A NEUROCOGNITIVE INVESTIGATION OF BILINGUAL ADVANTAGES AT ADDITIONAL LANGUAGE LEARNING

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

By

Sarah Elizabeth Grey, M.S.

Washington, DC
June 14, 2013
A NEUROCOGNITIVE INVESTIGATION OF BILINGUAL ADVANTAGES AT ADDITIONAL LANGUAGE LEARNING

Sarah Elizabeth Grey, M.S.

Thesis Advisors: Cristina Sanz, Ph.D., Michael T. Ullman, Ph.D.

ABSTRACT

This study investigated bilingual advantages at adult additional language learning. It aimed to address some of the limitations in previous research and provide a multi-dimensional perspective on potential advantages by measuring both behavioral and neural outcomes (using event-related potentials, or ERPs), at different points in the learning trajectory, under different instructional contexts, and for different linguistic targets.

Early, highly proficient Mandarin-English bilinguals were trained and tested on their learning of a Romance language-like artificial language called Brocanto2 (e.g., Morgan-Short, Sanz, Steinhauer, & Ullman, 2010). Bilinguals were trained either under an Instructed condition, which contained metalinguistic information about Brocanto2 together with meaningful examples; or under an Uninstructed condition: meaningful examples and no metalinguistic information. Both groups of bilinguals engaged in comprehension and production practice and at two points, low and high experience, were given grammaticality judgment tasks during which ERP data was simultaneously gathered. The data from these bilinguals was compared to a subset of previously-tested English monolinguals who were trained and tested under the same conditions (e.g., Morgan-Short, Steinhauer, Sanz, & Ullman, 2012).

The behavioral results showed that under an Instructed condition bilinguals and monolinguals did not perform differently, either on practice or grammaticality judgment. This
may be at least partly due to a leveling effect of the Instructed context. In the Uninstructed context, monolinguals outperformed bilinguals. This may have been related to differences in the way the two groups capitalized on practice, and the function of such practice in promoting language development.

The ERP data indicated that at low experience bilinguals and monolinguals recruited distinct neural mechanisms as indexed by P600s and late anterior negativities, but that at high experience differences in processing were less apparent. However, some suggestive evidence showed that at high experience bilinguals may have been recruiting the mechanisms underlying P600s and late anterior negativities more strongly than the monolinguals. Only the Uninstructed monolinguals appeared to have undergone a neurocognitive shift characterized by early, automatic signatures of processing, whereas no other group evidenced such a shift and the Uninstructed bilinguals may have still been relying on lexical/semantic mechanisms at the same level of experience. The results of this study have implications for research on bilingualism, neurocognition of late-learned language, and language pedagogy.
Dedication

To my family- especially David Grey, Mariana Grey, Crystal Grey-Hewett, and Amanda Walker- who have been beacons of support all these many years of my graduate school career. Whatever I achieve is owed to you and your belief in me. Thank you for understanding that I needed to leave and pursue this. I love you.
Acknowledgments

I find it difficult to justly measure the length and depth of my gratitude to the many important people who have contributed to this body of work and everything that led up to it. First and absolute foremost, I would like to express my most emphatic and heartfelt thanks to Cristina Sanz and Michael Ullman, the best support and guidance team that I could have imagined. Your drive, knowledge, and passionate dedication as mentors have inspired, challenged, and motivated me in indescribable ways and left deep, lasting impressions. I am endlessly honored to know and work with you both. In a similar vein, I would like to thank Kara Morgan-Short, who inherited me as I followed in her multidisciplinary footsteps, which has made all the difference in my personal and professional development. Kara, your research energy, focus, balance, and unfailing responsiveness during this project have been invaluable and you are a true role-model.

To the professors that have spread their knowledge and love of linguistics and language research over the years - Héctor Campos, Rusan Chen, Rod Ellis, Rhonda Friedman, Elena Herburger, Darlene Howard, Ron Leow, Allison Mackey, Alfonso Morales-Front, Patrick Rebuschat, Chandan Vaidya, and Elizabeth Zsiga - I thank all of you for your generosity in sharing your expertise and for providing me with a dynamic and robust skill set. I would also like to thank Michael Leeser and Gretchen Sunderland at The Florida State University, who were the first to encourage me to pursue a doctoral degree, and who planted the seeds of confidence in my ability to achieve such a thing.

To those with whom I have worked in the Brain and Language Lab - Laura Babcock, Ingrid Finger, Kaitlyn Litcofsky, Jarrett Lovelett, and Mariel Pullman - thank you for adding light to each and every in-lab day with your humor, camaraderie, and inspiring commitment to good science.
To my fellow graduate students and the dear friends I have made over the years - with whom countless laughs, grumbles, encouraging words, coffees, and cocktails have been shared - I could not have done this without you. Julio Torres, from the very first day of this unique journey you have been the most grounding and therapeutic friend and colleague. Thank you.

Rachael Allbritten, Jessie Cox, Elizabeth Kissling, Aubrey Logan-Terry, Ellen Johnson Serafini, and Kaitlyn Tagarelli, I have no words for how wonderful it has been to be surrounded by and have the support of women with the strength, kindness, and intelligence you all possess. You are each amazing and I am ever thankful to have you ladies in my corner. To Sergio Adrada Rafael, for being by my side for the majority of this academic pursuit, thank you. Also, to the dear women who were with me before, and have been with me throughout, this experience – My-Van Le-Talpasz, Jennie Bellone, and Sterling McMahan- I thank you endlessly for tolerating, supporting, and encouraging me in unbelievable ways and for always being there.

I also thank María-Elvira Daza, Michael Ferreira, Kristen Hall, Gwen Kirkpatrick, and Elena Herburger and Alfonso Morales-Front again, for their organizational support during my graduate career at Georgetown University. Having had such excited, unmitigated, and confident departmental support during my time here has been a valued gift and has not only immensely enriched my experience but has also instilled in me important and lasting professional values. Thank you.

I would also like to briefly acknowledge two other important sources of support. Studying, writing, and to a certain extent research, are often carried out in solitude and, at least in my experience, there are supporting mechanisms that make the solitude seem less solitary. During my journey through a bachelor’s degree, two master’s degrees, and a doctoral degree, both of the acknowledged entities have been indescribably helpful at various and continuous
points in my pursuit of higher education. Both served calming and energizing purposes, depending on my mood and productive status. Both helped me to separate, focus, and ground myself, and both encouraged me to gather my energetic, optimistic, and confident spirits and move forward. So, to Jim Adkins, Tom Linton, Rick Burch, and Zach Lind of Jimmy Eat World, thank you. And to Elvis Presley, who has been my most consistent musical therapy since I was a young girl, thank you.

Finally, I would like to acknowledge the Broconto2 bilinguals who took part in this study; your patience, interest, and engagement during lengthy testing sessions was very much appreciated and none of this would have been possible without you. This research was supported by a National Science Foundation Doctoral Dissertation Improvement Grant (SF 1124144) and a Language Learning journal Dissertation Grant.
# Table of Contents

## CHAPTER 1 - INTRODUCTION

- **Bilingual third language acquisition in adulthood**  
  1
- **Statement of the Problem**  
  2
- **Definition of Key Terms**  
  4
- **Bilingual**  
  4
- **L3 Acquisition and Learning**  
  5
- **Instructional Condition**  
  6
- **Explicit: Learning, Exposure, and Knowledge**  
  6
- **Implicit: Learning, Exposure, and Knowledge**  
  7
- **Processing**  
  8
- **Proficiency**  
  9
- **Declarative Memory**  
  12
- **Procedural Memory**  
  13

## CHAPTER 2 - REVIEW OF THE LITERATURE

- **Introduction**  
  17
- **Bilingual L3A**  
  17
- **Dual-language environments**  
  17
- **Positive Effects**  
  17
- **No or Negative Effects**  
  22
- **Single-language environments**  
  23
- **Positive Effects**  
  23
- **No or Negative Effects**  
  24
- **Summary: Dual-language and single-language environments**  
  25
- **Laboratory Studies**  
  26
- **A note on the effectiveness of explicit and implicit training in monolingual L2A**  
  27
- **Positive Effects**  
  28
- **No or Negative Effects**  
  39
- **Summary: Laboratory-based research**  
  42
- **Conclusions: Bilingual L3A compared to monolingual L2A**  
  43
- **Neurocognitive theories of late-learned language**  
  44
- **The declarative procedural model of language**  
  45
- **Declarative and procedural determinants of second language**  
  47
- **The shallow structures hypothesis**  
  50
- **The convergence hypothesis**  
  51
- **The competition model**  
  54
- **Associative learning in L2**  
  56
- **Summary: Neurocognitive theories of late-learned language**  
  58
- **Event-related potentials and language**  
  59
- **Introduction**  
  59
- **Event-related potentials**  
  60
- **N400s**  
  62
- **N400 summary**  
  68
- **(left) anterior negativities**  
  68
- **LAN summary**  
  72
- **P600s**  
  73

ix
P600 SUMMARY 78
CONCLUSIONS: ERPs in late-learned language processing 79
RATIONALE FOR THE CURRENT STUDY 79
RESEARCH QUESTIONS 81
PREDICTIONS 82
BEHAVIORAL 82
ELECTROPHYSIOLOGICAL 82

CHAPTER 3 – RESEARCH METHODS AND DESIGN 84
INTRODUCTION 84
PARTICIPANTS 84
MATERIALS 88
THE ARTIFICIAL LANGUAGE: BROCANTO2 88
DESIGN 91
INSTRUCTIONAL CONDITIONS 91
COMPREHENSION AND PRODUCTION PRACTICE IN BROCANTO2 92
BEHAVIORAL/ERP ASSESSMENT 93
PROCEDURE 95
DATA ANALYSIS 96

CHAPTER 4 – RESULTS 98
INTRODUCTION 98
PRACTICE PERFORMANCE 98
PRACTICE PERFORMANCE SUMMARY 108
INSTRUCTED LEARNERS: GRAMMATICALITY JUDGMENT PERFORMANCE 109
INSTRUCTED LEARNERS: ELECTROPHYSIOLOGICAL PATTERNS 110
WORD ORDER 110
GENDER AGREEMENT 118
INSTRUCTED LEARNERS: SUMMARY OF ELECTROPHYSIOLOGICAL RESULTS 124
UNINSTRUCTED LEARNERS: GRAMMATICALITY JUDGMENT PERFORMANCE 125
UNINSTRUCTED LEARNERS: ELECTROPHYSIOLOGICAL PATTERNS 127
WORD ORDER 126
GENDER AGREEMENT 134
UNINSTRUCTED LEARNERS: SUMMARY OF ELECTROPHYSIOLOGICAL RESULTS 140

CHAPTER 5 – DISCUSSION AND CONCLUSION 142
DISCUSSION 142
BEHAVIORAL PERFORMANCE 142
ELECTROPHYSIOLOGICAL OUTCOMES 151
LIMITATIONS AND FUTURE RESEARCH 167
CONCLUSION 170

APPENDIX A 174
APPENDIX B 177
APPENDIX C 183
REFERENCES 189
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>DESCRIPTIVE INFORMATION OF LANGUAGE STATUS IN BILINGUALS</td>
<td>87</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>DESCRIPTIVE CHARACTERISTICS OF INSTRUCTED AND UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>87</td>
</tr>
<tr>
<td>TABLE 3</td>
<td>INFORMATION ON ADDITIONAL LANGUAGE EXPERIENCE IN INSTRUCTED AND UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>88</td>
</tr>
<tr>
<td>TABLE 4</td>
<td>DESCRIPTIVE CHARACTERISTICS OF BROCANTO2</td>
<td>90</td>
</tr>
<tr>
<td>TABLE 5</td>
<td>PRACTICE PERFORMANCE IN INSTRUCTED AND UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>100</td>
</tr>
<tr>
<td>TABLE 6</td>
<td>PRACTICE ACCURACY OVER THE 44 BLOCKS OF PRACTICE (TIME SERIES) IN INSTRUCTED AND UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>102</td>
</tr>
<tr>
<td>TABLE 7</td>
<td>COMPREHENSION ACCURACY OVER THE 22 BLOCKS OF PRACTICE (TIME SERIES) IN INSTRUCTED AND UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>104</td>
</tr>
<tr>
<td>TABLE 8</td>
<td>PRODUCTION ACCURACY OVER THE 22 BLOCKS OF PRACTICE (TIME SERIES) IN INSTRUCTED AND UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>105</td>
</tr>
<tr>
<td>TABLE 9</td>
<td>REACTION TIME OVER THE 22 BLOCKS OF COMPREHENSION PRACTICE (TIME SERIES) IN INSTRUCTED AND UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>107</td>
</tr>
<tr>
<td>TABLE 10</td>
<td>PERFORMANCE ON WORD ORDER JUDGMENT IN INSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>110</td>
</tr>
<tr>
<td>TABLE 11</td>
<td>PERFORMANCE ON GENDER AGREEMENT JUDGMENT IN INSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>110</td>
</tr>
<tr>
<td>TABLE 12</td>
<td>SUMMARY OF ANOVA F-VALUES FOR THE COMPARISON BETWEEN INSTRUCTED BILINGUALS AND MONOLINGUALS’ PROCESSING OF WORD ORDER</td>
<td>114</td>
</tr>
<tr>
<td>TABLE 13</td>
<td>SUMMARY OF ANOVA F-VALUES FOR THE COMPARISON BETWEEN INSTRUCTED BILINGUALS AND MONOLINGUALS’ PROCESSING OF GENDER AGREEMENT</td>
<td>121</td>
</tr>
<tr>
<td>TABLE 14</td>
<td>PERFORMANCE ON WORD ORDER JUDGMENT IN UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>126</td>
</tr>
<tr>
<td>TABLE 15</td>
<td>PERFORMANCE ON GENDER AGREEMENT JUDGMENT IN UNINSTRUCTED BILINGUALS AND MONOLINGUALS</td>
<td>127</td>
</tr>
<tr>
<td>TABLE 16</td>
<td>SUMMARY OF ANOVA F-VALUES FOR THE COMPARISON BETWEEN UNINSTRUCTED BILINGUALS AND MONOLINGUALS’ PROCESSING OF WORD ORDER</td>
<td>130</td>
</tr>
<tr>
<td>TABLE 17</td>
<td>SUMMARY OF ANOVA F-VALUES FOR THE COMPARISON BETWEEN UNINSTRUCTED BILINGUALS AND MONOLINGUALS’ PROCESSING OF GENDER AGREEMENT</td>
<td>137</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. Accuracy across all 44 practice blocks 102

Figure 2. Accuracy for comprehension performance 104

Figure 3. Accuracy for production performance 106

Figure 4. Reaction time performance on comprehension blocks 108

Figure 5. ERP data. Processing of word order targets in instructed bilinguals and monolinguals 112

Figure 6. Topographical maps 113

Figure 7. ERP data. Processing of gender agreement targets in instructed bilinguals and monolinguals 119

Figure 8. Topographical maps 120

Figure 9. ERP data. Processing of word order targets in uninstructed bilinguals and monolinguals 128

Figure 10. Topographical maps 129

Figure 11. ERP data. Processing of gender agreement targets in uninstructed bilinguals and monolinguals 135

Figure 12. Topographical maps 136
Chapter 1 - Introduction

Bilingual third language acquisition in adulthood

Learning an additional language as an adult is a undeniably challenging task. All learners bring with them diverse sets of skills to the acquisition table and while the learning challenge is unavoidable, some learners seem better equipped for the task. Bilinguals in particular appear to have a comparative advantage over monolinguals on a range of cognitive functions such as executive control, conflict resolution, a protective cognitive reserve in the face of aging (Bialystok, 2009), and also for the learning of an additional language as an adult (Bialystok, 2001; Cenoz, 2003, 2011).

