing, in particular in neurocognitive studies on L2 processing, emphasizes the role of L2 proficiency in rate and success of L2 learning. In the Proficiency-based perspective on L2 learning as described most extensively by Steinhauer et al. (2009), every language learner will exhibit certain ERP patterns and cognitive processes while going through different stages of L2 proficiency. Steinhauer et al. (2009) distinguish 6 levels of L2 proficiency.

These six stages are posited to occur systematically, and that while language learning factors such as L1 background, individual differences, particular structure examined, L1 proficiency, and learning environment might affect the trajectory through the proficiency levels or ultimate endpoint, L2 learners will pass through all of these stages to approach native proficiency. The cognitive processes associated with each proficiency level are thought to impact the co-occurring ERP pattern. At the first proficiency stage, the L2 learner is considered a novice learner. These learners are not predicted to show any sensitivity in neurocognitive processing or accuracy to L2 constructions. The second proficiency stage L2 learners reach is that of very low proficiency. These learners are predicted to rely more heavily on semantics and pragmatics than syntax during L2 processing. As a result, N400 ERP patterns are predicted during processing of ungrammaticalities in the L2. In the third stage, L2 learners would be considered low to intermediate. They would just be beginning to use grammatical cues in sentence processing, thus an attenuated P600 would be expected, if not still N400s as they are just beginning grammaticalization. Use of syntax in L2 sentence processing would increase as the L2 learner advances into the fourth proficiency stage, the intermediate stage. In this stage N400s would no longer be expected, as the L2
learners would not be relying on compensatory strategies, rather large P600s would be expected. As an L2 learner is further exposed to the L2, L2 processing is predicted to become nearly automatic. Thus, a robust P600 would be predicted. As this automatic processing is occurring, anterior negativities may be seen prior to the P600, but this would be predicted to be bilateral. In the final stage of proficiency, L2 learners would be considered native-like. These learners would pattern as native speakers do, with highly automatized language processing, corresponding to a now lateralized negativity (LAN), followed by a robust P600 to grammatical errors.

Many previous studies that have investigated how early versus late L2 learners process ungrammatical utterances have confounded proficiency with AoA and level of exposure to the L2 (Hahne & Friederici, 2001; Weber-Fox & Neville, 1996). However, in studies in which these can be teased apart, the potential importance of proficiency on how L2 ungrammaticalities are processed is becoming apparent.

In one of the few experiments controlling for the confound of AoA and proficiency, Steinhauer White, Cornell, Genesee and White (2006) tested Chinese and French learners of English (AoA ~15 and 19 yrs old respectively), from high and low proficiency groups, in their processing of their L2 (English). In response to ungrammatical sentences in a grammaticality judgment task, high proficiency and native speakers demonstrated ERP patterns of LAN followed by P600, while low proficiency learners demonstrated only a P600. This lends support for the Proficiency-based model, in that the high-proficiency speakers demonstrated ERP patterns of native-like processing, while the low proficiency learners patterned like predicted for low to intermediate L2 learners. In contrast to what the Competition Model would predict, these
sensitivities to ungrammatical sentences were found irrespective of typological similarity/dissimilarity between L1-L2, for both the Chinese and French learners of L2 English.

A similar trend was reported in Rossi, Gugler, Friederici and Hahne (2006) in which late German-Italian and Italian-German L2 learners of differing proficiencies were compared in ERP patterns during L2 processing of word category and morphosyntactic violations. Each group of bilinguals included high proficiency learners and low proficiency learners. In both groups of bilinguals, and thus irrespective of L1, high proficiency learners demonstrated a (E)LAN and P600 effect, while low proficiency learners demonstrated no (E)LAN and attenuated P600 effects. Again, these findings support the Proficiency-based model of L2 learning, as L2 learners of low proficiency demonstrated ERP patterns similar to one another, independent of their L1. This finding also supports the Competition Model, as the same pattern was found regardless of L1 for both word category and morphosyntactic violations, structures similarly represented between the L1-L2. However, the Competition Model would predict L2 learners to show sensitivity to these structures even at low proficiency, as transfer of L1 representation to L2 is facilitated.

While there is evidence in support of the Proficiency-based model, there are some critical components that still need to be addressed. This model does not account for the role the L1 plays in L2 learning (even at high proficiencies, where there is often still differential performance based on structure tested (Hahne, Mueller and Clahsen, 2006)). Second, this model is based on adult speakers, and lists as supporting evidence only studies with adult L2 learners. This Proficiency-based model also brings to light a long-
standing issue of how to categorize a learner's proficiency, and whether each level is as distinct as presented in the model. Further parameters are needed to allow this model to accurately describe the stages L2 learners may pass through in learning, and how these may be affected by the variables highlighted above.

The Competition Model and the Proficiency-based model presented above are feasible theories of how L2 learning occurs, although the evidence is not yet conclusive for either model. However, neither theory accounts for a developmental perspective, L1-L2 relationship, or individual difference factors such as L1 and L2 proficiency, executive functions, working memory, and attitude and motivation towards learning an L2. As few studies have controlled these factors, there is not enough systematic evidence to be able to speak to the role these variables play in L2 learning. The present study examined how these variables correlate with L2 sentence processing, and the models above may be able to be refined as a result.

There is little research on how children process grammatical structures in their L2 relative to adults (Marinova-Todd et al., 2000). As adults form the basis for the Competition and Proficiency-based Models, it is unknown whether they will display the same trend of sensitivity to similar structures between the L1 and L2 more than those that are dissimilar. It is also unknown how they will perform in the L2 at various proficiency levels.

In a recent ERP and behavioral study (measuring off-line grammaticality judgments and online self-paced reading times), Brenders, van Hell, and Dijkstra (in prep) recruited Dutch-English child beginning L2 (English) learners at about 60 hours of L2 instruction, and presented them with L1-L2 similar (first person singular and past
tense (e.g., “The bad person killed/*kilt all the tigers.”)) and unique (present progressive (e.g., “The dog is *barking/*bark in the kitchen during dinner.”)) structures. They found that these child beginning L2 learners demonstrated P600s (see methods for discussion of ERP components) to the 1st person singular similar structure, and LAN/small P600s to the past tense similar structure. Moreover, a very small LAN effect was found to the unique construction where a LAN-P600 has been reported previously in native speakers. Even though the ERP data showed an emerging sensitivity to violations of L1-L2 similar structures and the L2 unique structure, the child L2 learners performed at chance on the behavioral measures. In a parallel study testing adult proficient Dutch-English bilinguals on the same structures, Brenders et al. observed a P600 effect for similar structures, and a LAN-P600 effect for the unique structure.

These findings suggest that child beginning L2 learners were already sensitive to violations of similar and unique syntactic structures at an early point in L2 learning. However, even though their responses reflect ERP components that are commonly found in response to syntactic violations, they do not mirror the findings from previously tested adult L2 beginning learners, such as those in Tokowicz and MacWhinney (2005). In Tokowicz and MacWhinney (2005), adult beginning L2 learners were found to have significant sensitivity to ungrammatical constructions in the unique structure, with a P600 reported. In contrast, beginning child L2 learners in the Brenders et al. (in prep) study demonstrated a very small LAN. With regards to similar structures, adult learners in Tokowicz and MacWhinney (2005) demonstrated P600s, and child L2 learners demonstrated P600s to one similar structure, and a LAN to another. These findings
suggest that there are (subtle) differences between child and adult L2 learners, which may be related to differences in AoA.

However, before reliable conclusions can be drawn as to the effects of L1-L2 similarity and age-related differences in learning, additional measures must be included. Brenders et al. included similar and unique structures, but not dissimilar structures (that are quite rare in English and Dutch, being both Germanic languages). The present study included similar, dissimilar and unique structures, so that the full Competition Model could be tested. Furthermore, Brenders et al. only tested child L2 learners, but not adult L2 learners. By presenting the same grammatical constructions to both child and adult beginning L2 learners in the current study, effects of age-related differences (while controlling for amount of L2 exposure) were explored further. Moreover, the current study examined how variability in working memory, executive functions, L1 proficiency, and attitude/motivation to learn an L2 affect L2 morphosyntactic processing. To the best of our knowledge, no previous studies have systematically examined the role that inter-individual variability in cognitive and linguistics factors plays on the processing of morphosyntactic structures in child and adult L2 learners.

While there is still much to learn about between group variation in morphosyntactic processing in child and adult L2 learners, there is also little known about variation within these groups of L2 learners. To date, most ERP studies of language processing have treated individuals within a group as uniform, averaging ERP responses of learners across trials, and using traditional grand mean analyses to inform how individuals of a certain proficiency level or L1-L2 composition pattern in language processing. In recent studies (Meulman et al., 2014; Tanner et al., 2013; Tanner & Van
Hell, 2014), traditional grand mean analyses have been dissected to investigate within group individual processing, with the finding that there are a range of N400-like, P600-like, and biphasic responses present across participants during processing – learner-specific patterns that are averaged out through traditional group grand mean analyses. For example, in Tanner et al. (2013), 20 novice adult L1 English classroom learners of L2 German, and 26 adult L1 German or L1 English classroom learners of L2 German were tested on morphosyntactic processing of a structure similar between German and English – Subject-Verb marking. In language processing of native and highly proficient German learners, there was a significant effect of grammaticality at 500-800 ms after stimulus onset – a P600. The novice group demonstrated a biphasic pattern of processing, with first an N400, followed by a P600 to ungrammaticalities in the L2. Instead of concluding that this grand average was representative of novice L2 learner language processing, Response Dominance Indices (RDI) were computed to obtain a more detailed view of individual learner processing. Tanner et al. (2013) found that some learners demonstrated a relatively large N400-effect whereas others demonstrated a relatively large P600-effect, and that the averaging of these participants had created a false biphasic N400-P600 pattern in the grand mean analysis. By correlating d-prime values for sensitivity in grammaticality judgments in the L2 to effect magnitude, they found a positive correlation between P600-like RDI and proficiency in the L2 in that as L2 proficiency increased L2 learners demonstrated a more typical P600 response to grammatical violations. In the present study I will be incorporating RDI analyses within each group, to capture more accurately individual L2 learners' responses to syntactic structures in L2 (and L1). Such analyses have not yet been conducted in child L2 learners, so this research will also
explore the extent to which this technique can be reliably applied to developmental populations.

Inclusion of the individual difference criterion above will allow for correlations with language processing measures that may further elucidate what contributes to language processing patterns in the L2, within and between groups. In the next section, I will discuss the potential role of each of these factors on L2 learning and processing.

The role of individual difference variables in L2 morphosyntactic processing

Working memory
Verbal working memory capacity refers to the ability to maintain goal-relevant verbal information in mind during processing, which can then be applied during cognitive processing tasks such as sentence comprehension. Verbal as well as visuo-spatial working memory capacity increases through development (Gathercole et al., 2004; Thomason et al., 2008), with children demonstrating functionally smaller working memory capacity than adults, even in studies testing children up to age 14 (Thomason et al., 2008). There are three different views on working memory related to L2 learning. Miyake and Friedman (1998) proposed the “more is better” hypothesis, which suggests that during L2 learning there is a large reliance on verbal working memory due to the non-automatic nature of L2 learning. Learners with higher verbal working memory capacity will be better able to retain information in working memory, which will enhance L2 learning. For example, Gathercole, Pickering, Ambridge and Wearing (2004) tested children aged 4 -15 on performance in various dimensions of working memory. From age 6 and above, working memory components increased linearly into early adolescence. By this regard it could be hypothesized that children would have decreased L2
morphosyntactic performance over that of adults, due to their lower working memory capacity.

A variant of the “more is better” has been proposed by DeKeyser (2012) who argues that “more is better" is only true for adults. He argues that individual differences in working memory do not play as large of a role for children as they do for adults. This perspective would imply that for adults high working memory confers advantages in L2 learning, while low working memory would put the adult L2 learner at a disadvantage. For children however – possibly due to less range in working memory capacity overall – individual differences in working memory would not affect L2 learning and processing.

The third perspective, the "less is more" perspective, is the most long-standing and widely-discussed of the three views (Newport, 1990; see also Goldowsky & Newport, 1993). This view argues that children are better L2 learners, because children have less-developed cognitive resources than their adult counterparts, which leads to an advantage in L2 learning (i.e., Cochran, McDonald & Parault, 1999; Kersten & Earles, 2001; Long, 1990; Ludden & Gupta, 2000). With less cognitive resources available to them (e.g., working memory), children are hypothesized to begin learning the L2 with the smallest/local constructions that form fundamental aspects of language, and acquire more complex constructions as time passes (i.e., as both L1 and cognitive resources develop). This is in contrast to adult L2 learners who, due to their more advanced cognitive resources, attempt to process grammatical complexity from very early learning, making early acquisition more difficult. This adult learning is thought to consist of memorization of larger units without the ability to accurately recombine these in language learning (Rohde & Plaut, 2002).
Elman (1993) tested the less-is-more hypothesis in a two-part study using networks to approximate L2 learning. The artificial language he used consisted of number agreement, different verb argument structures, and embedded clauses. In the first experiment, the effects of complexity on learning were tested through presentation of sentences with recurrent patterns, in order to facilitate word prediction. As the networks can only make predictions if the structure of the language is learned, this was viewed as parallel to the experience of an early L2 learner. Elman found that when presented with complex sentences from the grammar immediately, the network could not engage in word prediction; however, when presented with simple sentences first, with increasing complexity as the ratio of simple to complex sentences diminished, the network could successfully predict words. This supported the less-is-more hypothesis in that smaller input resulted in more successful learning. However, this scenario is not very realistic in that children are not exposed to only simple sentences with increased complexity, but rather receive complex input (albeit less than adult counterparts) from initial learning. More realistic is the idea that they would receive complex input throughout, but that their memory capacity would change with development. In part 2 of this study, Elman simulated the effects of changes in working memory by initially reducing the amount of memory available to the network. Elman found that when memory was limited and gradually increased, successful learning could occur, as opposed to when the full language was presented to the full memory system. Through slowly increasing network memory, representations could be formed on nuances of the language, rather than memorization of overall language trends (i.e., learning even irregularities by means of incremental presentation).
In the studies using artificial languages (e.g., Elman, 1990; 1991; 1993) or small subsets of natural language (e.g., Cochran et al., 1999; Goldowsky & Newport, 1993; Kareev, Lieberman, & Lev, 1997), participants have been presented with a finite set of words/structures, and their learning has been measured. This is far from the rich environment of a real language, and the rich input they would come across in a classroom environment. Thus, the effects of memory or sentence complexity could change with more input. As it took many more trials for the lower working memory condition to learn the artificial language (Elman, 1993), it may be that ultimate attainment for children would be better, but initial learning (the general rule knowledge shown by the “full” memory manipulation) would be better for adults. While these simulated artificial environments may result in support for the “less-is-more” model, the variance is much lower than that which would be encountered in real language, thus limiting the applicability to real-world L2 learning scenarios. The current study examined L2 morphosyntactic processing of children and adults in real language environments to gain better insight as to how L2 learning and processing is influenced by differences in working memory capacity.

Executive Functions and Cognitive Control

Much research on adult L2 learners has shown that the L1 and L2 are both active during language use (for reviews, see Kroll, Bobb, Misra, & Guo, 2008; Van Hell & Tanner, 2012), however errors from the non-target language are not common while speaking the target language. It has been suggested that cognitive control is largely responsible for this -- the ability to inhibit (Green, 1998) the non-target language and suppress it to the extent that the target language can be accessed without interference.
Through cognitive control irrelevant information and incorrect responses can be inhibited. In lower proficiency L2 learners, more cognitive control is needed to suppress the dominant L1 in order to successfully access and use the L2. As proficiency in the L2 increases strength of inhibition needed for non-target language suppression decreases. In adults we can assume that greater abilities in cognitive control should aid in language performance, particularly in these earlier levels of L2 proficiency, as they can better inhibit the dominant L1. As children have significantly lower levels of attention and cognitive control (Davidson, Amso, Anderson & Diamond, 2006), they may not be as successful in inhibiting the L1, resulting in possible interference during L2 learning. By measuring variations in executive functions in child and adult L2 learners, we can gain a better idea of how L2 performance may be influenced by this variable, particularly in child L2 learners where this is largely unexplored.

**L1 Proficiency**

While native language performance is often the golden-standard to which non-native processing and production is compared, there is much variability in L1 proficiency and resulting ERP patterns (Pakulak & Neville, 2010) that may partially explain variance in L2 performance (van Hell & Tanner, 2012). A more proficient L1 speaker is likely to have highly automatized language retrieval, and well-established language cues. This variability likely influences processing patterns in the L2, particularly for children who are still developing the L1 – a protracted process (Berman & Nir-Sagiv, 2004; Nippold, 2007; van Hell, Verhoeven, Tak, & van Oosterhout, 2005), while trying to learn an L2. The potential effects of variation in L1 on L2 learning patterns makes it is necessary to measure proficiency in both the L1 and L2 of learners. Remarkably, this is something that
is rarely done in L2 research and few if any studies examined how individual differences in L1 may affect L2 learning. In the present experiment variability in L1 proficiency (in terms of vocabulary and grammatical knowledge) was related to L2 grammatical processing, in both children and adults. Furthermore, variability in L2 vocabulary knowledge was related to L2 grammatical processing, again in both children and adults.