The cognitive benefits of bilingualism are most notoriously associated with work done by Ellen Bialystok and colleagues. This body of research suggests that the consistent mental balancing of two linguistic systems is related to improved cognitive function, specifically on executive control and attention-related tasks. Bilinguals have been found to outperform monolinguals on these non-linguistic cognitive tasks in adulthood (Bialystok, Craik, Klein, & Viswanathan, 2004; Costa, Hernandez, & Sebastian-Galles, 2008; Feng, Bialystok, & Diamond, 2009), in childhood (Bialystok & Martin, 2004; Martin-Rhee & Bialystok, 2008), and in advanced aging (Bialystok, Craik, & Luk, 2008; Ellen Bialystok, Martin, & Viswanathan, 2005). However, other researchers have not found a comparative advantage for bilinguals over monolinguals (e.g., Hilchey & Klein, 2011; Paap & Greenberg, 2013), suggesting that there are important complexities associated with the bilingual advantage that must continue to be elucidated in order to better profile the positive effects of bilingualism.

One particular benefit that has been associated with bilingualism, however, is its facilitative effects on third, or additional, language acquisition (L3A). Such positive effects have
been documented in both classroom and laboratory settings (e.g. Cenoz & Valencia, 1994; Lado, 2008) and for both early and late bilinguals (Stafford, Sanz, & Bowden, 2010). Potential benefits of bilingualism on additional language learning have been associated with greater metalinguistic awareness (Dillon, 2009; Jessner, 2008), better use of learning strategies (Bowden, Sanz, & Stafford, 2005; Kemp, 2007; Lin, 2009), and the availability of a broader linguistic repertoire (Cenoz, 2011; De Angelis, 2007), though no research has addressed whether there may be neurocognitive correlates of this advantage. All of the research on bilingual L3A compared to monolingual acquisition of an additional language (L2A) has investigated potential benefits of bilingualism using only behavioral measures, the majority of which targeted broad language performance. Nonetheless, some studies have investigated specific sub-domains of language, and this research suggests that bilinguals may outperform monolinguals for vocabulary (Keshavarz & Astaneh, 2004) and syntax (Klein, 1995; Zobl, 1993) as well as pragmatics (Safont-Jorda, 2003).

**Statement of the Problem**

Though there exist quite a few studies which have investigated the effects of bilingualism on learning an additional language, there are important limitations to this research. First, results have been reported only in terms of broad language proficiency measures or exclusively behavioral tests. Thus, we know little of the linguistic extent of potential bilingual advantages and nothing of the mechanisms underlying L3 performance. Second, control over amount and onset of additional language exposure as well as language transfer effects has been largely absent (but see Lado, 2008; Lin, 2009). Therefore, the influence of these factors cannot be teased apart from bilingualism and its effect on the outcomes. Third, the existing work on bilingual L3A has examined learners only at one point in their L3 learning stage or proficiency which does not inform the field about the trajectory of bilingual L3A compared to monolingual L2A. Fourth,
bilingual L3A studies have generally not considered the context - such as traditional grammar-based classrooms versus more communicative classrooms - in which the L3 is learned as a potential factor in the emergence of facilitative effects. Moreover, the few studies which have investigated the effects of instructional context have found conflicting patterns of results (Lado, 2008; Nation & McLaughlin, 1986; Nayak, Hansen, Krueger, McLaughlin, 1990).

The current research on bilingual advantages in adult additional language acquisition is underspecified and important, as yet unresolved, issues must be addressed. First, research must investigate bilingual L3A within a process-, not just product-, oriented level of analysis. This can be achieved by considering not only the behavioral measures of learning, but also using online measures of neurocognitive processing as complementary and potentially more powerful tools in assessing bilingual L3A. This is especially relevant for any potential bilingual advantages that may not be captured using broad measures of performance or exclusively behavioral outcomes. Second, different linguistic domains need to be empirically examined in order to provide a more detailed perspective on the scope of bilingual L3A. These linguistic domains need to be examined in an experimental context which allows for control over amount, type, and timing of language input as well as language transfer in order to reliably assess the impact of bilingualism on additional language acquisition. Third, the effectiveness of different types of L3 instructional contexts on not only outcomes, but also L3 processing must be considered. Research on adult monolingual L2A has shown that the type of instructional context (more explicit or more implicit) to which learners are exposed affects both the outcomes (Norris & Ortega, 2000; Spada & Tomita, 2010) and also the processing of L2 (Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012). How such conditions affect bilingual L3A would provide valuable information not only on L3A in isolation, but would, crucially, elucidate potential differences in
the effects of instruction on monolingual L2A and bilingual L3A. Finally, all of these factors must be considered along the trajectory of bilingual L3A, from low to high L3 experience, in order to provide a more refined understanding of how learning actually proceeds, and what differences there may be between bilingual L3A and monolingual L2A during the course of additional language learning in adulthood.

We currently have a reliable body of evidence supporting measurable cognitive benefits of bilingualism and burgeoning evidence that one dimension of these benefits may concern the learning of additional languages in adulthood. However, limitations in previous research and gaps in our knowledge preclude making strong conclusions about how robust this advantage is. More research is needed to determine (a) how extensive the potential advantage may be (i.e., behavioral and neural components of the advantage), (b) what learning contexts might constrain or, conversely, enhance this advantage, and (c) what characterizes the trajectory of learning in terms of both performance-based outcomes and neural indices of online language processing. Before detailing the theoretical and methodological components needed to address the above issues, definitions and operationalizations of key terms related to the study of the neurocognition of bilingual L3A are provided.

**Definition of Key Terms**

**Bilingual**

Within the Second Language Acquisition (SLA) and bilingualism literature, the term 'bilingual' has been privileged to a myriad of definitions (see, for example, Grosjean, 2001) and it is by no means monolithic. The cognitive benefits that have been found for bilingualism are proposed to stem from the unique situation that fluent bilinguals are in when engaging in use of
both languages regularly (Bialystok, 2009). Additionally, the facilitative effects of bilingualism on L3A have been found for bilingual populations that were either in dual-language environments or at particularly high levels of proficiency (Sanz, 2000; Lado, 2008). For purposes of this study on bilingual L3A, the term bilingual will refer to an individual that acquired both of their languages before the age of 6 years old (Fabbro, 1999; Fabbro, 2001; Paradis, 2004, 2009) and reports comparably high levels of proficiency in both.

**L3 Acquisition and learning**

The study of third language acquisition in bilinguals is a relatively new area of research and as such suffers from terminological confusion among researchers (for a review see Hammarberg, 2010). Third language acquisition can refer both to the process of learning a third language and also the study of such learning (Cenoz, 2011). Some researchers define the third language sequentially, or in other words "a language acquired chronologically after the first and second or after the first two in the case of early bilinguals" (Cenoz, 2011, p. 1). However, other researchers have considered the third language any language acquired after the first two, from third onwards (De Angelis, 2007; Hammarberg, 2010). In the present study, third language will refer to an artificial language, Brocanto2, that was designed to be Romance-like (Morgan-Short, 2007). Third language acquisition and learning will refer to the learning of Brocanto2 in an experimental setting; such learning does not require that Brocanto2 be the sequential third language of the bilinguals but that it be their only Romance (like) additional language. Additionally, within the confines of the current study, 'learning' or 'acquisition' will be assumed to have taken place when the bilinguals reach predetermined thresholds of performance or experience on comprehension and production in the L3.
Instructional condition

In SLA there has been a general tendency to associate the nature of a given instructional condition with the type of knowledge formed following exposure. However, type of learning, type of exposure, and the resultant type (or types) of knowledge are not isomorphic. Williams (2009) states that "It is useful to distinguish the nature of the learning process from the status of the resulting knowledge as assessed by a test of learning...the issue of the existence of implicit or explicit knowledge in the mind of the learner is distinct from how it got there" (p. 320). Thus, in order to set a clear path for the remainder of this paper, explicit and implicit learning, exposure, and knowledge, as well as the relevant terminology for the present study, are defined below.

Explicit: Learning, Exposure, and Knowledge

*Explicit learning* is a learner-internal process, one that involves a conscious intention to learn the material at hand, the use of conscious knowledge at the point of learning, and also awareness that there will be an assessment of learning (e.g., Williams, 2009). Explicit learning can take place independent of the input condition to which a learner is exposed, be it implicit or explicit in nature. Input conditions are, necessarily, learner-external factors in the acquisition landscape and Spada & Tomita (2010) followed Norris & Ortega (2000) in defining input conditions that provide rule presentation or direction to attend to forms as explicit conditions. Such conditions can also involve instruction to search for a rule (Rebuschat & Williams, 2006, 2012; Robinson, 1995). These will be the defining characteristics of the term *explicit exposure* used in the present study. Much in the same way that explicit learning can be independent of the exposure condition, the type of exposure condition with which a learner engages can be independent of the resultant knowledge that the learner develops. *Explicit knowledge*, following
Williams (2009), refers to knowledge that the learner is conscious of and aware of using, and is often verbalizable (see also (R. Ellis, 2008; Hulstijn, 2005; Paradis, 2004, 2009, among others).

**Implicit: Learning, Exposure, and Knowledge**

Compared to the term explicit, the use of the term implicit is less clear-cut in SLA research. Nonetheless, operational and converging definitions are emerging in the field. With respect to the term implicit learning, the following definition will be used: *Implicit learning* refers to process of learning something without intending to, being unaware that an assessment of learning will take place, and also having no awareness of the regularity to be learned at the point of learning (Williams, 2009; see also R. Ellis, 2008). Note again that the nature of learning process may be distinct from that of the input, or exposure, condition. An *implicit exposure* condition involves meaning-oriented exposure to the language with no instruction on the language rules, no instruction to search for rules, no direction to attend to target forms (Norris & Ortega, 2000; Spada & Tomita, 2010) and also no information that there will be an assessment of learning (Williams, 2009). Finally, the term *implicit knowledge* is considered to be knowledge that is not accessible to conscious awareness (or not verbalizable) and is deployed automatically (Paradis, 2004, 2009; Williams, 2009). Such knowledge has proven difficult to measure in both SLA and psychology (Rebuschat, 2012), though recent developments in cognitive neuroscience may be useful in reducing this difficulty (Williams, 2009).

From a behavioral perspective, researchers have attempted to identify implicit knowledge by using offline measures of verbalizable knowledge (Rebuschat, 2012; Rebuschat & Williams, 2006; Williams, 2005), speeded performance measures such as reaction time (Leung & Williams, 2011) and timed grammaticality judgments, or spontaneous production measures (R. Ellis, et al., 2009; Serafini, 2013; Stafford, 2011). Recently, subjective measures of awareness,
such as confidence ratings and source attributions on acceptability judgment tasks, have also been used as indices of implicit knowledge (Dienes & Scott, 2005; Rebuschat, 2012).

Having defined these terms as they pertain to the study of adult language acquisition, the terms relevant for the current study will now be clarified. Previous research done with Brocanto2 (the artificial language subjects learn in the present study) used the terms 'implicit condition' and 'explicit condition' to refer to the conditions under which the language input was provided to the learners (Morgan-Short, 2007; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012). The explicit condition used by Morgan-Short and colleagues consisted of the provision of metalinguistic information regarding the rules of Brocanto2 in conjunction with meaningful examples. The implicit condition provided no metalinguistic information to learners. This condition was characterized by exposure to meaningful examples with no direction to attend to forms or search for rules (for more information on Brocanto2 and these exposure conditions see Chapter 3). However, considering the definition of implicit exposure provided above, it is not clear that Morgan-Short's implicit condition wholly fits the description, specifically with respect to knowing there will be a test of learning. In an effort to avoid possible conceptual confounds with the terms explicit and implicit as they are understood in psychology and especially SLA, these terms will not be used to refer to the exposure conditions provided to the bilinguals in the present study. Instead, and following the terminology used in Williams (2009), Instructed and Uninstructed will be used to refer to what Morgan-Short called explicit and implicit, respectively.

Processing

Processing refers to the use of language, either for input or output situations, and in this study L3 processing is operationalized as the evidence gathered from behavioral and
electrophysiological data in response to violation versus correct sentences in the participants' additional language, Brocanto2.

**Proficiency**

Assessing proficiency in any given language (first, second, or third) is not a straightforward process. The notion of ‘proficiency’ often applies to the assessment of skills in or knowledge of an additional language, though it is also relevant in a first, or native, language (L1). For example, L1 speakers may have varying levels of proficiency depending on the linguistic skill (written versus oral, for example) and level of education or other variables (Mulder & Hulstijn, 2011; Pakulak & Neville, 2004; Pakulak & Neville, 2010). In studies on bilingualism, proficiency has been operationalized in various ways, not all of which are discussed here. For example, proficiency has been operationalized as the course level that participants were enrolled in at the time of the study (Kaushanskaya & Marian, 2009a; Lado, 2008; Lado, Lin, & Sanz, 2011). Alternatively, proficiency has been assessed by using participants’ self-ratings of language skill across the four modalities of reading, writing, speaking, and listening and contexts of language use (Bialystok, 2006; Bialystok, Craik, & Ruocco, 2006; Fernandes, Craik, Bialystok, & Kreuger, 2007; Gillon Dowens, Vergara, Barber, & Carreiras, 2009; Stafford, 2011). Proficiency has also been assessed independently by using either native-speaker raters (Abrahamsson & Hyltenstam, 2009; Gillon Dowens, et al., 2009) or a standardized test (Gillon Dowens, Guo, Guo, Barber, & Carreiras, 2011; Sanz, 2000). Performance on psycholinguistic tasks such as picture-naming or grammaticality judgment tasks has also been used as a measure of proficiency in some studies (Abrahamsson & Hyltenstam, 2008, 2009; DeKeyser, 2000; Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2011; Johnson & Newport, 1989). Thus, it should be clear from the various ways in which proficiency has been
operationalized in studies on language learning and processing that it is not a unitary concept and
does not have a single, all-encompassing operationalization.

For the purposes of the present study, two operationalizations of proficiency must be
distinguished. Specifically, it is necessary to differentiate between proficiency in the target
language of the study and proficiency in the already-learned languages of the bilingual subjects
involved in the study. With regards to proficiency in the target language, Brocanto2, a distinction
between low and high experience or proficiency is made. Low proficiency/experience refers to
above-chance accuracy on two consecutive blocks of practice in Brocanto2. High experience
refers to the completion of all practice blocks, regardless of accuracy performance on these
blocks (total number of blocks = 44). Though the criteria for low and high experience differ (i.e.,
one is based on accuracy and the other is based on amount of practice), these operationalizations
will allow the study to investigate the neurocognitive effects of different instructional contexts in
learners that receive the same amount of practice but may, in fact, differ in their L2 or L3 ability
(as assessed by accuracy over the course of practice blocks).

As mentioned above, this study focuses on bilinguals that demonstrate similar levels of
high proficiency in their two languages, which in this study are Mandarin and English.
Proficiency in these already-learned languages was determined by considering data from self-
reports. Self-report was used to assess proficiency in this study for several reasons. First, self-
reports have been suggested to be accurate indexes of actual language proficiency (Bachman &
Palmer, 1985; MacIntyre, Noels, & Clement, 1997; Marian, Blumenfeld, & Kaushanskaya,
2007; Ross, 1998; Shameem, 1998; Stefani, 1994), with high correlations being reported for self-
assessment and external measures of proficiency (Chincotta & Underwood, 1998; Dunn & Fox
Tree, 2009; Flege, MacKay, & Piske, 2002; Flege, Yeni-Komishian, & Liu, 1999; Jia, Aaronson,
& Wu, 2002; Marian, et al., 2007). This suggests that this method of gathering proficiency information has high reliability and validity. Second, the use of self-reports fits well with previous studies on bilingualism, and in particular cognitive advantages of bilingualism (Bialystok, 2006; Bialystok, et al., 2006; Paap & Greenberg, 2013). Third, the standardized tests available for Mandarin and English are likely not fit for testing the bilinguals who are participating in this study. For example, both the HSK (Hanyu Shuiping Kaoshi, Mandarin standardized test; HSK Center) and TOEFL (Test of English as a Foreign Language, English standardized test; Educational Testing Services) are designed for adult L2 learners of the respective languages and thus pose problems of reduced validity in using such measures to assess bilingual language skill, as well as being vulnerable to ceiling effects in the highly proficient, relatively balanced bilinguals sought in the present study. One available standardized proficiency test that has been used on both native speakers and adult L2 learners, the TOAL-3 (Test of Adolescent and Adult Language; Pearson Education, Inc.; Newman, Tremblay, Nichols, Neville, & Ullman, 2012; Pakulak & Neville, 2010), would not have the problem of ceiling effects or reduced validity, but has no Mandarin equivalent. Therefore, native-like performance in English for that test would have no comparable Mandarin measure with which to establish similar high proficiency in the two languages.

The self-reports were gathered using the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, et al., 2007) which has been shown to correlate with independent measures of proficiency such as reading fluency, productive vocabulary, and grammaticality judgment accuracy and reaction time, among others. These self-report measures of language skill were used to determine proficiency. High proficiency in the already-learned languages thus refers, somewhat broadly, to high self-ratings of proficiency in speaking,
Declarative and procedural memory systems

Two of the main neurocognitive theories of late-learned language reviewed in Chapter 2 discuss differences in L1 and L2 processing in terms of two memory systems: declarative and procedural. Therefore, information about each system is provided here to facilitate understanding the two neurocognitive theories discussed later.

Declarative memory

The declarative memory system is involved in the encoding, storage, and retrieval of information about semantic (facts) and episodic knowledge (events). Information in declarative memory is consciously accessible and learning within the system can occur after a single exposure, although the system itself appears to be associative in nature and thus frequency and attention to input strengthen encoding (Eichenbaum & Cohen, 2001)

Medial temporal lobe (MTL) brain regions – the hippocampus and related structures-with connections to temporal and parietal neocortical regions, subserve declarative memory processing. Learning of new information involves all parts of the system whereas consolidation of this information appears to rely on the MTL components (Eichenbaum & Cohen, 2001; Squire & Knowlton, 2000). Information eventually becomes independent of these structures and reliance shifts to neocortical regions in the temporal lobes (Squire, 2004; Squire, Clark, & Knowlton, 2001; Squire, Stark, & Clark, 2004; Squire & Zola, 1996). Other brain regions have
also been proposed to underlie declarative memory including the anterior ventral-lateral prefrontal cortex (PFC), areas of the frontal cortex corresponding to Brodmann's Areas (BAs) 45 and 47, and the right cerebellum (Ullman, 2004, 2013) as well as the cingulate cortex (Eichenbaum, 2002). Neurological insult to these regions impairs an individual's ability to create new memories (anterograde amnesia) and also results in partial to severe loss of old memories (retrograde amnesia), as in the famous case of patient HM following resection of his MTL structures as a treatment for intractable epilepsy (Corkin, 2002).

Declarative memory ability varies within a single individual throughout the lifespan and across individuals depending on sex, genes, handedness, disease, or other biological variables (Eichenbaum & Cohen, 2001; Ullman, 2005, 2013). Enhanced declarative memory ability, whatever the origin, is indexed by superior performance on tasks that involve the conscious storage and retrieval of information (verbal and non-verbal), usually from short-term memory but also associated with performance on working memory tasks (Carpenter, 2008). The knowledge held in declarative memory is largely accessible to conscious awareness and traditionally conceived of as being explicit in nature, although it does not have to be (Chun, 2000). Additionally, knowledge in declarative memory, or access to this knowledge may be more prone to weakening, or forgetting, than knowledge in the procedural memory system (Foerde, 2007; Takashima, et al., 2006)

**Procedural memory**

Procedural memory has been less thoroughly studied compared to declarative memory, due in large part to the implicit or unconscious nature of procedural memory functioning. Procedural memory refers to the skill and habit learning system (Squire & Knowlton, 2000) involved in the learning and control of long-established motor and cognitive functions. It
essentially underlies the 'how' of task or action performance (Beggs, et al., 1999) and may be specialized for the learning of sequences and rules (Henke, 2010; Poldrack & Foerde, 2008; Squire & Schacter, 2002). This system learns gradually and requires practice – the representations that result from learning are processes or routines as opposed to the "items" stored in declarative memory (Christiansen, Dale, Ellefson, & Conway, 2002; Eichenbaum, 2002; Ullman, 2004). The knowledge is not available to conscious awareness (Squire & Zola, 1996; Tulving, 2000), though once learned, procedural knowledge, or access to it, applies automatically and reliably (Morgan-Short & Ullman, 2011).

Procedural memory is subserved by frontal, temporal, parietal brain regions, the cerebellum, and basal-ganglia structures (Gabrieli, 1998; Squire & Zola, 1996; Ullman, 2004, 2013). The basal ganglia underlie motor, emotional, and cognitive processes (Gabrieli, 1998) and enjoy a generous amount of connectivity throughout the brain by virtue of the striatal-thalamic-cortical loop (Eichenbaum, 2002). The frontal lobes also subserve the procedural system (outputs from the basal ganglia via the thalamus project to frontal lobes), particularly the pre-motor, supplementary motor, pre-supplementary motor, and motor areas as well as inferior frontal lobes, including BA 44 (Eichenbaum, 2002; Ullman, 2004). The cerebellum is involved in motor control and learning (Eichenbaum, 2002). Neurological abnormalities within the procedural memory system have been shown to be related to deficits in an individual's ability to carry out normal motor tasks, as with the motor difficulties involved in late-stage Parkinson's disease (Ullman, 2004), and also to compromise language functions related to grammatical processing, as in non-fluent aphasics or children with Specific Language Impairment (Hedenius, et al., 2011; Hedenius, et al., In Press; Ullman & Pierpont, 2005).
Procedural memory ability, like declarative memory, varies as a function of individual differences such as sex, handedness, genes, and other biological variables (Ullman, 2005, 2013). Enhanced procedural memory ability has been indexed using tasks such as motor sequence learning, serial reaction time, and probabilistic classification learning (Carpenter, 2008).

Declarative and procedural memory are dissociated brain systems that carry out different cognitive functions, but the functioning of each memory system in relation to the other is in fact quite dynamic (Packard & Knowlton, 2002; Poldrack & Packard, 2003; Ullman, 2004, 2013). For example, deficits in processing ability for one system can be compensated for by the other system. In a study on amnesiacs and matched controls, Bayley and Squire (2002) showed that when declarative-like information is involved in the learning task, but declarative memory is compromised, procedural memory can do the learning – although it is slower and outside of awareness. Similar compensatory behavior by declarative memory in the face of procedural deficits has been found (Eichenbaum, 2002).

Declarative and procedural memory are also competitive and when either system can carry out the same task it is often the strength of representations or speed of processing that determine which system is engaged. Interestingly, when a particular system does become engaged for the certain learning task there appears to be concomitant impairment of the other system with regards to performance on that task (Poldrack & Packard, 2003).

Finally, the two memory systems are cooperative. Impaired declarative memory, for example, may compromise the ability to learn the cues relevant for implicit pattern learning in procedural memory (see Knowlton, Squire, & Gluck, 1994) on the Weather Prediction Task, WPT). When considering the dynamic interplay between reliance on each of these memory systems, it is important to remember that shifts in reliance do not constitute transfer of
Chapter 2 – Review of the Literature

Introduction

The following sections review studies on bilingual L3A compared to monolingual L2A. Following these reviews, various neurocognitive theories of late-learned language processing are provided. Finally, the use of event-related potentials (ERPs) in the study of language is discussed. The empirical and theoretical content reviewed will motivate the research questions and predictions offered at the end of this chapter and the research design discussed in Chapter 3.

Bilingual L3A

Regarding bilingual L3A compared to monolingual L2A, first studies conducted in dual-language environments will be reviewed, followed by those conducted in single-language environments. Finally, studies conducted in laboratory settings will be discussed. Where suitable, distinctions among the nature of the effects (positive, negative, or none) will be made.

Dual-language environments

Positive Effects

In one of the first studies to directly examine the effects of bilingualism on additional language acquisition Cenoz and Valencia (1994) compared the English performance of 320 Basque-Spanish bilinguals and Spanish monolinguals high-school students in the Basque Country. English performance was measured using an English oral interview, listening comprehension, reading comprehension, a 250-word writing sample, and 3 multiple-choice vocabulary and grammar tests. The authors also gathered information on intelligence, socioeconomic status (SES), exposure to English, and attitudes towards the target language environment. Their regression analysis indicated that the bilingualism of the Basque-Spanish
students predicted English performance independently of other factors that were also found to be predictive on the outcome measures, such as motivation and exposure to English. This study was among the first to provide indicate that facilitative effects of bilingualism on additional language acquisition may exist, but the sociolinguistic context of data acquisition made it difficult to control for possible confounds in amount and type of English input. Also, because the performance measures were mostly scholastic and entirely product-based, strong conclusions about the roots of the facilitative effects are difficult to make.

Lasagabaster (2000) also examined English performance in the Basque Country, but on younger learners (10-11 and 13-14 years old, no monolinguals) in three educational contexts: (1) a majority language program where Spanish was the language of instruction and Basque was taught as a second language, (2) a partial immersion program where both Spanish and Basque were the languages of instruction, and (3) a minority language program where Basque was the language of instruction and Spanish was taught as a second language. English performance was measured using vocabulary and grammar tests as well as tests of reading, writing, speaking, and listening. Lasagabaster also collected information on SES, exposure to English outside of school, age, IQ, gender, and motivation. The results revealed that the 10-11 year olds in the immersion program performed better on English measures than age-matched students in the majority language program. The results also showed that 13-14 year olds in the immersion program outperformed age-matched students in both of the other educational programs (minority and majority design). This study suggests that the variable bilingualism may not be enough to facilitate L3A, but that relatively high bilingualism, or balance, between the two languages is important in terms of bilinguals L3A outcomes.
Sanz (2000) conducted a study in Catalonia, Spain and replicated the above results for facilitative effects of bilingualism on English performance. One-hundred and twenty-four biliterate Catalan-Spanish bilinguals and 77 literate Spanish monolinguals took part. All participants were high-school aged and English proficiency was measured using the vocabulary and structure sections of the CELT English proficiency test. Additional information about age, gender, motivation, SES, intelligence, exposure to English, and attitudes towards British and US populations was also gathered. Mirroring the results of Cenoz and Valencia (1994), bilingualism was a significant predictor in English language achievement, independent from the other variables included in the regression analysis. The study suffers, however, from the same limitations as those provided for Cenoz and Valencia (1994) (and also relevant for Lasagabaster, 2000), namely that only coarse and academically-centered outcome measures were used and type and amount of English input were not reliably controlled.

Thomas (1988) tested French proficiency in 16 adult English-Spanish bilinguals and 10 adult English monolinguals in an unofficial dual-language environment. The study was conducted in south Texas in the United States of America, where 51% of the population is Hispanic, so although Spanish is not an official partner language with English, it is clearly present in the demography of the region. The bilinguals in the study were divided into two groups according to the type of Spanish exposure they had received; 10 learners has been exposed to Spanish in a formal (classroom) setting whereas 6 learners had learned Spanish informally at home. Motivation, SES, aptitude, amount of French exposure, French instructor, textbook, and teaching method were controlled for when evaluating French proficiency in the two groups. Proficiency was measured after one semester of French using vocabulary
(translation) and grammar (sentence completion) tests in addition to a 10-sentence writing sample.

The results showed that bilinguals outperformed monolinguals in both vocabulary and grammar. Additionally, the results showed that the bilinguals with formal education in both languages outperformed those with formal education only in English for grammar; this difference was not found for vocabulary performance. The author attributed this difference in grammar versus vocabulary performance to a difference in metalinguistic awareness resulting from formal language training. This study attempted to control for amount of exposure by gathering information on teacher, textbook, and classroom methodology, but the small number of participants in the study weakens the impact of the results. Moreover, the study likely suffers from confounds of language transfer, given that French and Spanish are both romance languages and this genetic similarity may have unduly conferred an advantage on Spanish speakers that is independent of the participants' bilingualism.

To my knowledge, there are three studies that have been conducted in a dual-language environment and have chosen a specific linguistic domain (as opposed to broad language proficiency) as the focus of the study. The first, Safont-Jorda (2003), measured pragmatic competence of request acts in English by 80 Spanish monolinguals and 80 Catalan-Spanish bilinguals (all female) enrolled in a university-level English for Academic Purposes (EAP) course at a Catalanian university. Pragmatic competence was measured using a Discourse Evaluation Test (DET) that consisted of 18 discourse situations for subjects to evaluate according to the appropriateness of the discourse in its provided context as well as a role-play scenario and open discourse completion task to assess pragmatic production. The results showed that bilinguals outperformed monolinguals on the DET and also pragmatic production for
English, confirming a facilitative effect of bilingualism on target language pragmatic awareness.

In a study conducted in Iran, Keshavarz and Astaneh (2004) investigated the impact of bilingualism on English vocabulary learning; measuring productive vocabulary ability in 30 Turkish-Persian bilinguals, 30 Armenian-Persian bilinguals, and 30 Persian monolinguals (all female). All subjects were high-school aged and considered to be at an intermediate English level. They were matched in terms of SES, hours of exposure to English, and teaching methodology used. The bilingual subjects differed with respect to the instructional status of their two languages. The Turkish-Persian bilinguals were enrolled in a school located in a Turkish-speaking city, but where the language of instruction was Persian. The Armenian-Persian bilinguals studied both languages academically. Subjects' vocabulary ability was measured using a Controlled Productive Ability Test at levels corresponding to 2000 and 3000-word lexicons (18 items per level). Results revealed, again, that bilinguals outperformed monolinguals and additionally that, like Thomas (1988) and Lasagabaster (2000), bilinguals receiving formal education in both of their languages outperformed those receiving formal education in only one of the two languages.

Finally, Zare and Mabarakeh (2013) investigated potential monolingual-bilingual differences in vocabulary learning. In their study, Iranian learners of English as a foreign language were divided as to whether they were Arabic-Persian bilinguals (n = 50, all male) or Persian monolinguals (n = 50, all male). The bilinguals used Persian at school and Arabic at home. Both monolinguals and bilinguals were given classroom instruction for 5 weeks on a list of 50 words (10 words each week) and two weeks following instruction, participants took recognition and production tests. For general L3 vocabulary performance (recognition plus production scores), the results showed that bilinguals outperformed monolinguals. This was also
true for the production test alone, though there were not group differences for recognition. The authors suggest that their results may stem from bilinguals’ ability to take advantage of their two languages and make associations between the L3 and their already-learned languages, at least for learning new words.

These studies provide important evidence that bilinguals may show advantages at learning an additional language as adults. However, nearly all of them have a number of methodological weaknesses, especially a lack of control over amount and type of additional language exposure, and for the final three a limited subject sample (all females or all males). Additionally, the majority of the studies tested only general language proficiency, and every study implemented only behavioral measures of additional language performance; thus they provide no insights into mechanisms underlying L3 development or processing and how it may compare with monolingual L2A.

No or Negative Effects

The studies reviewed above suggest that bilingualism may facilitate additional language learning, but another study conducted in a dual-language environment failed to find clear evidence for such an advantage. González-Ardeo (2000) compared lexical learning and phonetic production between 12 Spanish monolinguals and 36 Basque-Spanish bilinguals. The subjects were college-aged students in the Basque Country who had been enrolled in three semesters of English for Science and Technology (EST) lessons. Lexical learning was measured using a fill-in-the-blank vocabulary task and phonetic production was measured by a reading task. There was no difference between monolinguals and bilinguals on either lexical learning or phonetic production, though there was a trend for bilinguals to outperform monolinguals. Follow-up analyses to explore this trend revealed that positive effects of bilingualism emerged when
considering when dividing the subjects according to whether Spanish or Basque was their L1 and what education program they were enrolled in. However, this resulted in small sample sizes (after dividing 36 bilinguals into 6 groups based on L1 and language-of-education), and thus the claims of a bilingual advantage in this study speculative.