**Attitude and Motivation**

While attitude and motivation are not factors which are typically included in (neuro)cognitive studies of individual differences in language processing, these factors are often included in studies of second language acquisition -- affective variables such as attitude and motivation to learn an L2 have been shown to have a significant effect on learning (Masgoret & Gardner, 2003). Attitudes a learner has towards a language or speakers of that language influence level of effort put forth towards language learning (Gardner, 1980; Nunan, 2000). Negative affect towards a language or speaker group can actively reduce effort put forth to learn that language, and can result in less successful L2 learning as compared to an individual with positive affect. Motivation to learn a language can play a large role in L2 learning trajectory. Intrinsic motivation to learn a language because of personal enjoyment or because of pride in learning is thought to push speakers forward the most in rate and success of learning. Extrinsic motivators such as learning a language for social use, or to be able to communicate for a job, are not as influential for learning (Crookes & Schmidt, 1991), but are more so than engaging in language learning solely for course credit, or for a grade in school.

In the present study the role of attitude and motivation towards the L2 (Spanish) were measured (see methods section). To our knowledge no studies exist that have taken
into account the role of affect in child L2 learners, so the influence of these variables on L2 learning rate and success is a novel contribution to this literature. Additionally, measuring affective variables for intermediate adult L2 learners will allow for insight into effects on L2 learning, something unknown for intermediate-level L2 learners.

**The Present Study**

Much of the behavioral and ERP research on classroom L2 learners has been gathered from highly proficient learners (for a review, see Clahsen & Felser, 2006). Far less is available on beginning and intermediate language learners, and even less so on morphosyntactic processing of child L2 learners. Furthermore, in the studies that have attempted to speak to the rate and success of L2 learning for children as well as adults, there has often been the inevitable confound of AoA with exposure to the L2, proficiency, and development of cognitive faculties (Birdsong & Molis, 2001; Johnson & Newport, 1989; Weber-Fox & Neville, 1996). Additionally, the potential influence of the typological relationship of L1 to L2 have largely been ignored in child L2 learning research, and little is known about how variability in working memory, executive functions, L1 proficiency and attitude/motivation to learn an L2 modulate L2 sentence processing.

The present study investigated morphosyntactic processing of adult and child intermediate L2 learners during L2 sentence reading. ERPs and grammaticality judgment accuracy scores were gathered, as well as the individual difference measures listed above (e.g., working memory, executive functions, L1 proficiency, and attitude and motivation in L2 learning). Through manipulation of structures that are manifested similarly or dissimilarly between the L1 and L2, or that are unique to the L2, I sought to investigate
mechanisms related to transfer and competition between L1 and L2 during L2 sentence processing. Inter-individual variation in sentence processing has been shown within groups of L2 learners (Meulman et al., 2014; Tanner et al., 2013), as well as in monolingual language processing (Tanner & Van Hell, 2014). Through calculations of individual Response Dominance in ERP, and by correlating the ERP data to individual difference measures, we may shed light on what contributes to this variation in language processing, in L2 learners at two different ages.

Through the combined use of behavioral and neurocognitive (ERP) measures, we seek to gain more insight into how L2 learning manifests in children and adults with similar levels of L2 exposure. As seen in studies above, in Tokowicz and MacWhinney (2005; see also Brenders et al., in prep; McLaughlin, Tanner, Pitkänen, Frenck-Mestre, Inoue, Valentine & Osterhout, 2010) when grammaticality judgments were at chance levels, ERP signatures indicated learning at very early L2 exposure. ERPs (see Methods section) provide an excellent time-sensitive measure of neurocognitive processes during language comprehension as they unfold over time, and thus provide insights that cannot be obtained through offline grammaticality judgment measurements.
Chapter 2: METHODS

Participants

Adult L2 learners

Adult participants were 20 right-handed native-speakers of English who were learning Spanish as a second language (L2), at Penn State University. They were aged 18-23 (8 male; mean age: 19.5). All students were intermediate-level classroom learners (third semester), and were tested at approximately 205 hours of L2 instruction. Students were screened for any significant previous experience with Spanish prior to taking Spanish language classes at time of testing. At most they may have had two courses of introductory Spanish in early high school. Adult learners exposed to any language other than English before age 5 were not included in this study. Participants all had normal or corrected-to-normal vision, and no history of neurological or language disorders. In addition to a prescreening process, each participant completed a language history questionnaire that compiled information regarding L1 and L2 language experiences, all languages participants were exposed to, the age at which exposure began, and amount of time spent using or interacting with Spanish outside the classroom.

Child L2 Learners

Child learners were 7 right-handed native-speakers of English who were learning Spanish as an L2 at a charter school in Pennsylvania at which Spanish instruction is provided from Kindergarten to 8th grade. They were aged 9-11 (2 male; mean age: 10.33). All students were fourth and fifth grade intermediate-level Spanish classroom learners, and were tested at approximately 400-500 hours of L2 Spanish instruction. Participants all had normal or corrected-to-normal vision, and no history of neurological or language
disorders. In addition to a prescreening process, each child’s parent completed a language history questionnaire that compiled information regarding their child’s L1 and L2 language experiences, all languages children were exposed to, the age at which exposure began, and amount of time spent using or interacting with Spanish outside the classroom. In the child L2 learner sample, one participant had only one year of Spanish classroom exposure, and one participant had had three years of prior experience in a Spanish immersion classroom, and was raised with bilingual exposure to Russian and English in the home from birth. Additional children will be tested for this experiment in Spring 2016 – one year after the initial cohort. Once child L2 learner data collection is completed, these two children will be excluded from the final dataset.

**Materials**

**Spanish sentence stimuli**

The Spanish component of the grammaticality judgment task consisted of stimuli from three grammatical constructions (determiner number agreement, auxiliary omission, and determiner gender agreement). These stimuli fell within one of three categories. Auxiliary omission is formed similarly in English and Spanish (i.e., there is high cue reliability in both English and Spanish for subject-verb agreement. In both languages subject-verb agreement is required in a sentence, and cues indicating agreement are available and heavily utilized). Determiner number agreement is formed differently in English and Spanish (i.e., English marks number only on the noun, while Spanish requires marking on both the noun and verb). Determiner gender agreement is unique to Spanish (i.e., gender marking on nouns is not used in English, so this is a cue which is only valid in the L2). See Table 1 for examples. A total of 120 Spanish sentences were
presented to each participant, 40 items for each experimental construction. Sentences were very similar to those presented in Tokowicz and MacWhinney (2005), but for both children and adults vocabulary was modified to mirror that encountered in their L2 classes thus far. A fluent Spanish speaker reviewed all sentences for clarity and grammaticality prior to testing.

<table>
<thead>
<tr>
<th>Relation to L1 (English)</th>
<th>Function</th>
<th>Sample Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similar</td>
<td>Auxiliary omission</td>
<td>*Su abuela cocinando muy bien. *His grandmother cooking very well</td>
</tr>
<tr>
<td>Dissimilar</td>
<td>Determiner number agreement</td>
<td>*Ellos fueron a una fiesta. *They went to a-MASC party-FEM</td>
</tr>
<tr>
<td>Unique to L2</td>
<td>Determiner gender Agreement</td>
<td>*La chicas bailaron. *The-SING girls-PL danced</td>
</tr>
</tbody>
</table>

Table 1: Examples of structure types tested in Spanish.

**English sentence stimuli**

The English grammaticality judgment task consisted of English stimuli from three grammatical constructions similar to the materials used by Tokowicz and MacWhinney (2005): Auxiliary Omission (Tense), Reflexive, and Subject-Verb agreement.

<table>
<thead>
<tr>
<th>English sentences</th>
<th>Function</th>
<th>Sample Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Auxiliary omission</td>
<td>*His grandmother cooking very well.</td>
</tr>
<tr>
<td></td>
<td>Reflexive</td>
<td>*The children enjoyed himself.</td>
</tr>
<tr>
<td></td>
<td>Subject-verb agreement</td>
<td>*The boys makes excellent ice cream.</td>
</tr>
</tbody>
</table>

Table 2: Examples of structure types tested in English.

A total of 120 English sentences were presented to each participant; all were experimental items (40 items for each experimental construction).

**Procedure**

Prior to the experiment each participant completed a language history questionnaire to ascertain eligibility. Child participants had had a parent or guardian complete the form prior to their participation. Participants were consented prior to
participation.

The experiment was conducted in two locations. Children were tested in a mobile ERP van, while adults were tested in ERP facilities at Penn State University, using the same EEG equipment. EEG recordings in both locations took place using an elastic cap with 30 active electrodes (Brain Vision), using the international 10-20 system (Jasper, 1958; Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, PO9, O1, Oz, O2, PO10). The electrode configuration and the recording protocol were identical in both locations. Two electrodes on the left and right mastoids were used as reference points for scalp electrodes (TP10 in online reference, TP9 in offline). Vertical eye movements were recorded using one electrode underneath the left eye. Horizontal eye movements were recorded using one electrode at the outer canthus of the right eye. Each participant read sentences in Spanish and English. Participants and experimenters stayed in the same room while the task took place, while the researcher monitored ERP recordings on an adjacent monitor. In all tasks participants were permitted to take breaks as needed.

In the grammaticality judgment task, participants made judgments for Spanish and English sentences after the presentation of the final word (see Figure 1), indicating whether the presented sentences were grammatically acceptable in the language of presentation. The language of presentation was blocked with 2 blocks of Spanish sentences and then two blocks of English sentences in the first testing session. In the second test session participants completed all behavioral tasks. The purpose for stimuli presentation in English as well as Spanish was twofold: 1) To validate ERP setup and to see whether the sentences in English show typical ERP signatures (see end of Method
section for a description) associated with syntactic anomalies, in the children and the adults' online processing; 2) To provide a measure of a participant’s proficiency in their L1, English.

Half of the critical sentences were grammatical and half were ungrammatical. Participants never saw the grammatical and ungrammatical version of the same sentence. Sentences were randomly assigned to two versions of the experiment, and a given sentence appeared in the grammatical form to some participants, and in the ungrammatical form to others. The critical word in each sentence was at the violation point. The sentences were presented in a random order through E-Prime, which also recorded the reaction time for providing grammaticality judgments, and sent critical word onset information to the ERP acquisition software (PyCorder). After reading the sentence, participants indicated their grammaticality judgment by pressing one of two buttons on a button box; they pressed a button marked “☺” with their left hand to indicate if they thought the sentence was unacceptable and a button marked “☺” with their right hand if they thought the sentence was acceptable.
Figure 1: Timeline of stimulus presentation during primary task.

Prior to each sentence, a fixation cross (+) appeared at the center of the computer screen, followed by a smiley face. Participants were told they could blink when the smiley face was on the screen. After the smiley face disappeared sentences were presented one word at a time at the center of the computer screen (to minimize eye-movements). Following the timing procedure of Brenders et al. (in prep) that has been used successfully in child L2 learners and adult bilinguals, each word remained on the screen for 400 ms with a blank screen appearing for 400 ms between words. After the offset of the final word of the sentence, a blank screen appeared for 1500 ms, followed by a question mark that served as a prompt for a grammaticality judgment. As soon as the prompt appeared, participants were instructed to respond by pressing either “☺” or “☺”. Following response, the screen appeared with ☺ again, indicating participants could blink prior to next trial start. After completion of the ERP sentence task, participants were presented with a list of the critical L2 words in the experiment (i.e., the words which
were the point of violation in each sentence), and asked to provide L1 translations. This
was used to assess the participants’ explicit L2 knowledge of the target stimuli.

In the second session, participants were asked to complete behavioral tasks that
measure working memory, cognitive control, L1 and L2 proficiency, and
attitude/motivation to learn an L2 (these tasks are described in the next section). Adults
completed the Simon task, ANT, Rey Task, Forward and Backward Digit Span Task,
Operation Span Task, PPVT, TVIP, Language History Questionnaire and the Attitude
and Motivation Assessment. Children completed the Simon task, Rey Task, TVIP, ANT
for children, Forward and Backward Digit Span Task, Reading Span Task, PPVT, Corsi
Task and Mini Attitude and Motivation Assessment. Where adults and children
completed the same tasks test materials were identical, except for the Attitude and
Motivation test, ANT, and Language History Questionnaire.

Once the behavioral tasks were completed, the participant was debriefed. Total
testing time of each session was between 1 to 2 hours. The ERP testing session included
10 minutes for application of the cap, 1 hour for the sentence reading task (including self-
timed breaks), and 10 minutes for participant clean-up.

**Individual difference measures**

After completing the primary ERP sentence reading tasks, participants completed the
cognitive tasks, L1 and L2 proficiency tasks, and a questionnaire that measured
attitude/motivation to learn an L2. For all tasks, practice trials preceded experimental
blocks.
Cognitive tasks

Working Memory

**Corsi Block Tapping Task.** The Corsi Task (Corsi, 1972), a measure of visuo-spatial working memory capacity, was presented to child participants. Children were shown a set of 9 blocks, and had one light up at a time, increasing in set size from 2-8 blocks. They were asked to then click on each block that was lit up, in the order they appeared. The test continued until they reached ceiling – two incorrect responses in a row for a given set size. Accuracy was recorded and memory span will then be calculated.

**Rey-Osterrieth Complex Figure Task.** The Rey Task (Rey, 1941) was administered to both children and adult participants. Participants are shown a picture, and asked to copy it. They are then asked to recall the figure 3 minutes later (immediate recall), 30 minutes later (delayed recall) and then to identify components from the Rey image from correct and incorrect images presented. Accuracy and copy times are recorded at each step. The Rey complex figure measures recognition memory for visuo-spatial details, and assesses the respondent’s ability to use cues to retrieve information.

**Operation Span Task.** The Operation Span Task (Turner & Engle, 1989) was presented to adults. Participants were presented with a simple arithmetic problem (e.g., (9+1)/2 = 4?), and asked to judge whether it was correct or not by pressing one of two buttons on a button box. Between every question, a word (from L1, English) was presented to the participant, which they were asked to remember. After 2-6 trials, they were asked to recall the words they saw, in whichever order, as long as the last seen word was not typed in first.
The Operation Span Task was divided into 5 blocks, with increasing complexity per block. In each block, there were three sets of trials. The task began with a set size (equation and word) of two and increased to a set size of six. Accuracy was recorded, and working memory capacity was calculated.

**Sentence Reading Span Task:** The Sentence Reading Span Task (Daneman & Carpenter, 1980) was presented to child participants. Participants were presented with a set of unrelated sentences on a computer screen, one at a time, and were prompted to answer a probe question about the sentence while memorizing the final word of each sentence. After a given set of sentences, they were asked to recall the sentence-final words in the correct order. The set size of the sentences increased throughout the task. A total of 56 sentences were presented to participants. Trials were randomized across participants. Accuracy and reaction time information were recorded.

**Forward and Backward Digit Span Tasks.** The forward and backward digit span tasks (Wechsler, 1997) were presented to children and adults. Participants were presented with a digit on a computer screen. The digit then disappeared, and the participant had to state what the digit was. The number of digits increased by block, until 7 digits was reached; each block contained six trials. In the forward digit span task participants were asked to recall the digits in the order they were presented (this began with a three digit trial). In the backward digit span task participants were asked to recall the digits in the opposite order than which they were shown (this began with a two digit trial). Once participants missed two trials in a row, the task stopped. Accuracy was recorded and memory span was noted.

**Executive Functions and Cognitive Control**

**Simon Task** (Simon & Rudell, 1967). In the Simon Task a blue or red square was
presented on the right or the left side of the computer screen. Participants were instructed to press a button corresponding to the square’s color, while ignoring location. A total of 28 trials were presented to participants. Accuracy and reaction time information were recorded. The Simon Task measures executive control and is a conflict-processing task that relies minimally on language and memory processes that may interact with language proficiency. The Simon task is based on stimulus-response compatibility, and assesses the extent to which a person’s response to task relevant nonspatial information is affected by the dominant association to task irrelevant spatial information.

**Attention Network Task, ANT** (Fan et al., 2002). In the Adult ANT task, participants were instructed to ignore a series of flanker items (congruent (>>>>>), neutral (--->--), or incongruent) to press a button corresponding to a target. They were alerted to target presence on certain trials by a beep or * appearing on the computer screen. A total of 144 trials were presented to participants, split across 3 blocks of 48 trials. Accuracy and reaction time information were recorded. In the child ANT task, participants were also instructed to ignore a series of flanker items, and to press a button corresponding to the target. Stimuli were presented as fish, that were all either facing in the same direction as the target center fish (congruent), in the opposite direction from the target center fish (incongruent), or as the target fish flanked by no fish (neutral). They were alerted to target presence on certain trials by a beep or * appearing on the computer screen. A total of 144 trials were presented to participants, split across 3 blocks of 48 trials. Accuracy and reaction time information were recorded.

The ANT is a conflict-processing task that relies minimally on language and memory processes that may interact with language proficiency. The ANT is a
combination of a cue reaction time task (Posner, 1980) and a flanker task (Eriksen & Eriksen, 1974), and measures three attentional networks: executive control (monitoring and conflict resolution), alerting (attainment and maintenance of an alert state), and orienting (selection of information from sensory input).

**Proficiency measures**

**The Peabody Picture Vocabulary Test (PPVT) & Test de Vocabulario en Imagenes Peabody (TVIP).** In the PPVT and TVIP, four numbered photos were presented on a sheet of paper, and upon a one-word description, participants indicated which photo is being referenced through number naming. Order of trial presentation was randomized for each participant. The PPVT was administered in Spanish as the TVIP (Dunn, et al., 1986) as well as in English (Dunn & Dunn, 2007). One-word Spanish descriptions were pre-recorded. Complexity of pictures increased across trials. When six trials in a row were missed, the task was concluded. A total of 228 items could have been presented to each participant. Accuracy, and ultimate level reached, were recorded.