To summarize, evidence for bilingual advantages at additional language learning have been found in the majority of studies conducted in dual-language environments, especially on measures of general language proficiency (Cenoz & Valencia, 1994; Sanz, 2000) and for balanced bilinguals (Keshavarz & Astaneh, 2004; Lasagabaster, 2000), but also for more specific linguistic domains such as pragmatics (Safont-Jorda, 2003). However, these studies also suggest that bilingual advantages may not apply to additional language acquisition in toto, particularly with respect to aspects of vocabulary (e.g., Gonzalez-Ardeo, 2000; Thomas, 1988).

**Single-language environments**

**Positive Effects**

Klein (1995) did not test bilinguals, but did compare monolinguals (of different language backgrounds) and multilinguals (also of different language backgrounds) on their acquisition of English as an additional language. She tested 17 monolinguals and 15 multilinguals within a Universal Grammar framework. The participants in this study were in junior high school (12-15 years old) and the author was specifically interested in examining parameter resetting for preposition stranding in English. Multilinguals had English as either their L3 or L4 and had started learning their L2 in early childhood whereas the monolinguals began learning English as an L2 in early adolescence. Klein attempted to control for possible transfer effects by only including multilinguals whose languages were equivalent in terms of wh-question formation (i.e., wh-movement), but also in not permitting preposition stranding. The results from a
grammaticality judgment and correction task containing equal numbers of declarative and interrogative sentences (6 each) showed that while both groups evidenced parameter resetting, the multilinguals were able to reset faster than the monolinguals, suggesting an advantage for experience in multiple languages over experience in just one. This advantage was not obvious in the outcome (parameter resetting), but was found to be process-related (rate of resetting). This study is one of only a handful that have focused on a specific linguistic domain (morphosyntax) when comparing the effects of language experience on additional language learning¹ and notably suggests that facilitative effects can emerge in single-language environments. However, the small number of participants and test items are a clear limitation in generalizing the results to other single-language environments.

No or Negative Effects

Two studies in single-language environments have failed to find facilitative effects of bilingualism on additional language acquisition. Gibson, Hufeisen, and Libben (2001) tested 64 subjects on their acquisition of German prepositions; subjects were divided into 4 bi-or multilingual groups and 2 monolingual groups. The researchers measured acquisition by using a fill-in-the-blank task and the results showed no differences among the groups. Participants in this study came from a variety of L1 backgrounds, however, and completed proficiency self-ratings for their additional languages. This implies that there was little to no control over the type of German exposure they had received prior to testing or over the extent of bilingualism or

¹ Several other studies have approached L3 processing from a Universal Grammar perspective, but these studies are interested in cross-linguistic influence for L1, L2, and L3 as opposed to the isolated role of prior language experience in subsequent additional language processing (Bardel & Falk, 2007; Falk & Bardel, 2010; Iverson, 2010; Rah, 2010; Rothman, Iverson, & Judy, 2011).
multilingualism in the subjects. Additionally, the researchers used only one (mechanical) measure of additional language performance. Therefore, the absence of facilitative effects of bi- (or multi-)lingualism in this study could be related to methodological shortcomings and not, or less so, to a true absence of facilitative effects.

Okita and Jun Hai (2001) also failed to find evidence for a bilingual advantage at additional language learning. Their study was conducted on the acquisition of Japanese kanji (one of the scripts in the Japanese writing system). Sixty Chinese monolinguals in their second year of Japanese study and 84 Chinese-English bilinguals in their second and third years of Japanese study (46 and 38 students, respectively) were tested. Writing performance was measured by stroke order and a kanji writing test. The results of this study revealed, contrary to most research on bilingual L3A versus monolingual L2A, that monolinguals outperformed the bilinguals.

**Summary: Dual-language and single-language environments**

All of the studies reviewed above for single-language environments focused on specific aspects of language (i.e., UG parameter resetting, preposition acquisition, writing) and only one found the predicted facilitative effects for previous language experience (Klein, 1995). Also, this facilitative effect was only found for processing (rate of parameter resetting) and not on the outcome measure itself, for which for mono- and multilinguals were the same. The dual-language environments reviewed above mainly focused on more general language outcomes and found evidence for advantages of bilingualism in additional language acquisition (Cenoz & Valencia, 1994; Lasagabaster, 2000; Sanz, 2000). However, when considering more specific aspects of language, advantages were found in some studies (Keshavarz & Astaneh, 2004; Safont Jorda, 2003; Zare & Mobarakhe, 2013) and not others (Thomas, 1988).
Based on these studies conducted in real-world settings, it is difficult to construct a definite conclusion as to whether bilinguals show an advantage over monolinguals at learning an additional language in adulthood. This is likely due to the methodological heterogeneity across the studies, and partly also to methodological concerns in their experimental designs. Of the eleven studies reviewed above, eight found bi-or multilingual advantages. However, two of these eight studies (one dual-language, one single-language) did not test adults, rather their subjects were middle-school or high-school aged learners (Klein, 1995; Lasagabaster, 2000), and neither of these specifically compared monolingual L2A to bilingual L3A. Of the remaining six studies, two (both dual-language) tested only females (Keshavarz & Astaneh, 2004; Safont Jorda, 2003) and one tested only males (Zare & Mobarakeh, 2013). In the three remaining studies that found bilingual advantages at additional language learning (all dual-language), one likely suffers from language transfer confounds (Thomas, 1988) and the remaining two used only general measures of proficiency (Cenoz & Valencia, 1994; Sanz, 2000), which does not provide any linguistic detail about the potential advantage. Finally, these studies tested L3 performance at only one stage in their learning, and thus fail to capture potential differences in monolingual L2A and bilingual L3A that may change as a function of L3 experience. Thus, in real-world language settings there is in fact no research which has empirically tested the question of whether bilinguals show an advantage at learning an additional language in adulthood whose findings are not weakened by the subject sample (all women, all men, multilinguals, not adults) or the methodology (language transfer, only broad L2/L3 proficiency).

**Laboratory Studies**

Controlling for external variables such as language transfer, or type and amount of L2 or L3 exposure, is often difficult in real-world learning settings. Sanz and Lado (2008) recognized
the inherent limitations of conducting L3A research in real-world settings and suggest that laboratory designs are the most useful method to use when attempting to empirically address the relationship between cognitive benefits associated with bilingualism and bilingual L3A. Specifically, laboratory settings offer precise control over input conditions, which allows for a more fine-grained analysis of bilingual L3A compared to monolingual L2A, and the studies reviewed below indeed provide precise control over aspects of the additional language input. In particular, this precise control allows one to compare the effects of different types of input, or instructional contexts, with high reliability.

**A note on the effectiveness of explicit and implicit training in monolingual L2A**

A large body of research in SLA has investigated the effects of explicit and implicit instructional conditions on monolingual L2A. In their meta-analysis of the effectiveness of these two instructional conditions across 49 studies, Norris and Ortega (2000) concluded that explicit conditions were more effective than implicit conditions in terms of L2 outcomes, but noted that the assessment tasks in the majority of these studies (90%) favored explicit memory-based performance as opposed to freely productive (more automatic) use of language. Thus the advantage of explicit over implicit instructional treatments may have been an artifact of the tasks used to assess language ability, and not a true, qualitative superiority.

However, the results from a more recent meta-analysis which measured the effectiveness of explicit and implicit instruction on simple and complex linguistic targets echo the findings from Norris and Ortega (2000). Spada and Tomita (2010) conducted a meta-analysis on 30 published studies (10 of which were also in the Norris and Ortega (2000) study) and found that explicit instruction was more effective than implicit for both simple and complex grammatical targets. Additionally, though many of the studies included in this meta-analysis were found to
use tasks favoring explicit knowledge, an equal number (33%) of less controlled tasks were employed, suggesting that the bias towards explicit knowledge is becoming less pronounced. However, Spada and Tomita (2010) note that the implicit conditions in the studies consisted of 10 hours or less of instruction and suggest that advantages of implicit instruction may not be detectable after such a short time, since learning under implicit conditions takes longer than learning under explicit conditions (N. C. Ellis, 2005a, 2005b). Thus, though these meta-analyses suggest advantages for explicit compared to implicit conditions, SLA researchers continue to caution that the debate has yet to be reliably resolved.

**Positive Effects**

Regarding the effectiveness of explicit and implicit conditions and their interaction with bilingual L3A, the studies reviewed here provide a mixed picture. One of the first laboratory studies to investigate potential advantages of bi-or multilingualism on language learning under different instructional contexts was Nation and McLaughlin (1986). They used an artificial grammar (AG) to test learning and exposed 14 monolinguals, 14 multilinguals, and 14 bilinguals to implicit and explicit conditions. The artificial grammar was composed of 20 exemplar strings of letters (3 to 6 in a string) and both learning conditions were exposed to the same AG input; the difference between groups was the instructions they were given prior to exposure. The learners in the implicit condition were given no specific instructions to learn the rules of grammar while those in the explicit condition were informed that the system was rule-based and they should discover the rule. At test all participants were shown 50 strings, 25 of which were ungrammatical, and were asked to judge the well-formedness of each item (yes/no response) based on what they had seen during exposure.
The results showed that multilinguals outperformed both bilinguals and monolinguals under the implicit learning condition, but that groups were not different when exposed to the explicit learning condition. The researchers argued that the multilinguals could be considered 'expert' language learners and their increased language learning experience made them uniquely capable of forming abstract representations from linguistic input; thus making them better than less experienced (novice) language learners at novel grammar learning. They went on to argue that multilinguals are able to employ automatic skills more than less experienced learners and are more flexible in the face of tasks requiring demanding, controlled processing.

This study is among the first to have compared learners with different levels of language learning experience across controlled learning conditions and suggests that prior language experience yields advantages under more implicit conditions. However, language learning experience in this study was operationalized according to self-ratings by the subjects so the results are not wholly clear and must be evaluated with caution. Additionally, the researchers used an AG paradigm to investigate their 'expert' versus 'novice' research questions and while AG paradigms are theoretically informative with regards to the human capacity to learn the rules of complex systems that are language-like (see Reber, 1967, 1989 for example), they lack the characteristic form-meaning mappings of natural language systems and are thus incapable of producing robust results concerning language learning and processing on a practical and generalizable scale.

Nayak, Hansen, Krueger, and McLaughlin (1990) included form-meaning mappings in an experiment in which they trained 24 monolinguals and 24 multilinguals under memory and rule-search exposure conditions for an artificial language. The language in this study consisted of 40 sentences with up to five consonant-vowel-consonant trigrams and the lexicon of the language
was mapped to specific referents; the sentences appeared randomly on a computer screen 3 times for 5 seconds each. Learners in the memory condition were asked to memorize the items and those in rule-search condition were instructed to look for the rules that determined word order in the grammar. Additionally, learners were asked to pause 3 times during the learning phase to verbalize instructions for another subject to do the task in the same way (yoked controls). After the training phase, learners completed vocabulary and syntax tests. The vocabulary assessment consisted of 15 words of five different types and the participants' task was to say aloud whether the word and its corresponding referent appeared during training. To assess syntax, participants performed a grammaticality judgment task.

The results showed that no group differences existed for vocabulary learning of the artificial language, but that multilinguals outperformed monolinguals in learning the rules of word order (syntax) when they were informed that such rules existed (rule-search condition). Under the memory condition, no multilingual-monolingual differences were found. These results run contrary to those found in Nation and McLaughlin's (1986) study and the researchers suggest this may stem from the increased complexity in their artificial language as compared to the artificial grammar used by Nation and McLaughlin (1986). Additionally, the studies differ as to what characterized their implicit conditions so the results are not strictly comparable. One strength of this study is that it attempted to incorporate the characteristic form-meaning mappings of natural language into its stimulus design, however the participants were again grouped according to self-rating of reading, writing, and speaking skills in order to assign monolingual or multilingual status which again blurs the reliability of the study's findings. Additionally, the authors used the learners' verbalizations to suggest processing differences between their subjects, namely that multilinguals were more capable of adjusting their strategies than monolinguals.
They did not, however, specifically isolate strategy-use as a variable in the study and thus the proposed explanation for group differences are based on an ad-hoc evaluation of the outcome rather than on focused measures of processing differences between individuals with more and less language experience. Nonetheless, both of the aforementioned studies have as a core strength their experimental control over amount and type of input as well as the learning condition under which participants were exposed to the input.

Lado (Lado, 2008; see also Lado, Bowden, Stafford, & Sanz, 2013) also used a laboratory-based design, but used a natural language as opposed to an artificial language or grammar. This study employed computer-based instruction and subsequent tests of learning for agent/patient case assignment in Latin. The aim of the study was to investigate how external conditions, operationalized as more and less explicit reactive feedback, interact with level of bilingualism, cognitive capacity, and L3 development. One-hundred and fifty-one native English-speaking learners of L2 Spanish were grouped into levels of bilingualism according to their current academic status (college-level Spanish course enrollment) and thus assigned to Basic, Intermediate, Advanced, and Native-like levels. Learners were exposed to a computerized Latin lesson consisting of task-essential practice and feedback-during-practice with the difference between groups being the degree of explicitness in the feedback provided. More explicit feedback was operationalized as information about the correctness (right or wrong) of the learner's response paired with metalinguistic information about the target form. Less explicit feedback consisted of providing learners with information only about the correctness of their response (right or wrong) with no additional metalinguistic information. After the treatment phase subjects were tested in 4 ways to measure L3 development: written interpretation, aural interpretation, written grammaticality judgment, and sentence production.
The results showed that the highest level of bilingualism (Native-like) was associated with an advantage over monolinguals (Basic level subjects) for aural interpretation, grammaticality judgment, and sentence production and that even an Intermediate level of bilingualism was related to advantages over monolinguals for written interpretation. When considering the effects of more and less explicit feedback, the results showed that under more explicit conditions, Intermediate and above were different from monolinguals, but that under the less explicit learning condition only the highest level of bilingualism conferred advantages for L3 development. The results are similar to those of Nation and McLaughlin (1986) in that greater amounts of language experience affected performance only under less explicit learning conditions. The results are also somewhat similar to the results in the Nayak et al. (1990) study that found the explicitness of the task to favor advantages in additional language performance for more experienced learners (i.e., even Intermediate levels were able to outperform monolinguals when provided with explicit information).

However, these three laboratory studies are not wholly comparable because the nature of the instructional conditions varied across each study. Both Nation and McLaughlin (1986) and Nayak et al. (1990) manipulated the external conditions of the input according to task instructions (i.e., memorize or search for the rule) whereas Lado (2008; Lado et al., 2013) had participants in equivalent task-instruction conditions, but manipulated the nature of reactive feedback provided during practice. Additionally, the learners in Lado (2008; Lado et al., 2013) were individuals whose L2 was Spanish, a romance language descendent from Latin. Thus, possible effects of transfer based on the genetic relatedness of the L2 and the target L3 cannot be divorced from the effects that level of bilingualism had on L3 outcomes. Finally, only one study actually compared bilinguals with monolinguals Nation & McLaughlin (1986) whereas the other
two compared multilinguals with monolinguals (Nayak et al., 1990) or L2 learners with different L2 proficiency levels (Lado, 2008; Lado et al., 2013). (Note that Nation and McLaughlin (1986) did not find differences between the bilinguals and monolinguals).

There are also a series of studies which found advantages for bilinguals in additional language learning (at the level of vocabulary), but did not compare instructional conditions. Kaushanskaya and Marian (2009a) investigated the effects of bilingualism on novel word learning. Their subjects were 20 English monolinguals, 20 early Spanish-English bilinguals and 20 early Mandarin-English bilinguals. The subjects were exposed auditorily to 48 disyllabic novel words while they viewed the English translation of the word on a computer screen. Listening was self-paced and subjects were asked to repeat each novel word and its translation aloud three times. The testing phase consisted of one recall and one recognition test, immediately after vocabulary learning and one week later. The results showed that the two bilingual groups were not different from each other in terms of performance, but were better than monolinguals on recall and recognition at immediate testing, and also after at the one-week delayed test. Thus, the authors concluded that early bilingualism is related to advantages in novel vocabulary learning later in life.