In addition to the PPVT, the grammaticality judgments of the sentence readings tasks in Spanish and English are also indices of L2 and L1 proficiency, respectively.

**Attitude/Motivation to Learn an L2**

**Attitude and Motivation Test Battery (AMTB) (Adults) (Gardner, 1985).** In the AMTB, completed by the adult learners, 98 multiple-choice questions were presented on a computer screen. Participants were asked to report their levels of agreement with a series of statements regarding their attitude and motivation towards the Spanish language and culture. For example, participants read statements like “Studying Spanish is important because it will allow me to be more at ease with people who speak Spanish.”
Then participants made responses indicating whether they strongly disagreed (1) to strongly agreed (7) with a particular statement on a Likert scale.

**Mini-AMTB (Children)** (Masgoret, Bernaus, & Gardner, 2001). In the mini-AMTB, 11 multiple-choice questions were presented, asking for levels of agreement or disagreement with a series of questions about Spanish learning and the Spanish language. This test has been shown to be approximately as accurate as the larger AMTB for populations who cannot complete the full test. Only participants aged 9-11 were administered this task. Children selected a number from 1 (strongly disagree) to 7 (strongly agree) on a Likert scale to indicate their opinions.

**Neurocognitive method: ERPs**

Electroencephalography (EEG) is the scalp-recorded electrical activity produced by the brain from which Event Related Potentials (ERPs) are obtained by time-locking the signal to a given stimulus. ERPs provide a record of the brain’s electrical activity during mental processing as it unfolds over time. The main ERP components that are considered to index (morpho)syntactic processing are the left anterior negativity (LAN) and the P600 (e.g., Friederici, 2002). However, N400 effects have been reported in syntactic processing for some learners (e.g., Bañon et al., 2014; Tanner et al., 2014).

The LAN is an anterior negativity, considered to reflect syntactic processing, particularly the detection of morphosyntactic errors. It is often left-lateralized, but can also be bilateral. A LAN is not always reliably found in syntactic processing (cf. Tanner & Van Hell, 2014; see review Van Hell & Tokowicz, 2010), at least less consistently so than the P600. The P600 is a positive going component, associated with (morpho)syntactic processing, generally when a violation in an expected structure occurs,
or when there is processing of particularly complex syntactic structures (i.e., garden path sentences) (Kaan et al., 2000; Osterhout & Holcomb, 1992). The P600 is thought to reflect a late process of language comprehension where syntactic repair and re-integration is attempted (Van Hell & Tokowicz, 2010). The N400 is a centro-posterior negativity, generally from 300-500ms, considered to reflect issues in semantic processing and integration into preceding context (Kutas & Hillyard, 1980; 1984). The N400 has however been reported in syntactic contexts, particularly for low proficiency L2 learners, and is thought to reflect reliance on lexical rather than semantic information (Steinhauer et al., 2009).
Chapter 3: ANALYSIS & RESULTS

Analysis

The primary questions addressed in this study are: 1) How is L2 sentence processing influenced by the (dis)similarity in syntactic structures between L2 and L1, in children and adults? 2) Does individual variation in working memory and executive functions, L1 and L2 proficiency, and attitude/motivation to learn an L2 affect morphosyntactic processing, in children and in adults?

The first question is analyzed using ANOVAs, both for the ERP data (focusing on the LAN/N400 and P600 components) and the behavioral data (accuracy grammaticality judgments), as well as t-tests on accuracy data. The second question is addressed using correlations, both for the behavioral data and the ERP data (see Tanner & Van Hell, 2014). The data for the children and the adults have been analyzed separately.

Accuracy Analysis

Accuracy in each condition was calculated for each participant, for grammatical and ungrammatical sentences. These data were analyzed with ANOVAs using accuracy in grammaticality judgment (correct, incorrect) and type of construction (similar, dissimilar, unique) as factors. To determine whether performance was at, above, or below chance for each condition, we compared each mean, of combined accuracy to grammatical and ungrammatical sentences, individually against 50% in one-sample t-tests.

Several studies have demonstrated that L2 learners may perform at chance during grammaticality judgment tasks (Tanner et al., 2005; Tokowicz & MacWhinney, 2005), with worse performance to ungrammatical sentences, while exhibiting sensitivities in
online ERP measures. In order to ascertain truly accurate response rates relative to
guesses (sensitivity), d-prime was calculated for each individual, as a measure of correct
responses (hits) versus incorrect responses to a given grammaticality (false alarms)
(Wickens, 2002). A d-prime score of 0 indicates no sensitivity; a d-prime score of 4
indicates near perfect sensitivity (maximum 6.93).

**ERP Analysis**

In order to obtain ERPs, EEG recordings were amplified with a .1 to 40 Hz
bandpass filter, and digitized with a 550 Hz sampling rate. An offline 30 Hz low-pass
filter (24 dB/octave roll-off) was applied to the EEG data, after re-referencing. EEG
gathered during language processing were time-locked to critical words (in both the
grammatical or ungrammatical sentences), and were averaged offline. ERP epochs were
defined as a period of activity beginning 200ms prior to and ending 1000ms after critical
stimulus onset. Trials that were artifact-free (without excessive muscle artifact, drift, or
eye blinks) were included in averages. For adults, an average of 2.13% trials were
excluded, for children 9.89%. These numbers did not significantly vary by condition type
in either group.

ERP components of interest (LAN/N400/P600) were investigated using a priori
defined time windows. Through visual inspection and previous research, the early
LAN/N400 window was defined as 300-500 ms, and the late P600 window was defined
as 600-800 ms for L2, and 500-800 ms for L1.

These EEG signals were averaged within each acceptability and cross-language
similarity condition for each participant. Trials associated with correct and incorrect
grammaticality judgments were included in the analyses (as is typically done with ERP
studies in beginning L2 learners, e.g., Brenders et al., in prep; McLaughlin et al., 2010; Tokowicz & MacWhinney, 2005). Within each time window ANOVAs were computed. Lateral and midline sites were analyzed separately. Data from lateral sites were split into four regions of interest (ROI) – left anterior (F7, F3, FC1, FC5), right anterior (F8, F4, FC2, FC6), left posterior (CP5, CP1, P7, P3), and right posterior (CP6, CP2, P8, P4). At lateral sites the ERP data were analyzed using a 3 (Condition: Similar, Dissimilar, Unique) by 2 (Grammaticality: Grammatical, Ungrammatical) by 2 (Anteriority: Anterior, Posterior) by 2 (Laterality: Left, Right) repeated-measures within-participant design. At midline sites the ERP data were analyzed using a 3 (Condition: Similar, Dissimilar, Unique) by 2 (Grammaticality: Grammatical, Ungrammatical) by 3 (Electrode: Fz, Cz, Pz) repeated-measures within-participant design. If interactions were shown between condition and grammaticality or a topographic variable, further ANOVAs were conducted for each condition separately (as done in Chen et al., 2007; Hahne & Friederici, 2001).

Separate ANOVAs were conducted for children and adults. The Greenhouse–Geisser correction for inhomogeneity of variance was applied to all repeated-measures on ERP data with more than one degree of freedom in the numerator, and the corrected p-value is reported.
Results

Adult

Accuracy

Mean accuracy was around chance when judging sentence grammaticality in the L2 (see Figure 2 and Table 3). The d-prime values of the similar, unique, and dissimilar conditions were -.50, -.62, and -.24, respectively.

Using one-sample t-tests, for both grammatical and ungrammatical sentences combined, the mean accuracy of grammaticality judgment in each condition was compared to 50% -- what one would score if grammaticality judgment accuracy was perfectly at chance. In the dissimilar condition performance exceeded chance ($t(39)=2.37, p<.05$). In the similar condition and the unique condition performance was at chance ($t(39)=1.99, p=.05$ and $t(39)=1.15, p=.26$, respectively).

An Accuracy X Condition ANOVA yielded no interaction, so no one condition elicited significantly more accurate responses than another.

![Figure 2](image-url) 

**Figure 2:** Accuracy by condition in L2 Spanish, similar (Tense), unique (Gender) and dissimilar (Number) conditions.
In contrast, participants were highly accurate in their judgments to grammaticality violations in the L1 (see Table 3 and Figure 3). The d-prime values of the Tense, Reflexive, and Subject-Verb conditions were 6.66, 4.80, and 5.95, respectively.

**Table 3:** Mean accuracy in correctly identifying violations in grammaticality, broken down by grammatical and ungrammatical sentences in the L2 and L1.

<table>
<thead>
<tr>
<th></th>
<th><strong>Accuracy:</strong></th>
<th><strong>Accuracy:</strong></th>
<th><strong>Overall Accuracy:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L2</strong></td>
<td>Ungrammatical Sentences</td>
<td>Grammatical Sentences</td>
<td>Mean and SD</td>
</tr>
<tr>
<td>Similar Dissimilar Unique</td>
<td>Mean = 41.00%; SD =18.32</td>
<td>Mean = 73.75%; SD =14.77</td>
<td>Mean= 57.40%; SD= 23.34</td>
</tr>
<tr>
<td>Mean = 46.00%; SD =21.31</td>
<td>Mean = 70.25%; SD =14.19</td>
<td>Mean= 58.13%; SD= 21.68</td>
<td></td>
</tr>
<tr>
<td>Mean = 39.50%; SD =20.83</td>
<td>Mean = 69.00%; SD =15.36</td>
<td>Mean= 54.25%; SD= 23.44</td>
<td></td>
</tr>
<tr>
<td><strong>L1</strong></td>
<td>Tense Reflexive Subject-Verb</td>
<td>Tense Reflexive Subject-Verb</td>
<td>Mean= 97.75%; SD=3.39</td>
</tr>
<tr>
<td>Mean = 97.75%; SD =3.43</td>
<td>Mean = 97.75%; SD =3.43</td>
<td>Mean= 97.75%; SD=3.39</td>
<td></td>
</tr>
<tr>
<td>Mean = 94.25%; SD =5.68</td>
<td>Mean = 95.75%; SD =5.68</td>
<td>Mean= 94.25%; SD=5.68</td>
<td></td>
</tr>
<tr>
<td>Mean = 95.75%; SD =5.68</td>
<td>Mean = 95.75%; SD =5.68</td>
<td>Mean= 95.75%; SD=5.68</td>
<td></td>
</tr>
<tr>
<td>Mean = 91.00%; SD =8.05</td>
<td>Mean = 91.00%; SD =8.05</td>
<td>Mean= 91.00%; SD=8.05</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3:** Accuracy by condition in L1 English to Tense, Reflexive and Subject-Verb conditions.
ERP

Spanish

In Figure 4 the grand mean waveforms for processing of ungrammatical and grammatical sentences in the L2 are shown, for a series of electrodes commonly implicated in N400 and P600 responses – centro-parietal (CP) electrodes: F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4. These analyses average all three conditions (similar, dissimilar, unique).

In the LAN/N400 window of 300-500 ms there are no significant main effects or interactions.

In the P600 time window of 600-800 ms there are no significant main effects or interaction in the midline region. At lateral sites there is a significant effect of grammaticality ($F(1,19)=4.661, p<.05$) with ungrammatical sentences eliciting greater positivity than grammatical sentences (mean amplitude difference: 1.07µV).
Figure 4: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of L2 Spanish grammatical and ungrammatical sentences. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10µV.

Grand mean waveforms for the three critical Spanish conditions similar (Tense), dissimilar (Determiner Number), and unique (Determiner Gender) across the 9 representative electrodes are presented in Figures 5A, 5B, and 5C, respectively.
5 (B)
Figure 5: Grand mean waveform from nine representative electrodes contrasting grammatical and ungrammatical sentences across the three L2 Spanish conditions: (5A) Tense (Similar), (5B) Determiner Number (Dissimilar), and (5C) Determiner Gender (Unique) conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10µV.

Table 4 presents results from the omnibus ANOVA. In the 300-500 ms window there are no significant main effects or interactions.

In the 600-800 ms window there is a marginally significant interaction of Grammaticality X Condition and of Grammaticality X Electrode X Condition at midline electrodes, with largest positivity to ungrammatical sentences at Fz. At lateral sites there is a significant effect of grammaticality, with more positive amplitude to ungrammatical relative to grammatical sentences (mean amplitude difference= 1.07µV, SE=.49), and a
marginally significant interaction of Grammaticality X Condition X Anteriority, with
greatest positivity to ungrammatical sentences in the dissimilar condition at anterior sites.

<table>
<thead>
<tr>
<th></th>
<th>300-500 ms</th>
<th>600-800 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality (1,19)</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Grammaticality X Condition (2, 38)</td>
<td>--</td>
<td>2.69*</td>
</tr>
<tr>
<td>Grammaticality X Electrode (2, 38)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Electrode (4, 76)</td>
<td>--</td>
<td>2.74*</td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality (1, 19)</td>
<td>--</td>
<td>4.79**</td>
</tr>
<tr>
<td>Grammaticality X Condition (2, 38)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority (1,19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Hemisphere (1, 19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Anteriority (2, 38)</td>
<td>--</td>
<td>2.73*</td>
</tr>
<tr>
<td>Grammaticality X Condition X Hemisphere (2, 38)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority X Hemisphere (1, 19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Anteriority X Hemisphere (2,38)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Degrees of freedom are reported in parentheses.

*.05<p<.1
** p<.05
*** p<.01

Table 4: F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for Spanish language processing in the 300-500 ms and 600-800 ms windows.

Subsequently, the same ANOVAs were conducted for each condition separately at midline and lateral electrodes, in the 600-800 ms window, as this is where interactions of Grammaticality X Condition and Grammaticality X Condition X Anteriority are found in the omnibus ANOVA.

**Similar Condition:** In midline electrodes, an interaction of Grammaticality X Electrode is found \( (F(2,38)= 5.66, p<.05) \), with greatest negativity to ungrammatical sentences at the central (Cz) electrode. In lateral sites, a significant interaction of Grammaticality X Anteriority \( (F(1,19)=4.36, p=.05) \) is found, with greater positivity to ungrammatical sentences at anterior sites.

**Dissimilar Condition:** There is a marginally significant effect of grammaticality \( (F(1,19)=4.25, p=.053) \), with greater positive amplitude in processing ungrammatical
sentences than grammatical (mean amplitude difference = 2.044µV, SE=.99). No significant effects or interactions were present at lateral electrodes.

**Unique Condition:** The ANOVAs yielded no significant effects.

**Response Dominance Index**

Further inspection of individual waveforms indicates that not all participants exhibit the trend shown in the grand mean waveforms in Figures 5A, 5B, and 5C. As found in Tanner and Van Hell (2014), certain individuals appear to exhibit an N400 while others appeared to exhibit a P600, to the same conditions.

We therefore computed a response-dominance index (RDI) for each participant to tease apart variability of positive or negative processing that appears to have been averaged out in grand mean waveforms. To compute the RDI, we calculated a participant’s mean processing activity across the electrodes most commonly implicated at peak N400 and P600-like processing – the Centro-Parietal electrodes: C3, Cz, C3, CP1, CP2, P3, Pz, P4, across the two a priori time windows representative of this processing: the P600 window (600-800 ms), and the N400 window (300-500 ms). To determine response dominance, an effect magnitude was calculated for P600 dominance, and for N400 dominance, using processing in CP electrodes in the P600 and N400 time windows, according to the formula below. P600 effect magnitude was calculated as the amplitude of response to grammatical sentences in the 600-800 ms window, subtracted from the amplitude of response to ungrammatical sentences in this window. N400 effect magnitude was calculated as the amplitude of response to ungrammatical sentences in the 300-500 ms window, subtracted from the amplitude of response to grammatical sentences in this window. Then, through subtracting the N400 effect magnitude from the P600
effect magnitude, response dominance could be calculated through the equation below. This process was repeated for each participant in each condition (see Tanner & Van Hell, 2014, for more details).

\[
RDI = \frac{(P600_{\text{Ungram}} - P600_{\text{Gram}}) - (N400_{\text{Gram}} - N400_{\text{Ungram}})}{\sqrt{2}}
\]

**RDI Results.** The effect magnitudes were significantly negatively correlated in the unique \((r=-.62, p<.005)\) and dissimilar \((r=-.69, p<.001)\) conditions – this indicates that participants with large N400 effects showed little positivity, and vice versa. Effect magnitudes in the similar condition were not significantly correlated. Scatterplots of RDI distributions for each condition are found in Figures 6, 7, and 8. In these scatterplots the solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered to have a positive RDI. In these scatterplots we see a continuous distribution of N400-dominant, biphasic, and P600-dominant processing. Biphasic RDI values indicate nearly equal levels of N400 and P600 dominance, whereas N400 or P600 dominance demonstrates positive or negative dominance across both time windows, respectively. These scatterplots show there is clear variability between participants for RDI bias in L2 processing, with some L2 learners showing a strong N400 bias and others showing a strong P600 bias. This individual variation in response dominance is averaged out in traditional ERP grand average analyses (See Figures 4A, 4B, and 4C).