Kaushanskaya and Marian (2009b) investigated the effects of bilingualism on native-language interference in novel word learning. The subjects in this study were 24 early English-Spanish bilinguals and 24 English monolinguals. The materials were similar to those used in Kaushanskaya and Marian (2009a), but the target of study was the learning of novel words that contained familiar and unfamiliar phonemic information, and both monosyllabic and disyllabic words were constructed. The artificial phonemic inventory consisted of four phonemes that do not exist in either Spanish or English and four that exist in English only. There was letter-sound
correspondence for the four familiar English phonemes (/n/-N, for example), but not for the four unfamiliar phonemes. Thus the experimental mismatch consisted of words that overlap with English orthography, but do not match English phonology (the four unfamiliar sounds).

There were two learning contexts in this study; each context corresponded to one list of vocabulary such that half of the novel words were learned in one context and half in the other context. In the hearing-only context subjects heard the novel word and saw the English translation on the computer screen. In the hearing-and-seeing context subjects both heard and saw the novel word (and saw the English translation). Like Kaushanskaya and Marian (2009a) subjects were asked to repeat the novel word out loud three times and vocabulary learning was assessed with recognition and production tests. These tests were administered immediately after testing and following a one week delay.

The results showed that bilinguals outperformed monolinguals on vocabulary learning for recognition and production at both immediate and delayed testing for the words learned bimodally (the hearing-and-seeing phase). In the hearing-only context, bilinguals outperformed monolinguals at immediate but not delayed testing, for both recognition and production. These results led the authors to conclude that the bilinguals were better able than the monolinguals to overcome mismatches between native-language orthographic-phonological information, particularly in a bimodal learning context.

In another study on word learning, Kaushanskaya (2012) used a similar paradigm as Kaushanskaya and Marian (2009a,b). In experiment 1, English-Spanish bilinguals (n = 18) and English monolinguals (n = 36) were trained on phonologically familiar novel words (English phonology). The monolinguals were divided into high (n = 18) and low (n = 18) phonological short term memory span as measured by a Digit Span task (note: not enough low-span bilinguals
could be recruited). The bilinguals and high-span monolinguals were not different from each other on this memory measure. Training was similar to that employed in Kaushanskaya and Marian (2009a): the novel word was presented twice auditorily while the English translation was displayed on a computer screen. Again, participants repeated the novel word and its translation aloud three times. Word learning was assessed by recognition and production tests, immediately after training and also after a one-week delay.

The results from experiment 1 showed that English-Spanish bilinguals outperformed both high and low-span English monolinguals on learning phonologically familiar novel words for recognition and production, at immediate and one-week delayed testing. The high and low-span monolinguals did not significantly differ from each on any measure at either testing session. The author concludes from these results that bilingual advantages at word learning are not based on phonological memory differences between bilinguals and monolinguals and that high phonological memory capacity does not necessarily lead to superior word learning in monolinguals. However, the author notes that the phonological familiarity of the novel words may have ‘hidden’ the effects of phonological memory in bilingual advantages and that perhaps such advantages are more obvious in contexts where the novel words are phonologically unfamiliar.

In experiment 2, English-Spanish bilinguals ($n = 18$) and English monolinguals ($n = 36$) were trained on phonologically unfamiliar novel words (these subjects did not participate in experiment 1). Again, the monolinguals were divided into high ($n = 18$) and low ($n = 18$) phonological short term memory span as measured by a Digit Span task. The procedure was the same as that used in experiment 1 and learning was assessed with immediate and one-week delayed recognition and production tests. The results showed that bilinguals outperformed high
and low-span monolinguals at immediate and delayed production tests, and that high-span monolinguals outperformed low-span monolinguals at both test sessions as well. This pattern of results was the same for recognition at immediate testing. For recognition at delayed testing, however, the bilinguals outperformed the low-span monolinguals, but were not different from the high-span monolinguals. High and low-span monolinguals were not significantly different from each other. Overall, the author concludes that bilingualism is beneficial for learning of both phonologically familiar and unfamiliar words, and bilingualism may facilitate phonological memory skills.

In a similar laboratory study, monolinguals and bilinguals were not compared in terms of either explicit or implicit input or differential additional language learning, but in terms of processing – specifically language competition effects in novel language learning. Similar to Kaushanskaya & Marian (2009a,b) and Kaushanskaya (2012), Bartolotti and Marian (2012) also focused on word learning. In this study, 12 early Spanish-English bilinguals and 12 English monolinguals were trained on Colbertian, an artificial language composed of 24 nouns. Training consisted of displaying a picture on the computer screen and presenting the Colbertian word auditorily; participants repeated the word aloud once. Participants in this study were trained until they reached a pre-set criterion of 90% naming accuracy, thus equating all participants in terms of Colbertian proficiency (bilinguals and monolinguals did not differ in number of blocks to criterion).

Following training, participants completed an interference task during which they were asked to choose the picture that corresponded to a Colbertian word presented auditorily. Twelve items in this task contained a picture of the target and also a competitor English word (i.e., a picture of an acorn (target) and a shovel (competitor) presented with the word *shundo*, which
means acorn in Colbertian), 12 control items contained no distractor (i.e., no overlapping item), and for 24 items neither picture depicted the presented Colbertian word, the correct response was a red X. During this task, eye- and mouse-movements were recorded as well as accuracy and reaction time.

For eye-tracking data, the results showed that both monolinguals and bilinguals evidenced early activation (200 milliseconds, ms) of native-language (English) competitors, but that bilinguals resolved this competition earlier (by 700 ms) than monolinguals (up to 1400 ms). Additionally, the monolinguals looked at competitors more than bilinguals and at competitors more than control items. The authors suggest that bilingual experience may affect how interference from non-target languages is managed.

For mouse-tracking data, the results showed that increased competitor activation curved mouse trajectories away from the target in monolinguals, but did not find this in bilinguals. Instead, the competitor activation affected vertical motion but did not disrupt the target approach, suggesting that increased competition actually led to facilitation of the target. The authors suggested that this pattern of results may index bilinguals’ enhanced effectiveness in managing cross-linguistic interference by better facilitating access to the new language material (the target).

Bartolotti et al. (2011) also investigated vocabulary learning, but did not compare bilinguals and monolinguals and instead focused on degree of bilingualism. In this study, participants were divided into low ($n = 11$) or high bilingual experience ($n = 11$) based on L2 proficiency, age of acquisition and frequency of use. Participants were exposed to two languages based off of Morse code (tones and dots); they heard one language under either a low interference condition where pauses between words were long (300 ms) and the other under high
interference condition where pauses between words were short (100 ms). Participants were trained over three blocks, each about 4 minutes long, and after training they completed a two alternative forced choice test where they were asked to decide which of two Morse code words was more familiar. The results revealed that word learning (above chance performance at test) in the low interference condition was influenced by bilingual experience whereby only the high bilingual experience group evidenced learning of the words. Learning in this condition was not correlated with bilingual experience, age of acquisition or frequency of use, but did correlate positively with L2 proficiency. In the high interference condition, neither the high nor low experience bilingual groups performed above chance.

The authors suggest that the influence of high bilingual experience in the low interference condition may be related to a superior ability to hold auditory sequences in working memory for statistical analysis. They also suggest that such superior learning of statistical (implicit) cues may be related to more effective implicit learning mechanisms in bilinguals. The lack of any differences in learning as a function of bilingual experience in the high interference condition was proposed by the authors to be related to difficulty on behalf of both groups in attending to the statistical and pause-based cues simultaneously. Additionally, as these cues did not correlate with each other and thus could not be effectively used together to learn the words, both groups of bilinguals may have been unable to attend to either or alternatively attended to and attempted to (erroneously) integrate both in the learning task.

Finally, Kaushanskaya & Rechtzigel (2012) investigated what the potential root of the bilingual advantage at word learning might be. The participants in this study were 22 English monolinguals and 22 English-Spanish bilinguals. All participants were trained on a set of novel words (selected from the Gupta, et al., (2004) non-word database) that were presented aurally
and paired with their English translation. Participants heard each novel word-English word pair twice during training. The novel words were paired with either concrete (helmet) or abstract (virtue) English words. The testing phase required participants to produce the English translation of the novel words and served as a test of associative learning. The results of the study revealed that both monolinguals and bilinguals retrieved concrete words more accurately than abstract ones, and that bilinguals were more accurate than the monolinguals for the concrete referents. Bilinguals and monolinguals were similar in their accuracy on retrieving abstract referents. The authors interpreted this effect for concreteness in bilingual as compared to monolingual novel word learning as indicating that bilinguals may be more sensitive to semantic information during learning. This heightened sensitivity was proposed by the authors be related to a wider lexical/semantic network in bilinguals. They concluded that bilingual advantages for learning of concrete words may not be related to different learning mechanisms in bilinguals and monolinguals but instead reflect the higher levels of semantic activation in the bilingual versus monolingual lexical/semantic system due to the availability of their two languages.

No or Negative Effects

Lin (2009) conducted a similar study as Lado (2008), but used 90 Mandarin L1-English L2 learners of different English proficiency levels (High, Mid, and Low); thus avoiding the language transfer confound in Lado (2008; Lado, et al., 2013), as neither Mandarin nor English are genetically related to Latin. Lin (2009) used the same instructional condition and feedback conditions as Lado (2008; Lado et al., 2013), but these were adapted for Mandarin speakers. The results showed that level of bilingualism had no differential effect on written interpretation, aural interpretation, or sentence production – this absence of effects was attributed to the experiment's small sample size and correspondingly low power in the statistical analyses. For grammaticality
judgment there was a difference found for Mid versus Low level bilingual groups in the less explicit feedback condition, where the Mid group outperformed the Low group from pre to post testing. For the effects of more or less explicit feedback conditions, the results showed that learners in the more explicit feedback condition outperformed those in the less explicit feedback condition on all four tests.

These results are different from those found in Lado's study with L2 Spanish learners (2008; Lado, et al., 2013), but when considering the low power due to small sample sizes in three out of four tests, it is not surprising that mostly null effects were found for level of bilingualism. One similarity, however, is that under the more demanding (less explicit) task, effects for previous language experience emerged. The learners in Lin (2009), however, were divided into experimental groups based on a non-standardized English proficiency test whereas the Lado (Lado, 2008; Lado, et al., 2013) learners were grouped based on academic status, so it is difficult to assume that the Intermediate learners in Lado (2008) are equivalent to the Mid learners from Lin (2009). Therefore, no strong conclusions can be drawn regarding the role of level of bilingualism in L3 learning. Additionally, Lado (2008) found clear effects for bilingualism with the Near-native learners under less explicit (more demanding) conditions, but Lin's (2009) experiment does not have a Near-native group that would be expected to evidence similar effects. Instead, Lin (2009) found effects of bilingualism between Mid and Low learners in the less explicit condition. Essentially, Lado's (2008) strength of having more levels of bilingualism to compare on L3 development is absent in Lin's (2009) study, which has as its strength a control for language transfer effects (absent in Lado, 2008; Lado, et al., 2013). Paralleled comparisons between the two studies create an opaque set of conclusions regarding bilingualism and L3 development.
Another study conducted within the same Latin laboratory paradigm investigated age of arrival and bilingualism in additional language acquisition, though no monolingual comparisons were made. Stafford, Sanz, and Bowden (2010) exposed 33 biliterate, bilingual Latinos living in the United States of America to Latin under a single instructional condition. The condition provided preemptive metalinguistic information about how case is assigned and reactive metalinguistic feedback on correct and incorrect responses during subsequent task-essential practice. Fifteen of the bilinguals were considered “early” based on their age of arrival in the US (less than 12 years of age) and 18 were considered “late” bilinguals based on having arrived in the US after the age of 16. The results showed that both groups of bilinguals successfully learned the relevant morphosyntactic cues for Latin case assignment and there were no significant differences between the groups early in development (at immediate posttest). The late bilinguals, however, outperformed the early bilinguals for knowledge retention of case morphology (delayed posttest 3 weeks later).

The results were contrary to the authors' hypothesis that early bilinguals would have an advantage over late bilinguals in L3 Latin learning. The authors suggest that the implementation of the most explicit instructional condition may "level the playing field" (p. 179) for learners with regards to individual differences, such as onset of bilingualism. The early bilinguals in this study, however, reported themselves as more proficient in English (their L2) and less proficient in Spanish (their L1), so the lack of effects in this study may be due to the fact that the bilinguals tested were not balanced (recall that other studies found pivotal roles for balance in the bilingual participants (e.g., Thomas (1988) for grammar, Keshavarz and Astaneh (2004) for vocabulary).
Summary: Laboratory-based research

One commonality of many of the laboratory-based studies reviewed above is that they sought to manipulate (e.g., Nation & McLaughlin, 1986; Lado, 2008; Lado et al., 2013) and control (Bartolotti & Marian, 2012; Bartolotti, et al., 2011; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a, b; Kaushanskaya & Rechtzigel, 2012; Stafford, et al., 2010) the conditions under which participants received additional language training. The results, however, are not wholly consistent and relatively difficult to directly compare. In three of the ten studies, effects of instructional context on L3 learning produced different outcomes regarding the potential benefits of bilingualism. One found beneficial effects of bilingualism under less explicit learning conditions (Lado, 2008; Lado et al., 2013), another found them in the more explicit condition (Nayak et al., 1990), and yet another found no clear beneficial effects of bilingualism in either more or less explicit conditions (Lin, 2009). Another study which compared early and late bilinguals on L3A under the same instructional context also failed to find differences, though balance in the two languages may have affected the outcomes (Stafford, et al., 2010).

Note only one of the studies on instructional context actually compared bilinguals and monolinguals and failed to find bilingual-monolingual differences (Nation & McLaughlin, 1986). However, given their use of an artificial grammar, the study’s generalizability to natural languages is reduced. The remaining studies (Bartolotti & Marian, 2012; Bartolotti, et al., 2011; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a, b; Kaushanskaya & Rechtzigel, 2012) indeed found advantages for bilinguals over monolinguals, but only tested vocabulary learning and processing, which provides a relatively narrow basis on which to make claims about a bilingual advantage in additional language learning. Finally, none of these studies followed monolinguals and bilinguals during additional language learning, from low to high proficiency,
and thus cannot provide information on the trajectory of language learning and potential bilingual-monolingual differences.

**Conclusions: Bilingual L3A compared to monolingual L2A**

The laboratory-based studies, like the studies conducted in real-world dual-or single-language environments, fall short of providing a reliable and clear answer as to whether bilinguals show advantages at additional language learning in adulthood, compared to monolinguals. However, the studies provide meaningful insights into how future research on bilingual L3A should proceed.

First, research should target the bilingual group most likely to show advantages, namely early (Cenoz & Valencia, 1994; Kaushanskaya & Marian, 2009a; Lasagabaster, 2000; Sanz, 2000) and balanced, or of particularly high (native-like) L2 proficiency (Bartolotti, et al., 2011; Lado, 2008; Lado, et al., 2013). Second, the experimental design should avoid potential confounds of language transfer that have weakened the impact of evidenced bilingual advantages (Lado, 2008; Lado, et al., 2013; Thomas, 1988), and provide control over amount, timing, and type of additional language exposure so that the internal variable of bilingualism can be separated from potentially intervening external variables associated with uncontrolled L3 input (Bartolotti & Marian, 2012; Bartolotti, et al., 2011; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a,b; Kaushanskaya & Rechtzigel, 2012; Lado, 2008; Lado, et al., 2013; Lin, 2009; Stafford, et al., 2010). Third, in order to provide a more detailed picture of the linguistic extent of hypothesized bilingual advantages, research should closely examine several domains of language, and not just one dimension such as general proficiency (Cenoz & Valencia, 1994; Sanz, 2000), pragmatics (Safont Jorda, 2003), or vocabulary (Bartolotti & Marian, 2012; Bartolotti, et al., 2011; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a,b; Kaushanskaya
& Rechtzigel, 2012). Finally, though no studies on bilingual L3A have attempted to do this, research needs to examine L3A compared to L2A across a developmental trajectory of learning, from low to high experience or proficiency, so that potential monolingual-bilingual differences can be explored as a function target language experience.