RDI is a measure to investigate individual variation in processing dominance in the L2. Through correlations of individuals' RDI within and across conditions, we obtain a clearer picture of how individuals process ungrammatical sentences in the L2, in
various morphosyntactic categories. RDI scores in the three conditions were not significantly correlated -- only between similar and dissimilar conditions did a correlation trend towards significance ($r=.39, p=.08$). This indicates that participants did not have consistent N400 or P600 dominance in each condition, but rather displayed N400-dominant or P600-dominant processing differentially for each condition. Mean RDIs were calculated across participants in each condition. To sentences in the similar condition, there was a negative mean RDI bias (group mean RDI=$-0.92, SD=3.46$; 13 participants displaying negative RDI bias). In the unique and dissimilar conditions, there was a positive RDI bias (unique: group mean RDI=$1.27, SD=4.71$, and 10 participants displayed positive RDI bias; dissimilar: group mean RDI=$1.48, SD=4.91$, and 13 participants displayed positive RDI bias).

(6)
Figures 6, 7 and 8: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 6) Tense (similar), 7) Determiner Gender (unique), and 8) Determiner Number (dissimilar) conditions. The solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered those with a positive RDI.
The scatterplots presented in Figures 6-8 show there is clear variability between participants for RDI bias in L2 processing, with some L2 learners showing a strong N400 bias and others showing a strong P600 bias, with yet others showing no clear negative or positive processing dominance. In Appendix A you can find averaged ERPs for those L2 learners who demonstrated a negative RDI bias and L2 learners who demonstrated a positive RDI bias, in order to investigate topography and time course of processing by dominance.

**Response Magnitude Index**

As the RDI only provides information regarding the type of ERP response that is dominant for each individual (i.e., quality of response), a separate measure of magnitude of response (RMI: Response Magnitude Index) was calculated for each individual (i.e., quantity of response). Greater RMI can be thought to indicate greater neural response (Tanner, Inoue & Osterhout, 2014), or sensitivity to grammaticality, regardless of type of response (N400 or P600 dominant processing). In the RMI each individual's Euclidean distance from zero was calculated for each time window using the following equation:

\[
RMI = \sqrt{(N400_{Gram} - N400_{Ungram})^2 + (P600_{Ungram} - P600_{Gram})^2}
\]

Separate scatterplots of response magnitude for each condition are shown in **Figures 9, 10 and 11**: Tense (similar, Figure 9), Gender (unique, Figure 10) and Number (dissimilar, Figure 11).
Figures 9, 10 and 11: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 9) Tense (similar), 10) Determiner Gender (unique), and 11) Determiner Number (dissimilar) conditions.

Sentences in the similar condition elicited the greatest response magnitude (mean = 3.90; SD=1.94; range =0-38), followed by the unique condition (mean = 2.94; SD=1.93; range =0.01-23.3) and the dissimilar condition (mean = 2.87; SD=1.90; range =0-17.40).

RMI was not significantly correlated across the three conditions. In examining the relationship between RDI (a qualitative measure of L2 processing) and RMI (a quantitative measure of L2 processing), we found no significant correlation. We also examined correlations between RMI in the N400 and the P600 windows. Only the unique condition demonstrated a significant negative correlation between the N400 and the P600 windows ($r=-.55$, $p=.04$), with higher RMI in the N400 window correlating with lower RMI in the P600 window.

These findings indicate that response dominance in individual grammaticality processing is not governed by sensitivity to a condition (i.e., magnitude of response), as there are no correlations present between RDI and RMI measures. Thus, it appears that
the assumption of P600-like dominance indicating greater sensitivity to a violation, as has been commonly assumed, would appear to not be the case. The lack of correlations between RMI in the N400 and P600 windows also indicate that magnitude of response in the 300-500 ms window did not generally deviate from magnitude of response in the 600-800 ms window, except for in processing of the unique condition, where greater sensitivity to a grammatical violation in the 300-500 ms window was associated with smaller sensitivity in the 600-800 ms window.

**ERP**

**English**

Neurocognitive measures of L1 syntactic processing were collected – results are described below.

In Figure 12 the grand mean waveforms for processing of ungrammatical and grammatical sentences in the L1 are shown, for representative centro-parietal (CP) electrodes. These analyses average all three conditions (Tense (Auxiliary Omission), Reflexive, Subject-Verb).

In the LAN/N400 window of 300-500 ms, there are no significant main effects or interactions.

In the P600 time window of 600-800 ms there is a significant interaction of Grammaticality X Electrode, with ungrammatical sentences eliciting greater positivity than grammatical at parietal sites. At lateral sites there are no significant main effects or interactions.
Figure 12: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of L1 English grammatical (green) and ungrammatical (dashed red) sentences. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10µV.

Grand mean waveforms for the three English conditions, Tense, Reflexive, and Subject-Verb Agreement, across the 9 representative electrodes are presented in Figures 13A, 13B, and 13C, respectively.
Tense (English) Condition, Ungrammatical
Tense (English) Condition, Grammatical
13(B)

Reflexive Condition, Ungrammatical
Reflexive Condition, Grammatical

Graph showing brainwave patterns at different electrodes (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4) with marked differences between ungrammatical and grammatical conditions.
Figure 13: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences across three L1 English conditions: (13A) Tense, (13B) Reflexive, and (13C) Subject-Verb conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10µV.

Table 5 presents results from the omnibus ANOVA. In the 300-500 ms window there is a significant interaction of Grammaticality X Electrode, with ungrammatical sentences eliciting greater positivity than grammatical at parietal sites. At lateral sites there are no significant main effects or interactions.

In the 500-800 ms window there is a significant effect of Grammaticality at midline electrodes, with more positivity elicited to ungrammatical than grammatical sentences (mean amplitude difference =2.69µV, SE=.49). There is a significant interaction of Grammaticality X Electrode and of Grammaticality X Electrode X
Condition at midline electrodes, with largest positivity to ungrammatical sentences at Pz.

At lateral sites there are no significant main effects or interactions.

<table>
<thead>
<tr>
<th></th>
<th>300-500 ms</th>
<th>500-800 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality (1,19)</td>
<td>--</td>
<td>29.24***</td>
</tr>
<tr>
<td>Grammaticality X Condition (2, 38)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Electrode (2, 38)</td>
<td>4.85**</td>
<td>12.84***</td>
</tr>
<tr>
<td>Grammaticality X Condition X Electrode (4, 76)</td>
<td>--</td>
<td>2.85**</td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality (1, 19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition (2, 38)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority (1,19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Hemisphere (1, 19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality x Condition x Anteriority (2, 38)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Hemisphere (2, 38)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority X Hemisphere (1, 19)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Anteriority X Hemisphere (2,38)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Degrees of freedom are reported in parentheses.

*.05 < p < .1
** p < .05
*** p < .01

Table 5: F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for L1 English language processing in the 300-500 ms and 500-800 ms windows.

Subsequently, the same ANOVAs were conducted for each condition separately at midline electrodes, in the 500-800 ms window, as this is where interactions of Grammaticality X Condition X Electrode are found in the omnibus ANOVA.

**Tense Condition:** An interaction of Grammaticality X Electrode is found ($F(2,38)=13.80, p<.001$), with greatest positivity to ungrammatical sentences at the parietal (Pz) electrode.

**Reflexive Condition:** The ANOVAs yielded no significant effects.

**Subject-Verb Condition:** An interaction of Grammaticality X Electrode is found ($F(2,38)= 8.32, p<.001$), with greatest positivity to ungrammatical sentences at the central (Cz) and parietal (Pz) electrodes.
Response Dominance Index

As was observed in the neurocognitive processing of sentences in the L2, variability in individual processing of the L1 is present, with some individuals exhibiting more N400-like or P600-like processing than their peers – though with what appears as much less variability between subjects than present in their L2 processing. In order to examine how this variability manifests, we again looked at RDI, computing each individual’s mean amplitude activity in each condition, over the 9 representative CP electrodes, and calculating N400 and P600 effect magnitudes.

**RDI Results.** The effect magnitudes were significantly negatively correlated in each condition (Tense ($r=\ -0.78$, $p<.001$); Reflexive ($r=-.74$, $p<.001$); Subject-Verb ($r=-.66$, $p<.005$)) – this indicates that participants with large N400 effects showed little positivity, and vice versa. Scatterplots of RDI distributions for each condition are found in Figures 13, 14, and 15. In these scatterplots the solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered to have a positive RDI. In these scatterplots we see a continuous distribution of N400-dominant, biphasic, and P600-dominant processing. Biphasic RDI values indicate nearly equal levels of N400 and P600 dominance, whereas N400 or P600 dominance demonstrates positive or negative dominance across both time windows, respectively. The scatterplots above show there is much less variability between participants for RDI bias in L1 processing than in L2.

RDI is a measure to investigate individual variation in processing dominance in the L2. Through correlations of individuals' RDI within and across conditions, we obtain
a clearer picture of how individuals process ungrammatical sentences in the L2, in various morphosyntactic categories. RDI scores in the three conditions were not significantly correlated. This indicates that participants did not have consistent N400 or P600 dominance in each condition, but rather displayed N400-dominant or P600-dominant processing differentially for each condition. Mean RDIs were calculated across participants in each condition. Each condition was overall positively biased in RDI processing (Tense: group mean RDI=3.67, SD=4.12, 15 participants displaying positive RDI bias; Reflexive: group mean RDI=1.99, SD=3.82, 16 participants displaying positive RDI bias; and Subject-Verb: group mean RDI=2.53, SD=4.17, 15 participants displaying positive RDI bias).

(14)
Figures 14, 15 and 16: Scatterplots presenting effect magnitudes in the N400 and P600 windows for individuals across the three L1 English conditions: 14) Tense, 15) Reflexive, & 16) Subject-Verb conditions. The solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered those with a positive RDI.
The scatterplots presented in Figures 14-16 show there is less variability between participants for RDI bias in L1 processing, than L2, with most individuals presenting as P600-like processors. In Appendix B you can find averaged ERPs for those L1 learners who demonstrated a negative RDI bias and L1 learners who demonstrated a positive RDI bias, in order to investigate topography and time course of processing by dominance.

**Correlational analyses**

**Individual Difference Measures**

To examine how individual difference variables may affect response dominance and magnitude during L2 processing, correlations were calculated between measures that have been implicated to play a role in L2 learning (Cognitive abilities (working memory, executive functions and cognitive control), affect (attitude and motivation towards L2 learning), and proficiency (L1 proficiency, and L2 proficiency); for discussion of theoretical motivation, see Introduction), and measures of neurocognitive language processing (Response Dominance Index; Response Magnitude Index). In Table 6 mean individual difference scores, standard deviations and score ranges are presented for measures related to age of acquisition, L1 Proficiency (sensitivity to grammaticality, and exposure), L2 proficiency (sensitivity to grammaticality, self-ratings, and exposure). In Table 7 mean individual difference scores, standard deviations and score ranges are presented for performance on tasks measuring the cognitive abilities (working memory; executive functions and cognitive control).
<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similar</td>
<td>-.5</td>
<td>1.05</td>
<td>(-2.56 – 2.07)</td>
</tr>
<tr>
<td>Unique</td>
<td>-.62</td>
<td>1.23</td>
<td>-3.29 – 1.68</td>
</tr>
<tr>
<td>Dissimilar</td>
<td>-.24</td>
<td>1.22</td>
<td>-2.56 – 3.29</td>
</tr>
<tr>
<td><strong>Sensitivity to grammaticality in L2 (d')</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tense</td>
<td>4.83</td>
<td>2.87</td>
<td>1.68 – 8.6</td>
</tr>
<tr>
<td>Reflexive</td>
<td>6.67</td>
<td>2.71</td>
<td>2.56 – 8.6</td>
</tr>
<tr>
<td>Subject-Verb</td>
<td>5.95</td>
<td>3.03</td>
<td>1.68 – 8.6</td>
</tr>
<tr>
<td><strong>Accuracy in L2 Vocabulary (%)</strong></td>
<td>69.6</td>
<td>13.65</td>
<td>43.50-90.60</td>
</tr>
<tr>
<td><strong>Hr/day reading in L1</strong></td>
<td>1.93</td>
<td>.96</td>
<td>.50-3.00</td>
</tr>
<tr>
<td><strong>Hr/day reading in L2</strong></td>
<td>.59</td>
<td>.44</td>
<td>0-1</td>
</tr>
<tr>
<td>% of time exposed to L2</td>
<td>7.3</td>
<td>5.35</td>
<td>0-20</td>
</tr>
<tr>
<td>% of time would choose to read in L2</td>
<td>5.42</td>
<td>8.8</td>
<td>0-30</td>
</tr>
<tr>
<td>% of time would choose to speak in L2</td>
<td>7.55</td>
<td>11.38</td>
<td>0-40</td>
</tr>
<tr>
<td>Others judgment of L2 accentedness (1-10)</td>
<td>9.1</td>
<td>1.62</td>
<td>5-10</td>
</tr>
<tr>
<td><strong>Hr/week of L2 outside of classroom</strong></td>
<td>.58</td>
<td>.88</td>
<td>0-3</td>
</tr>
<tr>
<td><strong>Hr/week of homework/study in L2 outside of classroom</strong></td>
<td>1.95</td>
<td>.78</td>
<td>1-3</td>
</tr>
<tr>
<td><strong>Self-rated L2 Proficiency</strong> (Out of 10 – most)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking</td>
<td>3.6</td>
<td>1.19</td>
<td>2-6</td>
</tr>
<tr>
<td>Reading</td>
<td>4.95</td>
<td>1.39</td>
<td>2-7</td>
</tr>
<tr>
<td>Writing</td>
<td>4.2</td>
<td>1.91</td>
<td>1-7</td>
</tr>
<tr>
<td>Understanding</td>
<td>4.6</td>
<td>1.54</td>
<td>2-7</td>
</tr>
<tr>
<td>Accentedness</td>
<td>8.4</td>
<td>1.79</td>
<td>5-10</td>
</tr>
<tr>
<td><strong>AoA of L2 (years)</strong></td>
<td>13.84</td>
<td>3.42</td>
<td>6-22</td>
</tr>
</tbody>
</table>

**Table 6:** Mean individual difference scores, standard deviations and score ranges for variables related to Sensitivity to Grammaticality in the L2 and in the L1, Age of L2 Acquisition, L1 Exposure, L2 Exposure, and Self-rated L2 Proficiency.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rey Task</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copy Time (s)</td>
<td>252.5</td>
<td>110.93</td>
<td>120 – 518</td>
</tr>
<tr>
<td>Copy Raw Score</td>
<td>34.3</td>
<td>2.39</td>
<td>26 – 36</td>
</tr>
<tr>
<td>(out of 36)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate recall (%)</td>
<td>38.75</td>
<td>34.69</td>
<td>1 – 92</td>
</tr>
<tr>
<td>Delayed recall (%)</td>
<td>35.8</td>
<td>33.75</td>
<td>1 – 90</td>
</tr>
<tr>
<td>Recognition (%)</td>
<td>22.53</td>
<td>19.75</td>
<td>1 – 58</td>
</tr>
<tr>
<td><strong>TVIP Score (%)</strong></td>
<td>7.34</td>
<td>6.45</td>
<td>0 – 20</td>
</tr>
<tr>
<td><strong>O-Span Score</strong></td>
<td>46.1</td>
<td>5.27</td>
<td>33 – 54</td>
</tr>
<tr>
<td>(maximum score = 60)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CELF Score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(maximum score = 28)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forwards</td>
<td>10.45</td>
<td>2.28</td>
<td>7 – 14</td>
</tr>
<tr>
<td>Backwards</td>
<td>7.2</td>
<td>2.46</td>
<td>3 – 13</td>
</tr>
<tr>
<td>Total</td>
<td>17.65</td>
<td>4.13</td>
<td>12 – 25</td>
</tr>
<tr>
<td><strong>PPVT Score</strong></td>
<td>202.9</td>
<td>8.19</td>
<td>187 – 216</td>
</tr>
<tr>
<td><strong>Simon Score</strong></td>
<td>44.22</td>
<td>17.24</td>
<td>11.85 – 88.12</td>
</tr>
<tr>
<td><strong>ANT (Grand Average)</strong></td>
<td>524.28</td>
<td>67.58</td>
<td>443.81 – 689.5</td>
</tr>
</tbody>
</table>

Table 7: Mean individual difference scores, standard deviations and score ranges for performance on the Rey, O-Span, TVIP, PPVT, ANT, and Simon Tasks.

Table 8 presents correlations between language processing measures (RDI and RMI in the L2) and individual difference measures implicated in L2 learning (L1 and L2 proficiency, working memory, executive functions and cognitive control, attitude, and motivation).
<table>
<thead>
<tr>
<th></th>
<th>RDI (L2)</th>
<th>RMI (L2)</th>
<th>L1 Proficiency</th>
<th>L2 Proficiency</th>
<th>Attitude</th>
<th>Motivation</th>
<th>Working Memory</th>
<th>Cognitive Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDI L2</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMI L2</td>
<td>-.25</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 Proficiency</td>
<td>.22</td>
<td>-.04</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2 Proficiency</td>
<td>.34</td>
<td>-.19</td>
<td>-.05</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attitude</td>
<td>-.22</td>
<td>-.23</td>
<td>.05</td>
<td>-.15</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivation</td>
<td>.22</td>
<td>.06</td>
<td>.05</td>
<td>.56*</td>
<td>.16</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working Memory</td>
<td>.18</td>
<td>.26</td>
<td>-.28</td>
<td>-.20</td>
<td>-.52*</td>
<td>.04</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Cognitive Control</td>
<td>.04</td>
<td>-.16</td>
<td>-.26</td>
<td>.15</td>
<td>.26</td>
<td>-.18</td>
<td>-.17</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 8: Correlations between L1 Proficiency, L2 Proficiency, Attitude, Motivation, Working memory, Cognitive Control, RDI in L2, and RMI in L2.

There are no significant correlations present among Response Magnitude during L2 processing (i.e., sensitivity to ungrammaticality), and individual difference measures of L1 proficiency, L2 proficiency, attitude, motivation, working memory or executive functions and cognitive control. There are no significant correlations present between Response Dominance during L2 processing (i.e., quality of response, or how positive or negative dominant processing is in processing ungrammaticality), and individual difference measures of L1 proficiency, L2 proficiency, attitude, motivation, working memory or executive functions and cognitive control. Response Dominance during L2 processing and Response Magnitude during L2 processing were also not significantly correlated. Individual difference variables implicated in L2 processing were not significantly correlated to one another, except for the measures of attitude and of motivation towards L2 learning. Motivation towards L2 learning correlated positively with L2 Proficiency, indicating that participants who demonstrated higher Motivation towards L2 learning also demonstrated higher L2 Proficiency (as measured by offline L2 vocabulary accuracy, score in the TVIP task, self-reported proficiency in the L2, and
language history questionnaire responses regarding L2 exposure). Attitude towards L2 learning negatively correlated with working memory, indicating that participants who demonstrated a more positive attitude towards L2 learning scored lower on the working memory O-Span task.

**Child**

**Accuracy**

Mean accuracy was around chance when judging sentence grammaticality in the L2 (see Figure 17 and Table 9). The d-prime values of the similar, unique, and dissimilar conditions were -2.57, -.97, and -.68, respectively.

Using one-sample t-tests, for both grammatical and ungrammatical sentences combined, the mean accuracy of grammaticality judgment in each condition was compared to 50% -- what one would score if grammaticality judgment accuracy was perfectly at chance. In each condition performance was at chance (Similar: \( t(13)=.05, p=.96 \); Unique: \( t(13)=-.36, p=.73, \) and \( t(39)=.18, p=.86 \)).

An Accuracy X Condition ANOVA yielded no interaction, so no one condition received significantly more accurate responses than another.
Figure 17: Accuracy by condition in L2 Spanish, similar (Tense), unique (Gender) and dissimilar (Number) conditions.

<table>
<thead>
<tr>
<th></th>
<th>Accuracy: Ungrammatical Sentences</th>
<th>Accuracy: Grammatical Sentences</th>
<th>Overall Accuracy: Mean and SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similar</td>
<td>Mean = 33.57%; SD =24.10</td>
<td>Mean = 65.71%; SD =21.10</td>
<td>Mean = 49.64%; SD = 22.60</td>
</tr>
<tr>
<td>Dissimilar</td>
<td>Mean = 37.86%; SD =15.77</td>
<td>Mean = 64.29%; SD =19.88</td>
<td>Mean = 51.07%; SD = 17.83</td>
</tr>
<tr>
<td>Unique</td>
<td>Mean = 32.86%; SD =16.04</td>
<td>Mean = 62.14%; SD =26.75</td>
<td>Mean = 47.5%; SD = 21.39</td>
</tr>
<tr>
<td><strong>L1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tense</td>
<td>Mean = 85.00%; SD =9.13</td>
<td>Mean = 88.57%; SD =10.29</td>
<td>Mean = 86.79%; SD =9.71</td>
</tr>
<tr>
<td>Reflexive</td>
<td>Mean = 94.25%; SD =5.68</td>
<td>Mean = 65.71%; SD =17.42</td>
<td>Mean = 75%; SD = 12.65</td>
</tr>
<tr>
<td>Subject-Verb</td>
<td>Mean = 70.00%; SD =14.14</td>
<td>Mean = 92.14%; SD =8.59</td>
<td>Mean = 81.07%; SD = 11.37</td>
</tr>
</tbody>
</table>

Table 9: Mean accuracy in correctly identifying violations in grammaticality, broken down by grammatical and ungrammatical sentences in the L2 and L1.

In contrast, participants were above chance in their judgments to grammaticality violations in the L1 (see Table 9 and Figure 18). The d-prime values of the Tense, Reflexive, and Subject-Verb conditions were 2.23, .92, and 1.19, respectively.
In Figure 18, the accuracy by condition in L1 English Tense, Reflexive and Subject-Verb conditions is illustrated.

**ERP**

**Spanish**

In Figure 19, the grand mean waveforms for processing of ungrammatical and grammatical sentences in the L2 are shown, for a series of electrodes commonly implicated in N400 and P600 responses – centro-parietal (CP) electrodes: F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4. These analyses average all three conditions (Similar, Dissimilar, Unique).

In the LAN/N400 window of 300-500 ms there are no significant main effects or interactions.

In the P600 time window of 600-800 ms there are no significant main effects or interactions.
Figure 19: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of L2 Spanish grammatical (green) and ungrammatical (dashed red) sentences. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16µV.

Grand mean waveforms for the three critical Spanish conditions Similar (Tense), Dissimilar (Determiner Number), and Unique (Determiner Gender) across the 9 representative electrodes are presented in Figures 20A, 20B, and 20C, respectively.
20 (A)

Similar (Tense) Condition, Ungrammatical
Similar (Tense) Condition, Grammatical
Figure 20: Grand mean waveform from nine representative electrodes contrasting grammatical (green) and ungrammatical (dashed red) sentences across three L2 Spanish conditions: (20A) Tense (Similar), (20B) Determiner Number (Dissimilar), and (20C) Determiner Gender (Unique) conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16 µV.

Table 10 presents results from the omnibus ANOVA. In the 300-500 ms window, there are no significant main effects or interactions. In the 600-800 ms window, there are no significant main effects or interactions.
Subsequently, the same ANOVAs were conducted for each condition separately at midline and lateral electrodes, in the 300-500 ms and 600-800 ms windows – no significant main effects or interactions are present.

**Response Dominance Index**

Further inspection of individual waveforms indicates that not all participants exhibit the trend shown in the grand mean waveforms in Figures 20A, 20B, and 20C. As found in Tanner and Van Hell (2014), and participants reported above, certain individuals appear to exhibit an N400 while others appeared to exhibit a P600, to the same conditions.

We therefore computed a response-dominance index (RDI) for each participant to tease apart variability of positive or negative processing that appears to have been averaged out in grand mean processing (see Introduction & Response Dominance Index above.)

<table>
<thead>
<tr>
<th></th>
<th>300-500 ms</th>
<th>600-800 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Electrode</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Electrode</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Hemisphere</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Anteriority</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Hemisphere</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Degrees of freedom are reported in parentheses.

**Table 10:** F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for L2 Spanish language processing in the 300-500 ms and 600-800 ms windows.
**RDI Results.** The effect magnitudes were significantly negatively correlated in each condition (Similar: $r=-.89, p<.01$; Unique: $r=-.94, p=.001$; Dissimilar: $r=-.78, p<.05$). Scatterplots of RDI distributions for each condition are found in Figures 23, 24, and 25. In these scatterplots the solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered to have a positive RDI. In these scatterplots we see a continuous distribution of N400-dominant, and P600-dominant processing. Biphasic RDI values indicate nearly equal levels of N400 and P600 dominance, whereas N400 or P600 dominance demonstrates positive or negative dominance across both time windows, respectively. These scatterplots show there is clear variability between participants for RDI bias in L2 processing, though more so towards P600-like processing than N400. This individual variation in response dominance is averaged out in traditional ERP grand average analyses (See Figures 20A, 20B, and 20C).

RDI scores in the conditions were only significantly correlated between unique and dissimilar conditions ($r=.76, p<.05$). This indicates that P600-like processing was consistent for individuals in those two conditions. Mean RDIs were calculated across participants in each condition. Each condition was overall positively biased in RDI processing (similar: group mean RDI=2.09, $SD=6.93$, 4 participants displaying positive RDI bias; unique: group mean RDI=4.09, $SD=15.70$, 3 participants displaying positive RDI bias; and dissimilar: group mean RDI=2.39, $SD=9.36$, 4 participants displaying positive RDI bias).
(21)

![Similar](image1)

(22)

![Unique](image2)
Figures 21, 22 and 23: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 21) Tense (Similar) Condition, 22) Determiner Gender (Unique) condition, & 23) Determiner Number (Dissimilar) condition. The solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered those with a positive RDI.

The scatterplots presented in Figures 21-23 show there is clear variability between individuals in RDI. In the Appendix C you can find averaged ERPs for those child L2 learners who demonstrated a negative RDI bias and child L2 learners who demonstrated a positive RDI bias, in order to investigate topography and time course of processing by dominance.

**Response Magnitude Index**

RMI was calculated for each individual (i.e. quantity of response), in addition to RDI (i.e. quality of response) (see Response Magnitude Index in Adult Results for formula).
Figures 24, 25 & 26: Scatterplots demonstrating effect magnitudes in the N400 and P600 windows for individuals across the three L2 Spanish conditions: 24) Tense (Similar), 25) Determiner Gender (Unique), & 26) Determiner Number (Dissimilar) conditions.

Sentences in the unique condition elicited the greatest response magnitude (mean = 11.83; SD=10.65; range =2.60-33.98), followed by the dissimilar condition (mean = 8.06; SD=5.53; range =1.13-17.59) and the similar condition (mean = 6.52; SD=3.41; range =2.51-11.42).

RMI was not significantly correlated across the three conditions. In examining the relationship between RDI (a qualitative measure of L2 processing) and RMI (a quantitative measure of L2 processing), there is a positive correlation present only in the unique condition ($r=.82, p=.02$), suggesting P600-like processing is associated with more sensitivity to ungrammaticality in the unique condition. We also examined correlations between RMI in the N400 and the P600 windows. All conditions demonstrated a significant positive correlation between the N400 and the P600 windows (similar: $r=.80, p=.03$; unique: $r=.97, p<.001$; dissimilar: $r=.87, p=.01$), with higher RMI in the N400 window correlating with higher RMI in the P600 window.

ERP

English
Neurocognitive measures of L1 syntactic processing were collected – results are described below.

In Figure 27 the grand mean waveforms for processing of ungrammatical and grammatical sentences in the L1 are shown, for representative centro-parietal (CP) electrodes. These analyses average all three conditions (Tense (Auxiliary Omission), Reflexive, Subject-Verb).

In the LAN/N400 window of 300-500 ms, there are no significant main effects or interactions.

In the P600 time window of 600-800 ms, there are no significant main effects or interactions.

(27)

**Figure 27:** Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of L1 English grammatical (green) and ungrammatical (dashed red) sentences. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical
stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16µV.

Grand mean waveforms for the three English conditions, Tense, Reflexive, and Subject-Verb Agreement, across the 9 representative electrodes are presented in Figures 28A, 28B, and 28C, respectively.
Reflexive Condition, Ungrammatical
Reflexive Condition, Grammatical
Figure 28: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences across three L1 English conditions: (28A) Tense, (28B) Reflexive, and (28C) Subject-Verb conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16µV.
Table 11 presents results from the omnibus ANOVA. In the 300-500 or 500-800 ms windows, there are no significant main effects or interactions.

<table>
<thead>
<tr>
<th></th>
<th>300-500 ms</th>
<th>500-800 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality (1,6)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition (2, 12)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Electrode (2, 12)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Electrode (4, 24)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality (1, 6)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition (2, 12)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority (1,6)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Hemisphere (1, 6)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Anteriority (2, 12)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Hemisphere (2, 12)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority X Hemisphere (1, 6)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Condition X Anteriority X Hemisphere (2,12)</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Degrees of freedom are reported in parentheses.

* .05<\(p<.1\)

** \(p<.05\)

*** \(p<.01\)

Table 11: F-statistics from the omnibus grand mean ANOVA on mean amplitude measures for L1 English language processing in the 300-500 ms and 500-800 ms windows.

Subsequently, the same ANOVAs were conducted for each condition separately in the 300-500, and 500-800 ms windows, no significant main effects or interactions are present.

**Response Dominance Index**

As was observed in the neurocognitive processing of sentences in the L2, variability in individual processing of the L1 is present, with some individuals exhibiting more N400-like or P600-like processing than their peers. This variability is again possibly averaged out in traditional grand mean waveforms, such as found in Figure 27, resulting in a possibly inaccurate view of L1 sentence processing. In order to examine how this variability manifests, we again looked at RDI, computing each individual’s mean amplitude activity in each condition, over the 9 representative CP electrodes, and calculating N400 and P600 effect magnitudes.
**RDI Results.** The effect magnitudes were significantly negatively correlated in only the Tense condition ($r=-.77, p<.05$) – this indicates that participants with large N400 effects showed little positivity, and vice versa, in the Tense condition. Scatterplots of RDI distributions for each condition are found in Figures 29, 30, and 31. In these scatterplots the solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered to have a positive RDI. In these scatterplots we see a continuous distribution of N400-dominant, biphasic, and P600-dominant processing. Biphasic RDI values indicate nearly equal levels of N400 and P600 dominance, whereas N400 or P600 dominance demonstrates positive or negative dominance across both time windows, respectively.

RDI is a measure to investigate individual variation in processing dominance in the L2. Through correlations of individuals' RDI within and across conditions, we obtain a clearer picture of how individuals process ungrammatical sentences in the L2, in various morphosyntactic categories. RDI scores in the three conditions were not significantly correlated. This indicates that participants did not have consistent N400 or P600 dominance in each condition, but rather displayed N400-dominant or P600-dominant processing differentially for each condition. Mean RDIs were calculated across participants in each condition. Participants were overall negatively biased in RDI in Tense and Subject-Verb processing (Tense: group mean RDI=$-1.58$, $SD=9.51$, 4 participants displaying negative RDI bias; Subject-Verb: group mean RDI: $-0.09$, $SD=6.84$, 3 participants displaying negative RDI bias. In processing of grammaticality in
reflexive sentences, participants displayed positive bias (group mean RDI=0.20, SD=7.96, 3 participants displaying positive RDI bias).

(29)

![Tense Diagram]

(30)

![Reflexive Diagram]
Figures 29, 30 & 31: Scatterplots presenting effect magnitudes in the N400 and P600 windows for individuals across the three L1 English conditions: 29) Tense, 30) Reflexive, & 31) Subject-Verb conditions. The solid line indicates the best-fit line from the correlation analysis of each condition. The dashed line is representative of equal P600 and N400 effect magnitudes. The individuals above the dashed line are presumed to have a negative RDI, while those below are considered those with a positive RDI.

In the Appendix D you can find averaged ERPs for child L1 learners who demonstrated a negative RDI bias and child L1 learners who demonstrated a positive RDI bias, in order to investigate topography and time course of processing by dominance.
Chapter 4: DISCUSSION

Overview

In this study I examined sentence processing in the L1 and L2 in two developmentally distinct groups of intermediate L2 learners, children and adults. Using ERP and behavioral techniques, I investigated the influence of similarity of morphosyntax in the L1 and the L2 on the processing of morphosyntax in the L2, as well as the impact of individual difference variables on L2 processing, including working memory, executive functions and cognitive control, attitude and motivation towards L2 learning, and proficiency in the L1 and L2. The results for child and adult learners will be discussed separately.

Adult Language Processing

In this section I discuss implications of findings for adult language processing, in both the L1 and the L2.

L2 Processing:

The Competition and Proficiency-based Models are two theories of how L2 learning and processing occur. In this section, I will focus on how my results patterns with predictions of these theories, as well as existing research.

Competition Model:

To investigate the relationship between morphosyntax in the L1-L2 on processing of morphosyntax in the L2, participants were presented with grammatical and ungrammatical sentences in the L2 that were similar in morphosyntax between the L1 and L2, unique to the L2, and dissimilarly expressed in the L1 and L2. Through ERP and
behavioral measures, we were able to fully test the assertions of a widely-cited theory in the field: the Competition Model (see Introduction). According to predictions of the Competition Model, intermediate learners should show significant sensitivity to ungrammaticality in the similar condition (a (LAN+) P600), an attenuated P600 to ungrammatical sentences in the L2 unique condition, and a very small, or nonexistent, sensitivity to structures dissimilar in morphosyntax between the L1 and L2.

We found no LAN/N400-like processing in the 300-500 ms window for any of the three conditions (as in Figures 5A, 5B and 5C). In the 600-800 ms window, there was only a P600 present when processing ungrammaticalities in the dissimilar condition. This finding does not align with predictions from the Competition Model highlighted above, where features of morphosyntax dissimilarly manifested between an L1 and L2 are expected to be in competition, and elicit least sensitivity to ungrammaticality, when compared to structures similarly manifested in L1 and L2 or unique to the L2. These findings also conflict with those of the closely related study of Tokowicz and MacWhinney (2005) who found an attenuated P600 to processing of ungrammatical sentences in the similar condition, a P600 to processing of ungrammatical sentences in the unique condition, and no effect in the dissimilar condition. However, even in the Tokowicz and MacWhinney (2005) study, findings do not fully align with the predictions of the Competition Model, as it is the unique condition rather than similar that elicits a robust P600 in response to grammaticality violations. The difference in findings between the present study and Tokowicz and MacWhinney (2005) may be partially related to differences in proficiency between the two L2 learner groups. In Tokowicz and MacWhinney (2005), the L2 learners ranged anywhere from one to four semesters of
university Spanish at the time of testing. The L2 proficiency levels of the learners in the present study were more uniform, and were all tested at an intermediate level, where learners may have had more time to build cues dissimilarly used between the L1 and L2, reducing L1 effects on L2 processing, and allowing for more “native-like” processing of dissimilar structures in L2.