Thus far, bilingual L3A (compared to monolingual L2A) has been explored from an exclusively behavioral perspective and this behavioral evidence has proven difficult to reliably interpret in terms of potential bilingual advantages. Though there is suggestive evidence of a behavioral advantage for bilinguals over monolinguals at additional language learning, there is no evidence that bilinguals are different from monolinguals in terms of the neurocognition of late-learned language. Observed differences in behavioral outcomes are incapable of informing researchers about underlying mechanisms of bilingual L3A, how they may compare to monolingual L2A, and whether there might be neural correlates of bilingual advantages at additional language learning. The next section reviews several (neuro)cognitive theories of late-learned language, the consideration of which may help to guide the interpretation of potential advantages of bilingualism on additional language learning in adulthood.

**Neurocognitive theories of late-learned language**

The following section summarizes several neurocognitive theories and models of late-learned language processing which have been formulated in an attempt to explain the conditions under which late-learned language may differ from that of L1 speakers. Note that each of the theories and models reviewed below have been developed primarily to account for monolingual L2A, and it remains to be seen whether and how these neurocognitive approaches to L2A might be applied to bilingual L3A.
The declarative procedural model of language

The Declarative/procedural (DP) model is a dual-mechanism domain general model, that is, it proposes distinct cognitive and neural mechanisms for the separate linguistic components of grammar and lexicon, but does not assume that the mechanisms subserving these components are dedicated exclusively to that type of processing (i.e., linguistic). As a model of language processing, it is informed by independent knowledge about the functioning of the declarative and procedural memory systems (reviewed in Chapter 1), and makes specific predictions regarding the neural correlates of language processing within these systems.

The DP model posits that in the L1 and the L2, the declarative memory system is involved in the learning and processing of lexical/semantic information: the stored knowledge about words, their sounds, their meanings, and other semantic information (Ullman, 2001a, 2001b, 2004, 2005, 2013). In this model, the consolidation of new words relies on medial temporal lobe structures, although dependence eventually shifts to the neocortex (temporal and temporoparietal regions). Temporal lobe regions may be particularly relevant for the storage of word meanings whereas the storage of word sounds may be more strongly associated with temporoparietal regions. In terms of electrophysiology, the DP model makes specific predictions about which event-related potential (ERP) effects might be related to processing in one memory system or another. Within the model, it is proposed that N400 ERP responses reflect lexical/semantic processing in declarative memory for both L1 and L2 (Ullman, 2001b, 2004, 2005).

The DP model proposes that the procedural memory system is involved in the implicit learning and use of aspects of grammatical processing (rule-like computations) across syntax, non-lexical semantics, morphology, and phonology, and may be especially important for
grammatical (phrase) structure building. Learning is at least partially dependent on basal ganglia structures and learned knowledge may depend on temporo-parietal regions (Ullman, 2001a, 2001b, 2004, 2005, 2013). For language in neurologically normal adults, the model suggests that declarative memory can, in principal, be engaged for aspects of low-complexity grammatical structures (local dependencies); procedural memory is not required and individual differences (e.g., sex) may determine whether a learner relies more on their declarative chunking or on procedural computations for these grammatical structures (Ullman, 2005, 2007). More complex structures, such as long distance dependencies however, are likely to require procedural memory processing in L1 and be particularly unsuitable for declarative memory. With respect to electrophysiological signatures of language, the model proposes that (left) anterior negativities (LANs) reflect processing in procedural memory (2001b, 2004, 2005).

It is with regards to aspects of grammatical processing (such as phrase structure and morphosyntax) in an L2 that the model proposes differences between L1 speakers and L2 learners. At low levels of L2 proficiency (or experience), the model proposes that learners process lexical/semantic and grammatical information in declarative memory. This prediction is motivated by studies which suggest that declarative memory functioning improves through childhood and into early adulthood (DiGiulio, Seidenberg, O'Leary, & Raz, 1994; Graf, 1990; Vaidya, Huger, Howard, & Howard, 2007). Conversely, procedural memory abilities seem to function better in childhood, with little to no improvement (and possible attenuation) up to early adulthood. Thus, late L2 learners are expected to rely on their ‘good’ declarative memory more than procedural memory for the processing of L2. From an electrophysiological perspective, the DP model predicts that this reliance on declarative memory at low L2 proficiency should be reflected in N400s in response to grammatical violations (where such violations would be
expected to elicit AN-P600 responses in L1 speakers, see below in the P600 section of ERPs and language).

However, reliance on declarative memory for the processing of aspects of grammar in L2 is predicted to decrease as a function of L2 experience, or proficiency. Specifically, as L2 experience increases learners may come to rely less on declarative memory and more on procedural memory, which fits well with studies showing the increased practice or experience can facilitate processing in procedural memory (Schacter & Tulving, 1994; Squire, Knowlton, & Musen, 1993). In terms of electrophysiology, learners with high levels of L2 proficiency are expected to show a LAN-P600 response for grammatical violations. Note that the LAN would reflect processing in procedural memory, but the P600 is not associated with procedural memory in the model. The biphasic LAN-P600 response is predicted based on its occurrence in L1 speakers for certain grammatical violations, including phrase structure and morphosyntax (Friederici, Pfeifer, & Hahne, 1993; Osterhout & Mobley, 1995; Steinhauer & Connolly, 2008). Thus one component of the biphasic response is taken as evidence of procedural memory processing, whereas the entire biphasic response is considered evidence of L1-like processing of the grammatical violations.

In sum, the DP model posits that, at any level of L2 experience (or proficiency) lexical/semantic processing relies on the same memory system (declarative) as L1, and can be similar in L1 and L2. However, the similarity between L1 and L2 for processing of grammar depends, at least in part, on amount of L2 experience. At low levels of experience, aspects of grammar are processed differently in L2 learners (in declarative memory) compared with L1 speakers (in procedural memory). Nonetheless, with increased L2 experience, processing of aspects of grammar in L2 can approximate that of L1.
Declarative and procedural determinants of second language

Michel Paradis (Paradis, 1994, 2004, 2009) has also proposed a dual-mechanism model of L2 that involves declarative and procedural memory. While there are similarities between the DP model and that proposed by Paradis, there are also fundamental differences that have important consequences for predictions regarding L2 learning and processing. First, Paradis associates procedural memory with implicit knowledge and declarative memory with explicit (metalinguistic) knowledge, a distinction that the DP model does not make. Instead, the DP model admits that knowledge in declarative memory can be, but does not have to be, explicit (Morgan-Short, Finger, Grey, & Ullman, 2012; Ullman, 2013).

Second, whereas the DP model proposes a dissociation between lexicon and aspects of grammar in terms of the memory system underlying their processing, Paradis considers both aspects of the lexicon and of grammar to be related to processing in procedural memory (Paradis, 1994, 2004, 2009). However, his model does incorporate a distinction between vocabulary and the lexicon (Paradis, 2004, 2009). For Paradis, the lexicon refers to “a network of interrelated elements with their inherent morphosyntactic properties” (Paradis, 2009, p. 16). This network is implicit and subserved by procedural memory. In contrast, vocabulary refers to “the set of explicit word sound-meaning (and word grapheme-meaning) pairings, that is, those aspects that are consciously observable or teachable” (Paradis, 2009, p. 21). These sets of pairings represent explicit knowledge about words in the language; such knowledge is processed in declarative memory. This vocabulary depends, similarly to the DP model’s lexicon, upon temporal lobe-based structures. Paradis’ grammar and lexicon depends upon the frontal/basal ganglia-based system (note that in the DP model, grammar also depends on these structures).
With respect to L2 learning and processing, Paradis suggests that there is an optimal period for language, after which L2 learning becomes more difficult and/or less successful (Paradis, 2004, 2009). This optimal period involves implicit linguistic skill (i.e., processing in procedural memory), is time sensitive (shows age effects) and depends on properties of the brain (such as plasticity). However, this optimal period is not categorical, meaning that it is not an “all-or-nothing” approach to late-learned language. The notion of this period also accepts that variability among individuals (such as maturational or developmental stages) will likely result in varying outcomes across L2 learners. The optimal period, according to Paradis (2009), affects the acquisition of language and is a gradual process which takes place between two and five years of age, with five being the age at which gradual losses in plasticity of the procedural memory system begin. Given this loss of plasticity, Paradis proposes that individuals who begin learning a language after age 7 will rely more on declarative memory to process the L2. It is noted in Paradis (2009) that the predicted decline in use of the procedural memory system to learn language is not necessarily due to deficiency in procedural memory for language, but may also be related to other psychological factors including general learning ability and the presence of the L1 system. He also notes that “some implicit linguistic competence in L2 can probably be acquired in certain aspects of linguistic structure (syntax, morphology, phonology, in that order) though not completely at any level” (Paradis, 2009, p. 118).

By and large, however, Paradis posits that late learners of language are more likely to use speeded up control-like processing (in declarative memory) rather than automatic processing since the use of procedural memory for language processing is either no longer possible, or is extremely difficult and time-consuming. Efficient use of declarative memory can lead to high L2 proficiency, which is characterized by speeded-up control of L2 output. This use of declarative
memory is related to the pervasive inter-individual variability in terms of success in L2, specifically with respect to working memory, level of education, IQ, and motivation. However, if the L2 is used extensively, portions of it may eventually be automatized, resulting in implicit linguistic competence for those portions (i.e., processing in procedural memory). In sum, late learners of L2 as seen from Paradis’ model, may (albeit rarely) approximate L1-like performance, but will not process the whole of their L2 like L1 speakers, either for aspects of the lexicon or of grammar.

**The Shallow Structures Hypothesis**

The shallow structures hypothesis (SSH) was proposed in 2006 by Harald Clahsen and Claudia Felser (Clahsen & Felser, 2006a, 2006b, 2006c) in an attempt to account for the L1/L2 differences found in various psycholinguistic and neuroimaging studies of late-learned L2 processing compared to adult and child native-speaker processing. This hypothesis supports a dual-system view of language, and thus, like the DP model (Ullman, 2001a, 2001b, 2004, 2005) and Paradis’ account of late-learned language (Paradis, 1994, 2004, 2009), assumes that language consists of two divisible elements: a lexicon of structured entries and a computational system of combinatorial (rule-like) operations used to form complex utterances from lexical items (grammar). In their formulation of the SSH, Clahsen and Felser (2006a, 2006b, 2006c) distinguish between two different routes for sentence interpretation in L1: full parsing and shallow parsing. Full parsing refers to a fully specified syntactic representation whereas shallow parsing refers to a less detailed representation based on lexical/semantic information, associative patterns and other surface-level cues relevant for interpretation (Ferreira, Bailey, & Ferraro, 2002; Sanford & Sturt, 2002; Townsend & Bever, 2001).
The basic premise of the SSH is that full parsing is fed by the grammar and that, in L2, the grammar is not specified enough to provide the kinds of detailed syntactic information that are needed to process complex, hierarchical syntax in an L1-like way. Thus, L2 learners parse L2 sentences shallowly; relying on lexical, pragmatic, semantic and other information, but not on the complex, hierarchical syntactic information that L1 speakers would use. Note that shallow processing is not unique to the L2. For example, L1 speakers can also use shallow parsing (Ferreira, et al., 2002), but L2 learners, unlike L1 speakers, are restricted to it according to the SSH. Clahsen and Felser do not specify what neurocognitive mechanisms might underlie full or shallow parsing. Nor do they provide information on why L2 learners’ grammar might be underspecified, though they do reject the hypotheses that increased declarative memory reliance or decreased procedural memory availability could play a role (Paradis, 1994, 2004, 2009; Ullman, 2001a, 2001b, 2004, 2005). Note, however, that in terms of electrophysiology, Clahsen and Felser suggest that the SSH accounts for the absence of (left) anterior negativities in studies on sentence processing in L2, given that the this neural effect could be indexing automatic, full parsing mechanisms.

In terms of the lexicon, Clahsen and Felser do not predict L1/L2 differences in processing and suggest that with increased proficiency word-level processing can become L1-like. The SSH also suggests that morphological processing in L2 can become L1-like, especially for feature-matching (local agreement) and particularly at higher levels of proficiency. However, Clahsen and Felser stipulate that high L2 proficiency does not necessarily lead to native-like processing and that there is thus far no evidence that L2 learners, even at high proficiency, process complex hierarchical structures in a native-like way.

**The convergence hypothesis**
The convergence hypothesis (Green, 2003, 2005) has been largely motivated by studies on monolingual and bilingual aphasia and is a dual-system model of language, though not in the exact sense of the three neurocognitive models of late-learned L2 reviewed above (Clahsen & Felser, 2006a, 2006b, 2006c; Paradis, 1994, 2004, 2009; Ullman, 2001a, 2001b, 2004, 2005). The convergence hypothesis proposes that language can be divided into two devices: a lexico-semantic system which represents the meanings of words, their syntactic properties, and word forms, and a system of control which operates on the outputs from the lexico-semantic system.

Green’s hypothesis regarding L1 and L2 maintains that as proficiency in an L2 increases, the representation and processing profile of the L2 converge with those of L1 users. Specifically, the same ERP components are expected in L2 as are found in L1 speakers, and similar blood oxygen level-dependent (BOLD) activations are expected in L2 as in L1. The convergence hypothesis predicts that at high levels of L2 proficiency N400s should be elicited in response to lexical/semantic processing difficulty and P600s in response to illicit syntactic or morphosyntactic information (Green does not discuss the third important language-related ERP response, left anterior negativities). For BOLD responses, the convergence hypothesis predicts that lexico-semantic processing should show activations in the pre-motor cortex, supplementary motor areas, and cerebellum for word production and left anterior insula as well as left frontal operculum (part of Broca’s area) for phonological retrieval. At the sentence level, syntactic processing (also part of the lexico-semantic system) is predicted to involve left frontal areas, including part of Broca’s area and adjacent parts of the middle frontal gyrus. Semantic processing, on the other hand, is expected to involve temporo-parietal regions (left extrasylvian temporal and left anterior inferior frontal cortices) and potentially also the anterior temporal pole.
The control system implicated in the convergence hypothesis is proposed to involve neural control circuits that include both frontal attentional and subcortical structures. Green proposes that the language task schema (a mechanism for coordinating which language to use in the given context) may be mediated by subcortical structures, in particular the basal ganglia, and that the activation of the schema could be modulated by external information and also frontal structures, including the anterior cingulate cortex.

In sum, the convergence hypothesis predicts that potential qualitative differences in L1 and L2 representation and processing disappear as a function of L2 proficiency. Note, however, that this hypothesis does not stipulate that L2 learners will necessarily reach native-like levels of performance. Additionally, the hypothesis suggests that the context of acquisition (such as classroom or naturalistic) may affect the initial registration of L2 representations (Green, 2003). Finally, though the lexico-semantic system is not expected to differ between L1 and L2 at higher levels of L2 proficiency, the control system may indeed show increased activation in L2 compared to L1, particularly when control mechanisms are needed to resolve L1/L2 competition (Abutalebi & Green, 2008).

Green’s convergence hypothesis contrasts most clearly with the shallow structures hypothesis, which posits that L2 processing cannot reach L1-like profiles, even with high proficiency (Clahsen & Felser, 2006a, 2006b, 2006c) and also with Paradis’ view that the L2 can never become entirely L1-like (Paradis, 1994, 2004, 2009). The convergence hypothesis is most similar to the DP model (Ullman, 2001a, 2001b, 2004, 2005) in that both views consider increased L1-like processing to be a function of increased proficiency (or L2 experience) but differs from the DP model (as well as Paradis’ view and the SSH) in that the convergence
hypothesis relies on fundamentally different neurocognitive distinctions regarding the architecture of language (i.e., a lexico-semantic and control system).