How do our findings compare to those of other ERP studies testing L2 morphosyntactic processing? It should be noted that a substantial number of studies in this literature have tested whether L2 learners can display “native-like” processing in the L2, and fewer studies explicitly manipulated cross-linguistic similarity between L1 and L2 to examine potential cross-linguistic effects to test transfer effects and the predictions of the Competition Model. A very recent review study by Caffara et al. (2015)\(^2\) discusses studies that have examined the effects of L1-L2 morphosyntactic similarity or dissimilarity on L2 processing. Tables 12-14 present a subset of these studies that have explicitly tested the influence of cross-linguistic similarity of L1-L2 in Tense-Marking, Gender Agreement, and Number Agreement, on L2 processing. While most studies have focused on highly proficient L2 speakers, there are a few which have focused on lower proficiency learners – these are also included in Tables 12-14.

In Table 12, the results of 6 studies are presented in which Tense-Marking was used to investigate L2 processing. Highly proficient L2 learners processing Tense-Marking as a structure *similarly* represented between the L1 and L2, generally exhibited P600 effects (as well as a N400+ P600 in Weber and Lavric, 2008). In other studies

---

\(^2\) In this review Caffara et al. (2015) did not differentiate between conditions unique to an L2, and conditions dissimilar between an L1-L2, magnitude of response (e.g., attenuated P600 versus P600) or differences in proficiency measurements, so power could be optimized in analyses.
where participants were of low or intermediate proficiency, sensitivity to Tense-Marking was attenuated (Tokowicz & MacWhinney, 2005) or not present (the present study), indicating similarity does not always result in sensitivity for low proficiency learners, as is predicted by the Competition Model. In a study in which Tense-Marking is dissimilarly manifested between two languages (i.e., -- Dutch L1 – English L2), testing highly proficiency L2 learners (Brenders et al., in prep), a LAN + P600 was observed. In sum, the six Tense Marking studies that are currently available do not show a clear pattern as to how the similarity-dissimilarity of tense marking between L1 and L2 affect morphosyntactic processing in the L2. In order to speak to how proficiency or cross-linguistic influence in Tense-Marking impact L2 processing, we need more L2 processing studies that focus on low or intermediate proficiency learners, and studies with highly proficient L2 speaker from whom Tense-Marking is a dissimilar construct.

<table>
<thead>
<tr>
<th>Study</th>
<th>Explicit testing of competition?</th>
<th>Structure tested</th>
<th>Proficiency</th>
<th>Similar</th>
<th>Dissimilar$^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokowicz and MacWhinney (2005)</td>
<td>X</td>
<td>Tense Marking</td>
<td>Low</td>
<td>attenuated P600</td>
<td></td>
</tr>
<tr>
<td>Sabourin and Stowe (2008)</td>
<td></td>
<td>Tense Marking</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td>Weber and Lavric (2008)</td>
<td></td>
<td>Tense Marking</td>
<td>High</td>
<td>N400 + P600</td>
<td></td>
</tr>
<tr>
<td>Schmidt-Kassow et al. (2011)</td>
<td></td>
<td>Tense Marking</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td>Brenders et al. (in prep)</td>
<td>X</td>
<td>Tense Marking</td>
<td>High</td>
<td>LAN + P600</td>
<td></td>
</tr>
<tr>
<td>Present Study</td>
<td>X</td>
<td>Tense Marking</td>
<td>Intermediate</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Findings from 6 studies on L2 learning contrasting findings based on cross-linguistic similarity of tense marking structure, and proficiency.

In Table 13 results of 11 studies are presented in which Number Agreement was used to investigate L2 processing. In the studies where learners were of low proficiency (the present study; Lemhöfer et al., 2014; Ojima et al., 2005; Tokowicz & MacWhinney,

---

$^3$ *Reflects a structure judged unique rather than dissimilar in Competition Model hierarchy.
dissimilarly represented across the L1 and L2. ERPs either reflected no sensitivity (Ojima et al., 2005; Tokowicz & MacWhinney, 2005) or a LAN + P600 effect (Lemhöfer et al., 2014). In studies in which participants were of high proficiency and number agreement was dissimilarly represented across the L1 and L2 (Bañon et al., 2014; Chen et al., 2007; Dowens et al., 2010, 2011; Ojima et al., 2005; Xue et al., 2013), participants showed either a LAN, LAN + P600, P600, or N400 in processing ungrammaticality. In studies where participants were of high proficiency and number agreement was similarly manifested between L1 and L2, participants largely showed a P600 in processing ungrammaticalities (Bañon et al., 2014; Brenders et al., in prep; Lemhöfer et al., 2014; Tanner et al., 2014), or an N400 + P600 pattern (Tanner et al., 2014). These findings demonstrate that low proficiency L2 learners (and highly proficient L2 speakers) demonstrate highly variable responses to violations of number agreement when this structure is dissimilarly manifested in L1-L2. However, more studies are needed of intermediate learners, as the present study appears the only study targeting intermediate-level L2 learners.
<table>
<thead>
<tr>
<th>Study authors</th>
<th>Explicit testing of competition?</th>
<th>Structure tested</th>
<th>Proficiency</th>
<th>Similar</th>
<th>Dissimilar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ojima et al. (2005)</td>
<td></td>
<td>Number Agreement</td>
<td>High</td>
<td>LAN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Low</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Tokowicz and MacWhinney (2005)</td>
<td>X</td>
<td>Number Agreement</td>
<td>Low</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Chen et al. (2007)</td>
<td>X</td>
<td>Number/Person Agreement</td>
<td>High</td>
<td>--*</td>
<td></td>
</tr>
<tr>
<td>Dowens et al. (2010)</td>
<td></td>
<td>Number Agreement</td>
<td>High</td>
<td>(LAN+)</td>
<td>P600</td>
</tr>
<tr>
<td>Dowens et al. (2011)</td>
<td></td>
<td>Number Agreement</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td>Xue et al. (2013)</td>
<td></td>
<td>Number/Person Agreement</td>
<td>High</td>
<td>P600</td>
<td>N400</td>
</tr>
<tr>
<td>Bañon et al. (2014)</td>
<td></td>
<td>Number Agreement</td>
<td>High</td>
<td>(LAN+)</td>
<td>P600</td>
</tr>
<tr>
<td>Lemhöfer et al. (2014)</td>
<td></td>
<td>Number Agreement</td>
<td>High</td>
<td>P600</td>
<td>LAN+P600</td>
</tr>
<tr>
<td>Tanner et al. (2014)</td>
<td></td>
<td>Number/Person Agreement</td>
<td>High</td>
<td>N400+</td>
<td>P600</td>
</tr>
<tr>
<td>Brenders et al. (in prep)</td>
<td>X</td>
<td>Number/Person Agreement</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td>Present Study</td>
<td>X</td>
<td>Number Agreement</td>
<td>Intermediate</td>
<td>P600</td>
<td></td>
</tr>
</tbody>
</table>

Table 13: Findings from 11 studies on L2 learning contrasting findings based on cross-linguistic similarity of number agreement, and proficiency.

Finally, in Table 14 results of 9 studies are presented in which gender agreement was used to investigate L2 processing. In studies with lower proficiency participants, when the structure was similarly manifested between L1-L2 (Lemhöfer et al., 2014), no sensitivity was reported. When this structure was unique to the L2 as in the present study and Tokowicz and MacWhinney (2005), no sensitivity was found, or a P600 was reported, respectively. In participants with high proficiency, where gender agreement was similarly expressed in the L1 and L2, participants demonstrated a P600 (Foucart & Frencke-Mestre, 2011; Sabourin & Stowe, 2008), or no sensitivity (Sabourin & Stowe, 2008). If structures were dissimilarly represented in the L1-L2, high proficiency learners showed a range of effects: a LAN+P600 (Dowens et al., 2010), P600 (Bañon et al., 2014; Dowens et al., 2011; Foucart & Frencke-Mestre, 2012), ELAN (Foucart & Frencke-Mestre, 2011) or an N400 (Foucart and Frenck-Mestre, 2012). These findings show that there seems to be much variability of processing gender agreement in L2, even when the
same structure is used, as a function of both proficiency and cross-linguistic overlap. A consistent pattern does not emerge from these studies. However, finding a lack of sensitivity, even in highly proficient learners, in a condition where there is L1-L2 overlap, clearly goes against the predictions of the Competition Model (Lemhöfer et al., 2014).

<table>
<thead>
<tr>
<th>Study authors</th>
<th>Explicit testing of competition?</th>
<th>Structure tested</th>
<th>Proficiency</th>
<th>Similar</th>
<th>Dissimilar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokowicz and MacWhinney (2005)</td>
<td>X</td>
<td>Gender Agreement</td>
<td>Low</td>
<td></td>
<td>P600*</td>
</tr>
<tr>
<td>Sabourin and Stowe (2008)</td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Dowens et al. (2010)</td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>(LAN+) P600</td>
<td></td>
</tr>
<tr>
<td>Dowens et al. (2011)</td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td>Foucart and Frenck-Mestre (2011)</td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>ELAN</td>
<td></td>
</tr>
<tr>
<td>Foucart and Frenck-Mestre (2012)</td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>N400</td>
<td></td>
</tr>
<tr>
<td>Bañon et al. (2014)</td>
<td></td>
<td>Gender Agreement</td>
<td>High</td>
<td>P600</td>
<td></td>
</tr>
<tr>
<td>Lemhöfer et al. (2014)</td>
<td></td>
<td>Gender Agreement</td>
<td>Low</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Present Study</td>
<td>X</td>
<td>Gender Agreement</td>
<td>Intermediate</td>
<td>--*</td>
<td></td>
</tr>
</tbody>
</table>

Table 14: Findings from 9 studies on L2 learning contrasting findings based on cross-linguistic similarity of gender agreement, and proficiency.

What can we conclude from the discussion of our findings and those of other studies when structures are considered in terms of L1-L2 similarity, L2 uniqueness, and L1-L2 dissimilarity (although this was not necessarily the original focus of these studies)? It is not a language specific syntactic feature of a structure that is causing our results, as sensitivity (or lack thereof) has been found in various learner groups for the same syntactic manipulation. In processing tense marking, the Competition Model would
predict a (LAN+) P600 effect for L2 learners for whom these structures are similar between L1-L2, and no sensitivity or small sensitivity when dissimilar. However, different sensitivity profiles emerge in other studies, with a mix of no sensitivity, P600, attenuated P600, and N400-P600 were found.

In the number agreement condition, for which L1-L2 structures were dissimilar in our study, we found a P600 (which goes against the Competition Model as no effects were found for processing in the similar or unique conditions). Our finding is not an isolated one -- if we relate this to other studies testing novice or intermediate learners for whom number agreement is expressed dissimilarly across the two languages we find that LAN + P600 effects have been found for low proficiency L2 learners (Lemhöfer et al., 2014). It is worth noting that the present study also matches Lemhöfer et al. (2014) in that no sensitivity was found in L2 processing for a structure similarly used in the L1-L2, directly in contrast to predictions of the Competition Model.

In sum, the above review of studies testing L2 morphosyntactic structures that are similar or dissimilar across L1-L2, or unique to the L2, does not yield a conclusive patterns regarding the predictions of the Competition Model, or the influence of cross-linguistic transfer on L2 processing more generally, other than that the pattern of ERP responses is highly variable, and particularly so for low and intermediate level L2 learners. In the next section, the present findings will be related to those of other studies, in the context of the Proficiency-based Model (Steinhauer et al., 2009).

Proficiency-based Model:

The second model that theoretically motivated the present study, the Proficiency-based Model (Steinhauer et al., 2009; see Introduction), posits that regardless of L1-L2
overlap, individuals at the same proficiency level will present the same neural responses in processing ungrammaticality; if correct, learners of intermediate proficiency should demonstrate attenuated and delayed P600s to ungrammaticality (and possibly an N400 effect if the intermediate learners still rely more heavily on semantics and pragmatics than syntax during L2 processing), irrespective of L1-L2 (dis)similarity. Our adult intermediate L2 learners demonstrated no LAN/N400 effect for any of the three conditions (see Figures 5A, 5B and 5C), and a P600 only in the dissimilar condition (but only present at lateral sites, not widely-distributed, see Figure 3). Additional evidence against the Proficiency-based model is found in Tables 12-14 above, where individuals of the same proficiency level do not always process the L2 with the same patterns, and the pattern varies widely across different L2 structures for low and intermediate proficient L2 learners as well as highly proficient L2 speakers. Additionally, higher proficiency speakers sometimes demonstrate N400-like processing in the L2 (Foucart & Frencke-Mestre, 2012; Tanner et al. (2014); Xue et al., 2013).

To conclude, in Tables 12-14 above, we find that there does not seem to be a consistent pattern in the processing of ungrammaticality in the L2 based on just proficiency effects (or effects of cross-linguistic (dis)similarity, for that matter). Regardless of structure examined, or L1-L2 overlap, or participant proficiency, participants showed either no sensitivity or sensitivity through LAN, ELAN, attenuated P600, LAN + P600, N400 or N400+P600 effects to sentences with similar or dissimilar representations in the L1-L2. If it is the case that neither the Competition Model nor the Proficiency-based model accounts for the findings of this study, or can explain the large variability in findings of studies that have tested L2 processing in a variety of L1-L2
overlap, what could account for these? To look into this further, through examining individual waveforms of L2 processing, we found that learners actually varied greatly in amplitude of processing to ungrammatical sentences in the L2, within and across conditions. Before discussing these findings, we first discuss the pattern of results of L1 morphosyntactic processing.

**L1 Processing:**

Participants in this study were tested in L1 processing as well as L2. L1 processing was tested in order to obtain neurocognitive signatures of the L2 learners in their native language, L1. Particularly when the pattern of findings in L2 deviates from the typical neurocognitive signatures of syntactic processing, as indeed seems to be the case in the present intermediate learners, it is important to know that this pattern is not driven by a more general deviant pattern in syntactic processing that extends to their L1 (Tanner et al., 2014; Tanner & Van Hell, 2014). Based on earlier ERP studies of L1 morphosyntactic processing, we expected that participants would demonstrate one of the following patterns that have been found in earlier studies: 1) a LAN+P600 biphasic responses (e.g., Friederici, Hahne, & Mecklinger, 1996; Hagoort, Brown, & Groothusen, 1993), thought to reflect the automatic detection of syntactic violations and later attempts to reanalyze or repair errors, 2) an P600, in the absence of the LAN effect preceding a P600 (e.g., Hagoort, Wassenaar, & Brown, 2003; Nevins, Dillon, Malhotra, & Phillips, 2007; Osterhout, Mckinnon, Bersick, & Corey, 1996), or 3) a P600 preceded by an anterior negativity (AN) that is not left-lateralized (Dillon, Nevins, Austin & Phillips, 2012; Osterhout & Nicol, 1999).
In the present study, I observed a greater negativity in anterior sites to ungrammatical sentences relative to grammatical sentences in the 300-500 ms time window, which could be classified as an anterior negativity (AN). This anterior negativity was followed by a typical P600 (greater positivity to ungrammatical than to ungrammatical sentences in the posterior electrodes in the 500-800 ms window). The finding that participants process ungrammatical sentences in the L1 with an AN + P600 effect in grand mean analyses confirms that processing is as predicted for native L1 speakers, with an early negativity that could reflect automatic processing, and a later positivity that is thought to reflect reanalysis and repair. These typical neural signatures of L1 processing also indicate that the unexpected pattern observed in their L2 morphosyntactic processing is not driven by a more general deviance in language processing. Remarkably, many studies on L2 morphosyntactic processing only examine L2 processing, and do not include L1 processing.

Studying individual variation in L1 processing within a given sample by assessing the Response Dominance Index (RDI) and the Response Magnitude Index (RMI), is a new technique, and is employed in a growing number of studies (Meulman et al., 2014; Tanner et al., 2014; Tanner & Van Hell, 2014; see Introduction for more details) in L1 processing. Individual variation, in particular in N400-dominance and P600-dominance, has been found to influence grand mean averages. In the present data, we also observed variation in dominance, but the N400 dominance did not impact the typical P600 grand average signature (unlike what was observed in the L2, as discussed above).
RDI and RMI processing in L1 and L2:

To examine how variation manifested in L2 processing, participant processing was divided into RDI and RMI, for dominance and sensitivity respectively (see Results for formulas). For RDI, in the similar condition, participants overall displayed more N400-like processing (13 participants negative). In the unique and dissimilar conditions, participants overall displayed more P600-like processing (Unique: 10 participants positive; Dissimilar: 13 participants positive). N400 and P600 windows negatively correlated for the unique and dissimilar conditions, indicating lack of biphasic processing. For RMI, sentences in the similar condition elicited the greatest RMI, followed by the unique and dissimilar conditions. Both RDI and RMI were not correlated across their conditions, indicating differential dominance and sensitivity by condition; RDI and RMI were not correlated with one another.

Finding that participants differed greatly in both dominance and magnitude of response within a sample could inform why there is so much variation in reported L2 processing across learners. If all learner groups have such variance, it is likely that through the grand averaging traditional in ERP research, reported findings are in fact misrepresentative of true learner patterns.

While most studies have not in fact examined individual variation, there are some that have taken this into account (Meulman et al., 2014; Morgan-Short et al., 2012; Inoue and Osterhout, 2012; Tanner et al., 2013). In Tanner et al. (2013), native German and native English learners of college-level German (at first-year or at third-year) were tested in processing of a structure similarly manifested in L1-L2. Upon traditional grand average analyses, they found that native German as well as third-year learners had a
significant P600 when processing ungrammaticality. In contrast, first-year learners showed a biphasic N400-P600 pattern in processing. Upon division into RDI, they found the same pattern as presented above – individuals varied along N400 and P600-like processing, and the degree to which one was N400-like negatively correlated to P600-like processing, as found in the present study. This meant participants were not actually biphasic in processing, but through averaging were misrepresented as such.