The competition model

The Competition Model (Bates & MacWhinney, 1989; MacWhinney, 2001, 2005) is a single-mechanism, domain general model of language. The components of the model are not differentiated according to lexicon and grammar, and the mechanisms underlying processing are not dedicated to linguistic processing only but rather involve a distributed neuroanatomical network. The model was originally conceived of as a psycholinguistic model of L1, but was eventually applied to the domains of late-learned and also lost language (aphasia research). The original formulation of the Competition Model included the notion of competition, and three basic components: arenas, cues, and storage (MacWhinney, 2001, 2005). Competition is the core of the model and refers to a processing system that selects between various options (or cues) based on the relative cue strength of the options. Arenas are the linguistic domains within which competition occurs and include the lexicon, phonology, morphosyntax, and conceptualization. Cues refer to the notion of a linguistic sign as a mapping between form and function; Storage happens as a result of the learning of new mappings, and takes place for both short-term and long-term memory.

In an effort to apply the model to late-learned language, various components were added in order to provide a more unified account of language processing within the core notion of competition. Thus, Brian MacWhinney incorporated the components of chunking, codes, and resonance to the model (MacWhinney, 2005). Chunking refers to the storage and processing of linguistic units as whole entities instead of in their combinatorial parts. Codes refer to the language, or languages (in the case of switching and mixing), relevant to the processing moment
and include the notions of transfer (both positive and negative) as well as interaction (as in the switching mentioned above). The final component, resonance, is crucial in the competition model, especially as it applies to late-learned language, and refers to the repeated coactivation of reciprocal activations in the network. As the resonant connections grow, or in other words, as the language is repeatedly accessed, the possibility for cross-activation grows and the language begins to form a coherent co-activating neural circuit (MacWhinney, 2005).

The basic premise of the Competition Model is that language learning is a cue-driven process, which is a “process of acquiring coalitions of form-function meanings, and adjusting the weight of each mapping until it provides an optimal fit to the processing environment” (Bates & MacWhinney, 1989, p. 59). This cue-driven learning involves competition among linguistic cues that provide information about the interpretation of a sentence; cue selection depends on cue validity which, in turn, depends on cue strength. The classic application of the Competition Model to language employs the example of assignment of agent status to nouns in a sentence. For instance, some languages, such as English, rely on the preverbal position of the noun, making word order the strongest cue and other languages (with freer word order) might rely on case information as the strongest cue. The extensive use of L1 cues leads to their entrenchment. The model postulates that the strongest entrenchment occurs in phonology and the weakest in the lexicon.

For L2, especially in situations where there are competing L1-L2 cues, learners must overcome their L1 entrenchment in order to properly process cues in the L2. According to the Competition Model, this entails that they rely on resonant processes. The ability to rely on resonant processes may become less available with age due to general cognitive decline or commitments to L1 interactions which reduce the resonant state of the L2; making L1 intrusions
in L2 processing (i.e., negative transfer) more difficult to overcome. Thus, within this model, L1 and L2 processing involve the same architecture, but differences between L1 and L2 arise from differences in cue processing, which is rooted in L1 entrenchment. This entrenchment, however, can be overcome if sufficient resources are available to promote resonance in the L2 system. Thus, at a neurocognitive level, no-or limited - L1/L2 differences in processing profiles are expected (Hernandez, Li, & MacWhinney, 2005), especially when the L1 and L2 cues are not in competition, and even for lower levels of proficiency or L2 experience (Tokowicz & MacWhinney, 2005).

**Associative learning in L2**

Nick Ellis’ theory of late-learned language processing (N. C Ellis, 2006a, 2006b; 2006c) draws from work in psychology on associative and contingency learning. Contingency learning involves “gathering information about the relative frequencies of form-function mappings” (N. C. Ellis, 2006a, p. 1). This learning, and cognition more generally, is sensitive to two factors: recency and frequency. The relative recency of an item will affect its subsequent recall or recognition, as will the frequency with which that item has been previously encountered. Ellis suggests that language is tuned to input frequency in all linguistic domains, including phonology, reading, spelling, syntax, grammaticality, formulaic expressions, etc. However, frequency is not enough to promote learning, the system must also make associations between form and function of the input. In language, there are often multiple informative cues which co-occur and the learner’s task is to determine and select which cues are predictive, or reliable. Thus, “the driving forces of language learning… are frequency, conditioned by contingency, conditioned by selection” N. C. Ellis, 2006a, p. 15). According to Ellis’ approach, such mechanisms for learning apply for both L1 and L2, and across all domains of language.
However, associative learning does not only involve selection. It also involves selective attention, which is related to overshadowing and learned attention. Indeed, it is these aspects of associative learning that account for differences in L1 and L2 according to Ellis (N. C. Ellis, 2006a, 2006b, 2006c). Overshadowing occurs when two (or more) cues are equally frequent and equally reliable, but one is more salient and thus becomes associated with the outcome. This overshadowing eventually produces perceptual blocking of other, reliably informative cues. Selective attention, therefore, results from a single cue overshadowing, and subsequently blocking others. Extensive experience in attending to certain cues over others results in learned attention (to these specific cues), which affects how learners process new (L2) input. Thus, experience with the L1 attunes the processor to attend to certain cues at the expense of others; to describe this, Ellis also uses the term entrenchment (N. C. Ellis, 2006b).

Within a theory of associative learning of L2, the differences between L1 and L2 are proposed to stem from a failure to adopt and use the L2 form-function information that is available in the input. This failure-to-adopt, according to Ellis (N. C. Ellis, 2006a, 2006b) might be related to several factors: the available cues are unreliable predictors of the attested input, the cues are of low salience or are only tangentially important for the overall interpretative outcome, they are redundant, or they are ignored because of learned attention (L1 entrenchment). Ellis does not consider L1/L2 differences to be a product of neurobiologically determined age-effects, but rather a product of the limitations that the L1 imposes on the L2 in terms of associative learning mechanisms.

This associative learning theory of L2, unlike the Competition Model, does not specify whether or how L1 entrenchment can be overcome. The theory also does not consider what consequences overcoming learned attention to certain L1 cues might have on L2 processing,
making it difficult to determine whether, and on what levels, L2 processing can approximate that of L1 within an associative learning approach to late-learned language.

**Summary: Neurocognitive theories of late-learned language**

The six theories and models reviewed above, though fundamentally different in their approaches to L1 and L2 processing, all converge in suggesting that aspects of L2 processing can in principle approximate that of L1, especially with respect to word-level processing. The approaches crucially differ in the proposed extent to which L2 may come to be L1-like at the sentence level (i.e., syntax, morphosyntax). Several propose that L1-like processing can be attained when L2 experience or proficiency has reached high enough levels (Bates & MacWhinney, 1989; Green, 2003, 2005; MacWhinney, 2001, 2005; Ullman, 2001a, 2001b, 2005) and others claim it may never be entirely attained (Clahsen & Felser, 2006a, 2006b, 2006c; Paradis, 1994, 2004, 2009) even for high L2 proficiency. The approaches also differ in whether they assume separate systems for distinct domains of language (Clahsen & Felser, 2006a, 2006b, 2006c; Green, 2003, 2005; Paradis, 1994, 2004, 2009; Ullman, 2001a, 2001b, 2005) or a single, broadly distributed system which encompasses all linguistic domains (Bates & MacWhinney, 1989; N. C. Ellis, 2006a, 2006b; Green, 2003, 2005; MacWhinney, 2001, 2005).

One pertinent issue is how bilingual L3A and bilingual advantages at additional language learning in adulthood might fit within these neurocognitive approaches to late-learned language. Based on the extant research on bilingual L3A versus monolingual L2A which suggests that, using behavioral measures, bilinguals may indeed be better at monolinguals at learning an additional language, one question in need of attention is whether there are neurocognitive correlates of bilingual advantages in additional language learning and how such advantages could be explained within these neurocognitive approaches to late-learned language.
Of the neurocognitive approaches which suggest that L2 neurocognition can reach L1-like levels, each suggests a decisive role for L2 proficiency or experience (Bates & MacWhinney, 1989; Green, 2003, 2005; MacWhinney, 2001, 2005; Ullman, 2001a, 2001b, 2005). Thus, one possible manifestation of bilingual advantages at language learning in adulthood could pertain to reaching higher levels of proficiency than monolinguals and concomitantly also showing more L1-like neurocognitive profiles than monolinguals. However, considering the proposals which suggest that L1-like processing in the L2 is not attainable, it might be the case that, despite potential behavioral advantages, the neurocognition of bilingual L3A is no different from monolingual L2A and both are immutably different from L1 processing (Clahsen & Felser, 2006a, 2006b, 2006c; Paradis, 1994, 2004, 2009). One way to determine how the neurocognition of L3A compares to that of L2A, and how it fits into models of the neurocognition of late-learned language more generally, is to use the event-related potential technique. This methodology provides precise temporal information about language processing and offers an empirically sound measure of potential differences in L1 compared to L2 compared to L3 processing.

**Event-related potentials and language**

**Introduction**

This section is dedicated to providing an overview of ERPs in the study of language. First, a brief description of ERPs is provided. Second, descriptions of the three main language-related ERP responses relevant to the present study are offered. For each of the three ERP responses (N400, LAN, P600) the structure of the description is the following: first, an overview of the ERP effect, and particularly its elicitation in L1 research; second, theoretical views of
what type or types of cognitive processes the effect may index; and third, its elicitation in L2 research.

**Event-related potentials**

Event-related potentials refer to small changes in the electrical activity of the brain and can be recorded from electrodes placed on the scalp. These changes in electrical activity, or voltage, can be brought about by either internal or external events (Handy, 2005; Luck, 2005; Luck & Kappenman, 2012) and through the study of ERPs, researchers are able to investigate neurocognitive processes with excellent temporal precision (millisecond timing) and without necessary recourse to a (potentially) interfering secondary task (Kaan, 2007).

Event-related potentials reflect the synchronous build-up of post-synaptic potentials of large groups of neurons that are relatively close to the human skull and are parallel to each other. Such neurons, which are likely pyramidal cells in the neocortex (Kaan, 2007; Kappenman & Luck, 2012), contribute the most activity measured via the electrodes placed on the scalp (Handy, 2005; Kappenman & Luck, 2012; Luck, 2005). Because the ERP signal itself is much smaller than the amplitude in the spontaneous electroencephalogram (EEG), the signal must be amplified and averaged over a large number of trials. The resultant ERP waveforms contain both exogenous and endogenous potentials, where exogenous potentials refer to obligatory responses that are determined by the physical characteristics of the eliciting event (external) and endogenous refers to information processing in the brain that may or may not be a consequence of the eliciting event (internal) (Picton, et al., 2000). Waveforms are typically characterized by

---

2 Note that I am aware of no studies which have examined bilingual L3 processing with respect to these three ERP effects.
their amplitude (wave peaks and troughs), polarity (positive/negative), latency (the time of maximal amplitude), duration (in milliseconds), morphology (wave shape), and spatial topography (on the scalp). They are generally labeled according to their polarity and peak latencies (e.g., the P300 is a positivity with a peak latency around 300ms following stimulus presentation), though ERPs can also be labeled according to the processes they are considered to index (Kaan, 2007; Kappenman & Luck, 2012; Luck, 2005). For example, the mismatch negativity (MMN) is elicited in response to an incoming deviant stimulus compared to an expected (correct) auditory stimulus.

Language research using ERPs typically employs violation paradigms in order to study the time-course of language processing in different populations (e.g. children, patients, adult additional language users, etc.), and across different languages and linguistic domains (syntax, morphosyntax, semantics, phonology, etc.). This violation paradigm is characterized by measuring (through statistical subtraction) the neural activations to correct/expected/standard stimuli compared to violation stimuli, where the violation may include a specific linguistic target (such as phrase structure or agreement) or a semantic anomaly (I drink my coffee with sugar and *tar; * marks the violation; (Friederici, 2004; Kaan, 2007; Kotz, 2009). Using this violation paradigm, the field of language processing ERP research has revealed a set a widely studied neural activation patterns\(^3\) (referred to here as effects\(^4\): the N400, (left) anterior negativities (LANs), and the P600. Descriptions of the these components, the neurocognitive theories underlying them, and their emergence in L1 and L2 processing are reviewed below.

\(^3\) The three language-related effects reviewed here are not the only language-related ERP patterns studied using the EEG technique, for more information consult Stemmer and Connoly (2011) and Luck & Kappenman (2012).

\(^4\) The term component refers to a deflection in the waveform elicited by a certain type of stimuli (i.e., content words) whereas the term effect refers to the differences between waveforms elicited by different experimental conditions (correct versus violation, for example) (Kaan, 2007; Molinaro, Barber, & Carreiras, 2011).
Since its discovery in a study on visual sentence processing by Kutas and Hillyard (Kutas & Hillyard, 1980) the N400 component has generally been taken as an index of lexical/semantic processing in the brain (Friederici, 2002, 2004; Kaan, 2007; Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008). The N400 is a monophasic negativity with a broad centro-parietal scalp distribution appearing 200-600ms post-stimulus onset (maximal at 400ms) for written, visual, and signed stimuli. It has been found in both monolinguals and bilinguals, and across more than a dozen languages (Kaan, 2007; Kutas & Federmeier, 2011). Though it does not seem to vary in timing or topography as a function of language characteristics or writing systems, it is sensitive to language proficiency (Kutas & Federmeier, 2011; McLaughlin, Osterhout, & Kim, 2004). The N400 has been most actively studied in language research, though is not an exclusively language-related component. For example, N400 effects have been elicited for non-language stimuli such as sounds, faces, and pictures (Ganis, Kutas, & Sereno, 1996; Olivares, Iglesias, & Bobes, 1994; Van Petten & Rheinfelder, 1995), with the common element of all such stimuli being that they constitute meaningful items.

While the latency of the N400 appears to be relatively stable across various experimental manipulations and stimuli (but see van den Brink & Hagoort, 2004) for the auditory latency shift), differences in both amplitude and scalp distribution are commonly found. The variation in amplitude is often elicited in semantic anomaly and semantic priming (e.g., table-chair, jaguar-chair) manipulations, with less related or more anomalous items eliciting larger amplitudes (commonly referred to as the 'N400 effect'). The N400's amplitude has also been found to vary depending on the pragmatic anomaly of a word out of context (van Berkam, 2009) and other factors related to discourse or contextual fit, item frequency, predictability and also interpretative
relevance of the target (Choudhary, Schlesewsky, Roehm, & Bornkessel-Schlesewsky, 2009; Friederici, 2004; Kutas & Federmeier, 2011; Lau, et al., 2008). Differences in scalp distribution can emerge depending on the modality of the stimulus presentation, with a slight right hemisphere bias for visually-presented stimuli and occasionally a slight left hemisphere bias for auditorily-presented stimuli (Friederici, 2004; Kutas & Federmeier, 2011; Lau, et al., 2008). Such variations have led researchers to conclude that although the N400 is not dependent on the modality of the stimulus, it is sensitive to it (Kutas & Federmeier, 2011). Scalp topography of the N400 is also affected by the nature of the stimuli. The N400 in response to familiar faces with mismatch features, for example, has an occipital maximum (Olivares, et al., 1994) whereas that for meaningful environmental sounds is left hemisphere dominant (Van Petten & Rheinfelder, 1995).