In Meulman et al. (2014) highly proficient L1 Romance language immersed learners, of L2 Dutch, were tested on the processing of a structure dissimilar to that of the L1 (grammatical gender), and of non-finite verbs (shown to be relatively simple for L2 learners), and compared in their processing to native L1 speakers. In processing the “simpler” verbs, a biphasic N400-P600 was found in L2 learners, as well as native speakers. To ungrammaticality in gender, L2 learners showed no sensitivity, while native speakers showed a P600. In examining RDI, both native speakers and L2 learners replicated the biphasic pattern in verb processing. In native speaker processing of gender, RDI largely patterned with positivity dominance, but L2 learners did not show significant sensitivity to ungrammaticality in gender. This finding contrasts with the present study of intermediate learners, and the RDI in first-year L2 learners in Tanner et al. (2013), where variation in RDI was substantial. This also suggests that individual variation in RDI may decrease as L2 proficiency increases.

However, this trend is contradicted in Tanner and van Hell (2014). In this study monolingual English speakers were tested on neurocognitive processing of ungrammaticality in morphosyntax in English. These participants showed variability in RDI not captured through the traditional grand average waveforms – where a biphasic
LAN + P600 was reported in grand mean waveforms, most participants actually displayed either N400 dominant or P600 dominant responses. From this we can conclude that even with high proficiency, significant variation exists among individuals.

In a final study, Tanner et al. (2014) examined the influence of a series of individual difference measures related to L2 learning in a group of proficient L1 Spanish – L2 English learners, processing L2 sentences of a construction similar in the L1-L2. As found in previous studies (Tanner & van Hell, 2014), biphasic N400-P600 processing in the grand mean waveforms was not representative of variation as found in RDI. In addition to RDI, Response Magnitude was also calculated, to measure sensitivity to structures in the L2 (see Introduction for review). Through including RDI and RMI as dependent measures in regression models, they tested the influence of a series of individual difference measures (AoA, Length of Residence, L2 use, proficiency, and motivation to speak like a native) on RDI or RMI. For RMI only offline L2 proficiency showed a relation, with higher L2 proficiency corresponding to higher RMI. In RDI, lower AoA and higher motivation corresponded to more P600-like processing.

The findings of the present study, and others reviewed above, have implications for future neurocognitive research. Variability in processing patterns among L2 learners (as well as L1 speakers) begs the question of how valid it is to consider the P600 as the golden-standard for “native-like” syntactic processing. Findings that even in L1 processing, as well as advanced L2 processing, some participants pattern with N400 requires a re-evaluation of the functional significance of ERP components. Moreover, low or intermediate level L2 learners' processing patterns are highly variable, so high
variability (rather than a consistent patterns) appears to be fundamental pattern of L2 morphosyntactic processing.

In the next section, I will discuss the pattern of results of individual difference variables that were studied to examined what variables may influence L2 morphosyntactic processing in this group of intermediate L2 learners.

**Individual Difference Measures and L2 Processing:**

As discussed above, there is significant variability in groups of speakers and learners in processing of both the L1 and L2, respectively. To examine what may account for this variation, several studies have investigated the role of individual difference factors in language learning using behavioral and ERP measures, examining association of factors such as working memory, L2 proficiency, executive functions and cognitive control, attitude, and motivation. In the present study the relationship of these factors, as well as L1 proficiency, were investigated in correlation with factors of L2 processing (RDI and RMI).

**Working memory** has been shown to play a role in L1 processing in many studies (Kim & Christianson, 2007; Just & Carpenter, 1992), with the relatively uniform finding that higher working memory correlates with better language processing abilities, particularly in sentences that are grammatically complex. The “more is better” hypothesis (see Introduction) suggests that higher WM is needed for more automatic processing. If this were correct, we would expect to see positive correlation between working memory and L2 processing abilities (e.g., Dussias and Piñar, 2010). In some studies N400 effects in L2 morphosyntactic processing have been associated with learners of lower working memory capacity (Nakano et al., 2010). However, in other studies working memory span
has not shown any correlation to dominance of response during L2 processing. In the present study we found no correlations between working memory and RDI and RMI in the L2. Both the present study and Tanner and Van Hell (2014) used non-verbal tasks to measure working memory, whereas studies (e.g., Nakano et al., 2010; Ye and Zhou, 2008) which have found influence of working memory on neurocognitive processing patterns have used tasks linguistic in nature (such as reading-span). As discussed by Miyake and Friedman (1998), working memory tasks that are linguistic in nature do not only measure working memory, but also linguistic processing, so these tasks may be a less accurate reflection of working memory performance.

**L1 proficiency.** While native language abilities are often assumed to be uniform, there is actually much variability present in L1 proficiency, and corresponding ERP patterns (Pakulak and Neville, 2010). Variability in L1 proficiency likely influences L2 proficiency and processing patterns, particularly if cross-linguistic transfer occurs as predicted by the Competition Model. We would predict that with variation in L1 proficiency in the current study, L2 processing should correlate with higher RDI. We did not find any correlation of L1 proficiency with L1 or L2 processing. This mirrors the fact that, while differential processing was found in lower proficiency L1 speakers tested in Pakulak and Neville (2010), in Tanner and Van Hell (2014) (see RDI section above), individual difference measures of L1 proficiency of a typical undergraduate student population were not associated with N400 or P600 dominance in L1 processing. Unlike the Pakulak and Neville (2010) study, the present study, as well as Tanner and Van Hell (2014), did not create low and high proficiency groups. It may be that the range in L1 proficiency of university participants is not substantial enough to impact the L2.
**L2 Proficiency:** Studies using longitudinal testing have shown that beginning proficiency N400 dominant participants later show P600-like effects with increased proficiency (McLaughlin et al., 2010). The Proficiency-based Model predicts that as proficiency in the L2 increases, neurocognitive processing becomes more “native-like”, for L2 learners. If this were correct, we would expect to see a positive correlation of RDI with proficiency. There was no correlation found between L2 proficiency and RMI and RDI in L2 processing. However, P600-like processing is not necessarily more likely as proficiency increases. An equivalent level of P600-like dominance has been found in participants across proficiency levels. In Tanner and Van Hell (2014), Tanner et al. (2014), and Tanner et al. (2013), similar patterns of variability in RDI were found in processing of native speakers, proficient immersed L2 learners, and novice L2 learners, respectively.

**Cognitive Control:** Greater cognitive control is thought to result in more successful L2 processing (Luo, Luk, Bialystok, 2011; Novick et al., 2005; Ye & Zhou, 2008). If greater cognitive control leads to better conflict resolution in processing of ungrammatical sentences, we would predict that higher cognitive control would correlate with RDI. In this study, there were no correlations found between cognitive control ability and L2 processing. Both the present study and Tanner and Van Hell (2014), where no association of cognitive control with RDI was found, used a task purely morphosyntactic in nature, whereas studies which have found correlation of cognitive control with neurocognitive processing patterns have used manipulations such as garden path sentences, which may be necessary to elicit conflict resolution in language processing.
Attitude and motivation towards L2 learning are not typically included in neurocognitive studies of individual differences in language processing. However, affective variables such as attitude and motivation to learn an L2 have been shown to have a positive correlation with learning in behavioral studies (Crookes & Schmidt, 1991; Gardner, 1980; Masgoret & Gardner, 2003; see Introduction for review). In the only study we are aware of that examines the role of motivation on neurocognitive L2 processing, Tanner et al. (2014) included the factor of motivation to speak like a native speaker (as measured through a language questionnaire, using a likert scale) as an independent variable with L2 neurocognitive processing. Rather than using a rating scale, in this study motivation and attitude was tested through the110-item ATM questionnaire. It appeared that motivation towards L2 learning correlated positively with L2 proficiency, with participants who demonstrated higher motivation towards L2 learning also demonstrating higher L2 proficiency (as measured by offline L2 vocabulary accuracy, score in the TVIP task, self-reported proficiency in the L2, and language history questionnaire responses regarding L2 exposure).

Child Language Processing

Most neurocognitive and psycholinguistic studies on L2 learning have focused on adult learners (for a review, see van Hell & Tokowicz, 2010), who may greatly differ from children in rate of learning and cognitive processes utilized in L2 learning, or even in L1 processing. Likewise, theoretical models on L2 sentence processing (Competition Model and Proficiency-based Model) are based on adult processing. In the present study I measured neurocognitive processing in both the L1 and L2 for children aged 9-11 years
(see Methods), as well as performance in a series of individual difference measures implicated in L2 learning.

**L2 Processing:**

The predictions of the Competition or Proficiency-based Models are based on adult processing, and there are currently no theoretical models that specifically focus on child L2 processing. However, we can assume that, if the Competition Model is based on cross-linguistic influence of an (adult) entrenched L1 on an L2, and children have an L1 that is not as entrenched, cross-linguistic influence would not be as influential as is predicted for adult learners. In the only study of child L2 learners, Brenders et al. (in prep; see Introduction) found beginning L2 learners displayed an AN + attenuated P600 in response to ungrammaticality in a structure similar between Dutch L1-L2 English, and a LAN to ungrammaticality in a structure unique to L2 English. This suggests that cross-linguistic features may play a role in L2 learning in child L2 processing. However, the child L2 learners tested in Brenders et al. (in prep) had exposure to their L2 outside of the classroom. The child L2 learners tested in the present study have little, if any, exposure to L2 Spanish outside the classroom.

So far, I have tested 7 child participants, and data collection will continue in Spring 2016 when a new cohort of L2 learners become available. The currently available data show no significant effects of grammaticality, in the 300-500 or 600-800 ms windows, in the grand mean or individual conditions. However, upon visual inspection, there is a positive going trend to ungrammaticality in midline electrodes. We also see emerging sensitivity to grammaticality in the similar condition at frontal and central midline electrodes, and in the unique condition, at frontal electrodes. These effects were
not statistically significant, and only with additional child participants can we speak to how reliable this trend would be.

**L1 Processing:**

To investigate processing of morphosyntax in the L1 in native English speaking children, we used sentences from three constructions: Tense, Reflexive, and Subject-Verb Agreement (see Methods). Research has shown that children have largely successfully acquired the use of auxiliary in tense agreement in the L1 by the age of 4 (Theakston & Lieven, 2008). Reflexives have been found to be acquired around 3 years of age (see Guasti, 2002, for a review), and subject-verb agreement is successfully used by age 4 (Brown, 1973), though recent studies demonstrate comprehension may not be fully in place until age 6 (Johnson et al., 2005). These features are expected to be well in place by the age of testing in the present study, and as proficient L1 English speakers, we would expect to see P600 effects in morphosyntactic processing of ungrammaticality in the L1.

However, the few ERP studies of child L1 processing indicate that there are differential patterns in child and adult language processing. Almost all studies of child L1 processing that are available focus on the auditory modality, and measure processing of semantic anomalies. In one of the very few neurocognitive studies of syntactic processing in children, Friederici and Hahne (2001) presented children (ages 6-9 yrs old) with phrase structure violations in German, in case-marking. Children were found to present what was called a “P600” in processing of syntactic errors, though the signal was significantly larger in amplitude for children than typically seen in adults, and presented with a latency of 750 ms, extending to 1500 ms post stimulus (in contrast to the common 500-700 ms in adult processing). Atchley et al. (2006) measured ERP responses to violations of subject-
verb agreement in English-speaking 8.5-13 years old children and adults. Both adults and children demonstrated significant P600 effects. As ERP in Atchley et al. (2006) was calculated using grand mean procedures, it may be that the processing of older participants was driving the P600 effect.

In the present study, I found no significant sensitivity to grammaticality, in midline or lateral sites, in both the grand mean (see Figure 27) or each individual condition (see Figures 28A, 28B & 28C). As seen above in adult L1 data, there is significant within group variability in processing patterns. As the current group of child participants was only 7 individuals, we cannot reliably say whether what we observed was a result of true L1 processing patterns for children in this age range, or if it is a result of averaging in a small and very variable dataset. Preliminary RDI analysis indicates that there is definitely variability in neurocognitive processing within this group, with an overall negative RDI in the Tense (4 participants N400-like) and Subject-Verb conditions (3 participants N400-like). RDI in the reflexive condition was positive (3 individuals P600-like).
Chapter 5: CONCLUSION & FUTURE DIRECTIONS

In this study I tested neurocognitive processing in the L1 and L2, in children and adults, and correlated language processing measures with individual difference measures, to examine potential associations. This experiment consisted of several novel elements, which have the potential to inform language processing research. For adults and children, there is a lack of data on how intermediate learners process an L2, and how well models such as the Competition Model or Proficiency-based Model do or do not support processing patterns. We found that the Proficiency-based Model did not account for the findings in either group in the present study. In adult L2 learners, we discovered that findings were not fully explained by the Competition Model either. Through use of new calculations to examine variability in language processing, we were able to demonstrate that findings in our study may be attributed to within group variability, disguised by traditional grand mean analyses. As variability has been shown to be substantial in both L1 and L2 processing, in both adults and children, future ERP studies should be aware of the potential issues of traditional ERP analyses as they currently are done. In order to examine this variation further, we correlated individual difference factors with language processing measures. In adult participants, only motivation was found to correlate with L2 proficiency, indicating the role of motivation may be important in language processing. Through investigating child L1 processing, we added to the very few studies of neurocognitive processing, and through preliminary RDI analyses demonstrated that variability within groups is just as large, if not larger, than that found in adult language processing.
The present study has many avenues for further development. Most language processing studies to date have measured L2 processing in highly proficient adult L2 learners, and have had potential confounds of L2 exposure and AoA among others. More research is needed for variable populations, both in age, and in proficiency. Without information as to how language develops, we cannot have a clear picture of what eventual L2 processing encompasses, and what mechanisms underlie ultimate processing. Obtaining ERP measures of L1 processing in experiments is necessary, as L2 processing may vary as a function of L1. Ultimately, findings from this line of research can have implications for building our knowledge of how language is learned, how it is processed, and how the L1 and L2 interact in the bilingual mind.
REFERENCES


Appendix

Appendix A

Table 15 presents the ANOVAs broken down across the N400 and P600 windows by negative response dominance index for each of the three conditions. Figure 32 presents grand average waveform for negative RDI bias participants for each condition. Table 16 presents the ANOVA analyses broken down across the N400 and P600 windows by positive response dominance index for each of the three conditions. Figure 33 presents grand average waveform for positive RDI bias participants for each condition.

Breakdown of participant RDIs into separate analyses showed the following. Individuals with negative dominance showed wide-spread negativities during processing sentences in the similar and unique conditions in the 300-500 ms window, at midline electrodes, commonly implicated in N400-like processing. To sentences of the dissimilar condition, these participants showed wide-spread negativity in the 300-500 ms window, with an interaction commonly found in the LAN component – with greatest activation at left anterior electrodes. In the 600-800 ms window these negativities continue, though with topographic changes. Individuals who display positive dominance show only early (300-500 ms) widespread positivity for sentences in the unique condition, but show P600-like processing in the 600-800 ms window for all conditions.

The finding that there are individuals in each condition who have negative or positive response dominance causes an issue when considering traditional grand mean analyses of participants, where the ERPs of all participants are averaged, and positivity and negativities are cancelled out, leaving an attenuated effect, or a different effect
altogether. In the grand mean analyses, separated by condition (as shown in Figures 5A, 5B and 5C) we find no significant effects or interactions in the 300-500 ms window for any group, overlooking the N400 in the 300-500 ms window in the negative dominant group for the similar and unique conditions (see Figures 32A and 32C), and the LAN found in processing of dissimilar sentences (see Figure 32B). In the grand mean analysis displayed in Figures 5A, 5B and 5C a P600 is only found for processing in the dissimilar condition, whereas positivity dominant processors show P600s to each structure in the 600-800 ms window. Again, the traditional grand mean analyses seems to be skewing and masking true patterns in L2 processing.

**Negative RDI Bias.**

<table>
<thead>
<tr>
<th></th>
<th>Similar Condition</th>
<th>Unique Condition</th>
<th>Dissimilar Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>300-500 ms</td>
<td>600-800 ms</td>
<td>300-500 ms</td>
</tr>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gram.</td>
<td>26.012*** (1,13)</td>
<td>8.29** (1,13)</td>
<td>14.18*** (1,7)</td>
</tr>
<tr>
<td>Gram. x Elec.</td>
<td>--</td>
<td>4.87** (2,26)</td>
<td>--</td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gram.</td>
<td>19.31*** (1,13)</td>
<td>--</td>
<td>10.78** (1,7)</td>
</tr>
<tr>
<td>Gram. X Ant.</td>
<td>--</td>
<td>4.297* (1,13)</td>
<td>--</td>
</tr>
<tr>
<td>Gram. X Hem.</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Gram. X Ant. X Hem.</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Degrees of freedom are reported in parentheses. Gram. = Grammaticality; Cond. = Condition; Elec. = Electrode; Ant = Anteriority; Hem. = Hemisphere.

*.05<p<.1
** p<.05
*** p<.01

**Table 15:** F-statistics from ANOVA on mean amplitude measures for participants of negative RDI in the three L2 Spanish conditions in the 300-500 ms and 600-800 ms windows.
In the 300-500 ms (N400 window) for each condition for the midline (Fz, Pz, Cz) and lateral (left frontal, right frontal, left posterior, right posterior) regions:

**Similar**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 2.23µV) and lateral electrodes (mean amplitude difference: 1.43µV), with more negative amplitude in processing ungrammatical sentences than grammatical.

**Unique**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 1.92µV) and lateral electrodes (mean amplitude difference: 1.44µV), with more negative amplitude in processing ungrammatical sentences than grammatical.

**Dissimilar**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 1.81µV) and lateral electrodes (mean amplitude difference: 1.89µV), with an interaction of Grammaticality X Anteriority X Hemisphere– most pronounced negative amplitude to ungrammatical sentences in the left frontal region.

In the 600-800 ms (P600 window) for each condition for the midline (Fz, Pz, Cz) and lateral (left frontal, right frontal, left posterior, right posterior) regions:

**Similar**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 1.6µV), and an interaction of Grammaticality X Electrode, with greatest negative amplitude in processing ungrammatical sentences in parietal sites. No significant interactions are found at lateral electrodes.

**Unique**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 2.54µV) and lateral electrodes (mean amplitude difference:
1.6µV), with more negative amplitude in processing ungrammatical sentences than grammatical.

**Dissimilar:** At midline electrodes there is no significant interaction. At lateral electrodes there is a significant interaction of Grammaticality X Anteriority X Hemisphere – most pronounced negative amplitude to ungrammatical sentences in the left frontal region.

32(A)
32 (B)

Dissimilar (Number) Condition, Ungrammatical
Dissimilar (Number) Condition, Grammatical
Figure 32: Grand mean waveform from nine representative electrodes contrasting grammatical (green) and ungrammatical (dashed red) sentences, for those individuals demonstrating negative RDI, across three L2 Spanish conditions: (32A) Tense (Similar), (32B) Determiner Number (Dissimilar), and (32C) Determiner Gender (Unique) conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10µV.
Positive RDI Bias.

<table>
<thead>
<tr>
<th></th>
<th>Similar Condition</th>
<th></th>
<th></th>
<th>Unique Condition</th>
<th></th>
<th></th>
<th>Dissimilar Condition</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>300-500 ms</td>
<td>600-800 ms</td>
<td>300-500 ms</td>
<td>600-800 ms</td>
<td>300-500 ms</td>
<td>600-800 ms</td>
<td>300-500 ms</td>
<td>600-800 ms</td>
</tr>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gram.</td>
<td>4.056* (1,5)</td>
<td>10.85** (1,5)</td>
<td>8.44** (1,11)</td>
<td>8.55** (1,11)</td>
<td>4.437* (1,12)</td>
<td>7.67** (1,12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gram. x Elec.</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gram.</td>
<td>--</td>
<td>16.56*** (1,5)</td>
<td>9.04** (1,11)</td>
<td>6.11** (1,11)</td>
<td>--</td>
<td>10.264*** (1,12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gram. x Ant.</td>
<td>4.13* (1,5)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Gram. x Hem.</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Gram. x Ant. x Hem.</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Degrees of freedom are reported in parentheses. Gram. = Grammaticality; Cond. = Condition; Elec. = Electrode; Ant = Anteriority; Hem. = Hemisphere.

*.05<p<.1
** p<.05
*** p<.01

Table 16: F-statistics from ANOVA on mean amplitude measures for participants of positive RDI in the three Spanish conditions in the 300-500 ms and 600-800 ms windows.

In the 300-500 ms (N400 window) for each condition for the midline (Fz, Pz, Cz) and lateral (left frontal, right frontal, left posterior, right posterior) regions:

**Similar:** There are no significant interactions.

**Unique:** There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 2.64µV) and lateral electrodes (mean amplitude difference: 2.75µV), with more positive amplitude in processing ungrammatical sentences than grammatical.

**Dissimilar:** There is a nearly significant effect of grammaticality in midline electrodes, with more positive amplitude in processing ungrammatical sentences than grammatical (mean amplitude difference: 1.81µV). No significant interactions are found at lateral electrodes.
In the 600-800 ms (P600 window) for each condition for the midline (Fz, Pz, Cz) and lateral (left frontal, right frontal, left posterior, right posterior) regions:

**Similar**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 2.67µV) and lateral electrodes (mean amplitude difference: 6.11µV), with more positive amplitude in processing ungrammatical sentences than grammatical.

**Unique**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 2.26µV) and lateral electrodes (mean amplitude difference: 2.75µV), with more positive amplitude in processing ungrammatical sentences than grammatical.

**Dissimilar**: There is a significant effect of grammaticality in midline electrodes (mean amplitude difference: 3.63µV) and lateral electrodes (mean amplitude difference: 3.09µV), with more positive amplitude in processing ungrammatical sentences than grammatical.
33(A)

Similar (Tense) Condition, Ungrammatical
Similar (Tense) Condition, Grammatical
33(B)
Figure 33: Grand mean waveform from nine representative electrodes contrasting grammatical (green) and ungrammatical (dashed red) sentences, for those individuals demonstrating positive RDI, across three L2 Spanish conditions: (33A) Tense (similar), (33B) Determiner Number (dissimilar), and (33C) Determiner Gender (unique) conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10µV.
Appendix B

Figure 34 presents grand average waveform for negative RDI bias participants for each L1 English condition, for midline electrodes (Fz, Cz & Pz). Figure 35 presents grand average waveform for positive RDI bias participants for each L1 English condition, for midline electrodes (Fz, Cz & Pz).

(34)
Figures 34 & 35: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences, for those individuals demonstrating negative RDI (Figure 34) and positive RDI (Figure 35), across three L1 English conditions: Tense, Reflexive and Subject-Verb conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 10µV.

Negative RDI Bias.

In the 300-500 ms (N400 window), overall there is a significant effect of Grammaticality ($F(1,3)=10.98, p<.05$, mean amplitude difference: 2.23µV), with more negative amplitude in processing ungrammatical than grammatical sentences in the midline region (Fz, Cz, Pz). In the lateral regions (Left Frontal (LF), Right Frontal (RF), Left Posterior (LP), Right Posterior (RP)) there is a significant effect of Grammaticality ($F(1,3)=10.98, p<.05$, mean amplitude difference: 2.22µV), with more negative
amplitude in processing ungrammatical than grammatical sentences. In the lateral region there is a significant interaction of Grammaticality X Condition X Anteriority X Hemisphere \((F(2,6)=11.88, p<.05)\).

In the 300-500 ms (N400 window) for each condition for the midline (Fz, Pz, Cz) and lateral (left frontal, right frontal, left posterior, right posterior) regions:

**Tense:** There is a significant effect of grammaticality in midline electrodes \((F(1,4)=17.69, p<.05, \text{mean amplitude difference: } 2.96\mu V)\), with more negative amplitude in processing ungrammatical sentences than grammatical. There were no significant effects in lateral regions.

**Reflexive:** There is a significant effect of grammaticality in midline electrodes \((F(1,3)=9.78, p<.05, \text{mean amplitude difference: } 1.92\mu V)\), and lateral electrodes \((F(1,3)=11.25, p<.05, \text{mean amplitude difference: } 2.31\mu V)\), with more negative amplitude in processing ungrammatical sentences than grammatical.

**Subject-Verb:** There is a significant interaction of Grammaticality X Electrode \((F(2,8)=5.8, p<.05)\), such that ungrammatical sentences were processed with significantly more negative amplitude than grammatical at the parietal electrode. In lateral sites there is no significant effect of grammaticality on amplitude of processing, though an interaction of Grammaticality X Anteriority X Hemisphere is present \((F(1,4)=15.24, p<.05)\) – most pronounced to ungrammatical sentences in the Left Frontal region.

In the **500-800 ms** (P600 window) there are no significant effects or interactions at midline electrodes. In the lateral regions there is an interaction between Grammaticality X Anteriority X Hemisphere \((F(1,3)=11.32, p<.05)\). Only in the Subject-Verb condition is there a significant effect of Grammaticality found \((F(1,4)=130.77,\)
with most negative amplitude in processing to ungrammatical sentences in the Left Frontal Region.

**Positive RDI Bias.**

In the **300-500 ms** (N400 window), overall there is a significant effect of Grammaticality ($F(1,14)=5.44, p<.05$, mean amplitude difference: .81µV), with more positive amplitude in processing ungrammatical than grammatical sentences in the midline region (Fz, Cz, Pz). Additionally there is an interaction between Grammaticality X Electrode ($F(2,28)= 8.02, p<.01$), with most positive amplitude in ungrammatical sentence processing at Pz. In the lateral regions (Left Frontal (LF), Right Frontal (RF), Left Posterior (LP), Right Posterior (RP)) there are no significant effects or interactions.

In the 300-500 ms (N400 window) for each condition for the midline (Fz, Pz, Cz) region:

**Tense:** There is a significant interaction of Grammaticality X Electrode ($F(2,28)=3.89, p=.056$), with most positive amplitude in processing of ungrammatical sentences in the parietal region.

**Reflexive:** There are no significant effects or interactions found.

**Subject-Verb:** There is a significant effect of grammaticality, with more positive amplitude in processing ungrammatical sentences than grammatical ($F(1,14)=7.39, p<.05$, mean amplitude difference: 1.36µV). There is an interaction of Grammaticality X Electrode ($F(2,28)=6.58, p<.05$), with most positive amplitude in processing of ungrammatical sentences in the parietal region.

In the **500-800 ms** (P600 window), overall there is a significant effect of Grammaticality ($F(1,14)=219.38, p<.01$, mean amplitude difference: 3.97µV), with more
positive amplitude in processing ungrammatical than grammatical sentences in the midline region (Fz, Cz, Pz). Additionally there is an interaction between Grammaticality X Electrode ($F(2,28)=15.14, p<.05$), with most positive amplitude in ungrammatical sentence processing at Pz. In the lateral regions (Left Frontal (LF), Right Frontal (RF), Left Posterior (LP), Right Posterior (RP)) there are no significant effects or interactions.

In the 300-500 ms (N400 window) for each condition for the midline (Fz, Pz, Cz) region:

**Tense:** There is a significant effect of grammaticality, with more positive amplitude in processing ungrammatical sentences than grammatical ($F(1,14)=59.29$, $p<.001$, mean amplitude difference: $4.79\mu$V). In a significant interaction of Grammaticality X Electrode ($F(2,28)=11.90$, $p=.001$), greatest positive amplitude in processing ungrammatical sentences was at central sites.

**Reflexive:** There is a significant effect of grammaticality, with more positive amplitude in processing ungrammatical sentences than grammatical ($F(1,15)=23.18$, $p<.001$, mean amplitude difference: $3.29\mu$V). In a significant interaction of Grammaticality X Electrode ($F(2,30)=4.61$, $p<.05$), greatest positive amplitude in processing ungrammatical sentences was at central sites.

**Subject-Verb:** There is a significant effect of grammaticality, with more positive amplitude in processing ungrammatical sentences than grammatical ($F(1,14)=20.92$, $p<.001$, mean amplitude difference: $3.77\mu$V). In a significant interaction of Grammaticality X Electrode ($F(2,28)=5.89$, $p<.05$), greatest positive amplitude in processing ungrammatical sentences was at central sites.
Appendix C

Table 17 presents the ANOVAs broken down across the N400 and P600 windows by negative response dominance index for child learners in each of the three L2 Spanish conditions. Figure 36 presents grand average waveform for negative RDI bias participants for each condition. Table 18 presents the ANOVA analyses broken down across the N400 and P600 windows by positive response dominance index for each of the three conditions. Figure 37 presents grand average waveform for positive RDI bias participants for each condition.

Negative RDI Bias.

<table>
<thead>
<tr>
<th></th>
<th>Similar Condition</th>
<th>Unique Condition</th>
<th>Dissimilar Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>300-500 ms</td>
<td>600-800 ms</td>
<td>300-500 ms</td>
</tr>
<tr>
<td>Midline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Electrode</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority</td>
<td>--</td>
<td>32.12**(1,2)</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Hemisphere</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Anteriority X Hemisphere</td>
<td>--</td>
<td>--</td>
<td>6.79*(1,3)</td>
</tr>
</tbody>
</table>

Degrees of freedom are reported in parentheses.

* .05 < p < .1
** p < .05
*** p < .01

Table 17: F-statistics from ANOVA on mean amplitude measures for participants of negative RDI in the three Spanish conditions in the 300-500 ms and 600-800 ms windows.

In the **300-500 ms** (N400 window) for each condition for the midline (Fz, Pz, Cz) and lateral (Left Frontal, Right Frontal, Left Posterior, Right Posterior) regions:

**Similar**: There are no significant effects or interactions found.

**Unique**: There are no significant effects or interactions found.

**Dissimilar**: There are no significant effects or interactions found.
In the **600-800 ms** (P600 window) for each condition for the midline (Fz, Pz, Cz) and lateral (Left Frontal, Right Frontal, Left Posterior, Right Posterior) regions:

**Similar**: There are no significant effects or interactions found at midline electrodes. At lateral sites there is a significant interaction of Grammaticality X Anteriority \((F(1,2)=32.12, p<.05)\) – most pronounced to ungrammatical sentences in the posterior region.

**Unique**: There are no significant effects or interactions found.

**Dissimilar**: There are no significant effects or interactions found.
Figure 36: Grand mean waveform from nine representative electrodes contrasting grammatical (green) and ungrammatical (dashed red) sentences, for those individuals demonstrating negative RDI, across three L2 Spanish conditions: Tense (similar), Determiner Number (dissimilar), and Determiner Gender (unique) conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16µV.
Positive RDI Bias.

<table>
<thead>
<tr>
<th></th>
<th>Similar Condition</th>
<th>Unique Condition</th>
<th>Dissimilar Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>300-500 ms 600-800 ms</td>
<td>300-500 ms 600-800 ms</td>
<td>300-500 ms 600-800 ms</td>
</tr>
<tr>
<td><strong>Midline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammaticality</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Grammaticality X Electrode</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

| **Lateral**         |                   |                  |                      |
| Grammaticality      | 16.75**(1,3)      | 21.46**(1,3)     | 9.12*(1,3)           |
| Grammaticality X Anteriority | --                | --               | --                   |
| Grammaticality X Hemisphere | --                | --               | --                   |
| Grammaticality X Anteriority X Hemisphere | --                | --               | --                   |

Degrees of freedom are reported in parentheses.
* .05 < p < .1
** p < .05
*** p < .01

**Table 18:** F-statistics from ANOVA on mean amplitude measures for participants of positive RDI in the three Spanish conditions in the 300-500 ms and 600-800 ms windows.

In the **300-500 ms** (N400 window) for each condition for the midline (Fz, Pz, Cz) and lateral (Left Frontal, Right Frontal, Left Posterior, Right Posterior) regions:

**Similar:** There is no significant effect or interaction at midline electrodes. At lateral sites there is a significant effect of grammaticality, with more positive amplitude in processing ungrammatical sentences than grammatical (mean amplitude difference: 2.28 µV).

**Unique:** There are no significant effects or interactions found.

**Dissimilar:** There are no significant effects or interactions found.

In the **600-800 ms** (P600 window) for each condition for the midline (Fz, Pz, Cz) and lateral (Left Frontal, Right Frontal, Left Posterior, Right Posterior) regions:

**Similar:** There is no significant effect or interaction at midline electrodes. At lateral sites there is a significant effect of grammaticality, with more positive amplitude in processing ungrammatical sentences than grammatical (mean amplitude difference: 4.66 µV).
**Unique:** There are no significant effects or interactions found.

**Dissimilar:** There are no significant effects or interactions found.

(37)

**Figure 37:** Grand mean waveform from nine representative electrodes contrasting grammatical and ungrammatical sentences, for those individuals demonstrating positive RDI, across three Spanish conditions: Tense (Similar), Determiner Number (Dissimilar), and Determiner Gender (Unique) conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16µV.
Appendix D

Figure 38 presents grand average waveform for child negative RDI bias participants for each L1 English condition, for midline electrodes (Fz, Cz & Pz). Figure 39 presents grand average waveform for postitive RDI bias participants for each L1 English condition, for midline electrodes (Fz, Cz & Pz).

**Negative RDI Bias.**

In the 300-500 ms (N400 window), overall there is a significant effect of Grammaticality at midline electrodes ($F(1,2)=75.48$, $p<.05$, mean amplitude difference: 3.20µV), and lateral electrodes ($F(1,2)=51.39$, $p<.05$, mean amplitude difference: 3.39µV), with more negative amplitude in processing ungrammatical than grammatical sentences. In the lateral region there are no significant effects or interactions.

In the 500-800 ms (P600 window) there is no significant effect of Grammaticality or interaction of Grammaticality, Condition or Electrode in midline electrodes. In the lateral regions there is a significant effect of grammaticality ($F(1,2)=23.56$, $p<.05$), with more negative amplitude in processing ungrammatical than grammatical sentences (mean amplitude difference: 3.75µV).
(38)
Figures 38 & 39: Grand mean waveforms from nine representative electrodes contrasting amplitude of processing of grammatical (green) and ungrammatical (dashed red) sentences, for those individuals demonstrating negative RDI (Figure 38) and positive RDI (Figure 39), across three L1 English conditions: Tense, Reflexive and Subject-Verb conditions. Presentation of the critical item is represented by the vertical bar (0 ms), and the waveform is shown from -200 ms pre critical stimulus to 1000 ms post stimulus presentation. Negative voltage is plotted up, and voltage is a range of 16µV.

**Positive RDI Bias.**

In the **300-500 ms** there are no significant effects or interactions found.

In the **500-800 ms** there are no significant effects or interactions found.