Though ERP research itself is not yet advanced enough to reliably localize the neural sources of N400 activity, other neuroimaging techniques have been able to provide insights into the brain areas associated with the same types of processing that elicit N400s in ERP research. Studies using functional magnetic resonance imaging (fMRI) and the semantic priming paradigm have found activation in the left inferior frontal gyrus (IFG), the left medial temporal gyrus (MTG), and superior temporal gyrus (STG) (Devlin, Jamison, Matthews, & Gonnerman, 2004; Frisch, Kotz, von Cramon, & Friederici, 2003; Giesbrecht, Camblin, & Swaab, 2004; Gold, et al., 2006; Matsumoto, Lidaka, Haneda, Okada, & Sadato, 2005; Rissman, Eliassen, & Blumstein, 2003; Wible, et al., 2006). Using the semantic anomaly paradigm, similar regions have also shown activation, though the results are less consistent (Friederici, Ruschemeyer, Hahne, & Fiebach, 2003; Kiehl, Laurens, & Liddle, 2002; Kuperberg, et al., 2003; Ni, 2000; Ruschemeyer, Fiebach, Kempe, & Friederici, 2005). Magnetoencephalography (MEG) research has implicated
the left mid-posterior MTG, the STG, and superior temporal sulcus (STS) in processing during semantic priming and semantic anomaly experiments, with an onset of 250ms for auditory stimuli (Helenius, et al., 2002; Uusviuori, Parviainen, Inkinen, & Salmelin, 2008) and 300-350ms for visual stimuli (Dale, et al., 2000; Maess, Herrmann, Hahne, Nakamura, & Friederici, 2006). Other localized areas, which comprise a later response to the anomaly, include the left anterior temporal and inferior frontal cortex as well as the right orbital and anterior temporal cortex (Lau, et al., 2008). An event-related optical signal study reinforced the implication of these areas (and their timing) in the generation of the N400. Specifically, the researchers found left hemisphere responses in the mid-posterior STS and MTG at 200-400ms post-stimulus onset, with involvement of anterior temporal and inferior frontal areas later, after 500ms (Tse, et al., 2007). The use of intracranial recordings has shown that activity in the anterior medial temporal lobe correlates with the N400 response, indicating that this area may be an additional generator (McCarthy, Nobre, Bentin, & Spencer, 1995; Nobre & McCarthy, 1995). Overall, these brain areas correspond well with the distributed network of semantic memory storage and online integration during language comprehension, and thus point to the import of the N400 response in the processing of semantic, meaningful information (Kutas & Federmeier, 2011; Lau, et al., 2008). However, despite the general consensus that N400s index lexical/semantic processing, the specific interpretations of this component are still a matter of debate.

Perhaps the most widely accepted theoretical account of the N400 is that it broadly reflects the integration of lexical/semantic information. This view has been supported since the discovery of the N400 (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980) and imaging data corroborate more specified accounts which attempt to link the N400 response with particular brain regions active during various stages of semantic processing such as automatic lexical
access, controlled retrieval of lexical information, and semantic integration with context (Lau, et al., 2008). However, the specific "stage" at which the N400 comes into language processing play is not yet clear. Some researchers associate N400 activity with semantic unification, or the binding of lexical information into a representation of multi-word utterances, and also meaning from non-linguistic modalities such as gesture. This is considered to be a late, or post-lexical account of the N400 (Brown & Hagoort, 1993; Hagoort, Baggio, & Willems, 2009).

Contrastively, researchers have also proposed that N400 activity reflects very early stages of processing, such as orthographical or phonological processing (Deacon, Dynowska, Ritter, & Grose-Fifer, 2004).

Kutas and Federmeier (2011) reviewed in detail the reasons for which the aforementioned views are too interpretatively narrow and cannot account for the variety of conditions under which N400 effects are elicited. For example, they point out that the post-lexical view of the N400 cannot account for effects that are elicited with pseudowords (i.e., items for which no representation exists and thus cannot be semantically unified, see (Laszlo & Federmeier, 2009) and the orthographical/phonological view cannot account for contextual or discourse-related N400 effects (van Berkam, 2009). Thus, the authors update their 'semantic memory' view of the N400 to suggest that it does not reflect the activation of lexical representations per se, but rather reflects the activity of a multimodal long-term memory system which is modulated by incoming stimuli as meaning is dynamically constructed over time. This synchronization between stimulus-driven activity and the multimodal long-term memory system is associated with a binding (occurring during the N400) into a multimodal conceptual representation which, rather than being stored and subsequently accessed, is created as the message unfolds in time and is highly context-dependent (Laszlo & Federmeier, 2010). This binding is both implicit and
Transient, even when explicit memory systems are compromised (amnesics) and is likely not directly responsible for the conscious detection of meaning (Kutas & Federmeier, 2011).

An alternative N400 view, pioneered by Bruno Debruille and colleagues (Debruille, 1998; Debruille, 2007; Debruille, Pineda, & Renault, 1996; Debruille, et al., 2008), suggests that generation of the N400 is not a function of any type of integration processes, but instead indexes inhibition. The inhibition hypothesis proposes that N400s arises when the knowledge activated during word priming or sentence processing is not verified and must be inhibited. For example, in the sentence “He was eating...”, knowledge related to food is activated and creates expectations about incoming material. If the next word is “books”, which does not match the established expectations, the food knowledge must be inhibited (Debruille, 2007). The resultant N400 is a function of the amount of knowledge that has to be inhibited; the more knowledge that has to be inhibited, the larger the N400 response will be. Accordingly, modulations in N400 amplitude are accounted for. Note, however, that only four studies have actively investigated this hypothesis (Debruille, 1998; Debruille, 2007; Debruille, et al., 1996; Debruille, et al., 2008) and it remains theoretically underspecified in terms of its ability to account for the various contexts in which N400 effects are elicited (i.e., high versus low cloze probability (Kutas & Federmeier, 2011).

While most research has found N400s in response to unarguably lexical/semantic material (see above), recent studies have also found N400 activity for stimuli that are not so obviously lexical/semantic (e.g., more rule-based or rule-like). Frisch and Schlesewsky (2001) for example, found N400s in response to case marking violations in German, where case-marking is traditionally viewed as being morphosyntactic. Additionally, Bornkessel, McElree, Schlewesky, and Friederici (2004) found an N400 for word order reanalysis in subject-object
ambiguities that were related to case-marking in German. Bartke, Rösler, Streb, and Wiese (2005) also found N400 effects for violations of German number marking which involved forms that were not completely irregular, but were also not the default plural form.

In response to these findings, some researchers have hypothesized that the N400 may involve more rule-like computation than previously believed, in particular for non-default (Bartke, et al., 2005) or interpretively relevant (Choudhary, et al., 2009) rule-based information. The latter view is supported by the finding of N400 activity for violations of Hindi case-marking, which constrains the range of possible interpretations of the subject. The authors concluded that their findings, in combination with the German case-marking N400 results (Frisch & Schlesewsky, 2001, 2005), provide evidence that the processing of rule-based linguistic information correlates with an N400 when the violation of a rule application is interpretative, as opposed to formal, in nature (Choudhary, et al., 2009). These findings somewhat compromise considering the N400 a marker of lexical/semantic processing that cannot be related to "anomalies in language that are non-semantic in nature such as grammatical violations" (Kutas & Federmeier, 2000), p. 464). Instead such findings suggest that the N400 may be more dynamically associated with rule-like procedures than has been previously believed, specifically when such procedures bear on the meaning of the input.

Regarding the elicitation of N400s in L2 research, N400s are often elicited in response to lexical/semantic anomalies, though occasionally with reduced amplitudes or later onsets (McLaughlin, et al., 2004; Steinhauer, White, & Drury, 2009; Ullman, 2001b). Ojima, Nakata, & Kakigi (2005) for example, found N400s in response to semantic anomalies in English in both high and low proficiency L2 learners, although it was delayed compared to native English speakers. Weber-Fox & Neville (1996) also found N400s in response to semantic anomalies in
their L2 English learners, and this response was independent of age of L2 acquisition (note that proficiency was not controlled for in this study). N400s were also reported for Japanese L1-German L2 speakers with less than 3 years of German instruction who were living in Germany (Hahne & Friederici, 2001) as well as for English L1-French L2 learners following just one month of formal instruction (McLaughlin, et al., 2004).

Thus, N400s, and perhaps L1-like lexical/semantic processing more generally, appear to be common in L2 processing, even at low levels of proficiency or L2 exposure.

**N400 Summary**

The centro-parietal negativity discussed above is considered to reflect aspects of lexical/semantic processing in L1, or the integration of meaning on a more general scale (Kutas & Federmeier, 2011). Other theories suggest that lexical/semantics may not be all that the N400 reflects and propose a larger role for rule-like computations (Choudhary, et al., 2009), or even that N400s index semantic inhibition (Debruille, 2007) as opposed to integration. Nonetheless, the brain regions that have been proposed as potential neural sources of N400 activity fit well with the processing of lexical/semantic information (Lau, et al., 2008). Finally, in L2, N400s have been found for lexical/semantic processing, even at low L2 proficiency or after minimal formal training (McLaughlin, et al., 2004; Ojima, et al., 2005). These N400s are similar to those found in L1, though perhaps with timing or amplitude differences (Ojima, et al., 2005), which suggests that lexical/semantic procedures in L2 are likely not qualitatively different than those which take place in L1.

*(Left) Anterior Negativities*
Anterior negativities (ANs) generally overlap in time with N400s, occurring around 300-500ms post-stimulus, and are distinguished from N400 activity by their scalp topography as well as the types of violations that elicit them (Friederici, 2002; Molinaro, et al., 2011; Osterhout & Mobley, 1995). Specifically, ANs are generally found in more frontal electrodes (anterior distribution) and are often left-lateralized. Note that many researchers therefore refer to left anterior negativities (Friederici, 2002; Kotz, 2009; Molinaro, et al., 2011; Morgan-Short & Ullman, 2011) when describing this component. Though the frontal distribution of ANs is typically their defining characteristic, they can also extend to more central sites and despite their tendency to be left–lateralized, ANs with more bilateral distributions have also been reported (Pakulak & Neville, 2010). The central extension, and perhaps also the bilateral distribution of ANs, may be related to lower levels of proficiency in the L1 (Pakulak & Neville, 2010).

Some researchers have proposed that ANs can be separated into an early component (eLAN, 100-200ms) and one which occurs later (LAN, 300-500ms; (Bornkessel & Schlesewsky, 2006; Friederici, 2002; Friederici, Hahne, & von Cramon, 1998). This eLAN has been found in response to phrase structure violations (Friederici, et al., 1993; Friederici, Steinhauer, & Frisch, 1999; Neville, Nicol, Barss, Forster, & Garrett, 1991), though it is less commonly reported for these contexts than (L)ANs (Friederici, Hahne, & Mecklinger, 1996; Hagoort, Wassenaar, & Brown, 2003; Newman, Ullman, Pancheva, Waligura, & Neville, 2007).

LANs are also elicited in response to agreement violations. Specifically, LANs have been found in response to violations of number agreement in the verbal domain (Kaan, 2002; Molinaro, Vespignani, & Job, 2008; Osterhout & Mobley, 1995; Silva-Pereyra & Carreiras, 2007) as well as the nominal domain for determiner-noun (Barber & Carreiras, 2005; Hagoort,
Based on the timing and violation contexts of the eLAN and LAN, it has been proposed that eLANs index early, automatic structure building during online sentence processing (“first pass parsing”) and LANs reflect the assignment of grammatical functions or relations to words and their thematic information (Friederici & Weissenborn, 2007; Friederici, 2002, 2004; Friederici & Kotz, 2003) or, more broadly, LANs reflect early detection of violations of morphosyntactic information which encodes such information (Molinaro, et al., 2011). However, other researchers consider the eLAN\textsuperscript{5} to be an index of failure to select an appropriate syntactic (verb argument) template and the LAN to reflect difficulties in computing agreement or syntactic hierarchical structure. In their extended argument dependency model (eADM), Bornkessel & Schleswksky (2006) suggest that LANs are found during a specific stage in sentence processing, specifically when *Compute Linking* fails. Compute linking in their model refers to the computation of agreement regarding the thematic roles of the nouns in a sentence. Linking occurs when the (verb) argument hierarchy of a given verb is associated with the arguments that have been processed and LANs are elicited when no agreement relation can be successfully computed.

Regarding the neural sources of the LAN, research on aphasic patients has implicated the anterior portion of the left hemisphere (Friederici, et al., 1998) and also the basal ganglia (Friederici, von Cramon, & Kotz, 1999). LANs have been posited to rely on the procedural memory system (Ullman, 2001a, 2001b) which is subserved by frontal-basal ganglia structures,

\textsuperscript{5} Although it seems as if there may be an early effect that may be dissociated from the LAN, the (L)AN will be the focus of consideration in this proposal, not least because it has been attested in a larger body of literature and is, to some extent, better understood. For a detailed treatment of eLAN theory and research, see Steinhauer and Drury (K. Steinhauer & Drury, 2011).
but little else is known about possible neural generators of LAN effects (see also Frisch, et al., 2003).

Another important point to note about LAN theories is that, in some L1 studies, an anterior negativity is found in later time-windows (600-2000ms) than the traditional LAN (Friederici, et al., 1993; Martin-Loeches, Munoz, Casado, Melcon, & Fernandez-Frias, 2005; Mueller, Hahne, Fujii, & Friederici, 2005). These later anterior negativities are often bilaterally distributed and some researchers suggest that they may be related to increased working memory demands (Fiebach, Schlesewsky, & Friederici, 2002; Martin-Loeches, et al., 2005). However, other researchers have suggested that it is not a separate component, but rather a continuation of the earlier LAN (Steinhauer & Drury, 2011). More research, however, is needed in order to determine whether these later anterior negativities and earlier LANs are, in fact, distinct neural responses or are instead a single response with a longer duration than previously thought.

In L2, LANs have been found for subject-verb agreement and phrase structure violations, including for artificial languages (Friederici, Steinhauer, & Pfeifer, 2002; Morgan-Short, Steinhauer, et al., 2012). However, unlike N400s, LANs have been found only at higher levels of L2 proficiency. Rossi, Gugler, Friederici, and Hahne (2006), for example, examined low and high proficiency late L2 German and L2 Italian learners on their processing of three types of violations: phrase structure (category), subject-verb agreement, or combined. For both groups of L2 learners, the high proficiency subjects evidenced LANs in response to all three types of violations. There were no LANs in the low proficiency group. Ojima et al. (2005) also tested low and high proficiency L2 learners on subject-verb agreement violations. The high proficiency learners showed a LAN similar to that found in the native English-speaker comparison group, though no LANs were found for the low proficiency learners. Other studies, however, have
failed to find LANs in L2 learners at high proficiency for phrase structure (Mueller, et al., 2005) or subject-verb agreement violations (Hahne, Mueller, & Clahsen, 2006). Instead, learners may show N400s or N400-like effects (Osterhout, et al., 2008; Tanner, McLaughlin, Herschensohn, & Osterhout, 2012; Weber-Fox & Neville, 2001).

Though LANs are generally found in response to nominal number and gender agreement violations in L1, most studies have failed to find them in the L2, either for noun-determiner or noun-adjective constructions (Gillon Dowens, et al., 2011; Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008; Tokowicz & MacWhinney, 2005). However, one L2 study which did find LANs in response to such nominal agreement violations, Gillon Dowens et al. (2009), tested high L2 proficiency learners who had been living in the target language environment for an extended period of time (range 12-33 years, average 22.1), suggesting an influential role for type of exposure. Nonetheless, there were slight differences in the onsets and scalp topographies of the LANs for the number and gender conditions; these differences were not found in the native speaker control group. Specifically, the LANs for the L2 group were more bilaterally distributed, extended into central-posterior electrodes, and occurred slightly later. The authors suggested these differences may have been related to effects of age of L2 acquisition, but conclude that the attainment of L1-like morphosyntactic processing depends on other important variables as well, such as L2 proficiency.

LAN Summary

In sum, anterior negativities which occur relatively early during sentence processing (300-500ms) are considered to reflect processing associated with automatic structure-building and have been found for phrase structure, verbal agreement, as well as nominal agreement in L1. These ANs are generally left-lateralized and may depend on sub-cortical structures such as the
basal ganglia (Friederici, 2004; Friederici, et al., 1998), and be related to processing in the procedural memory system (Ullman, 2001a). Additionally, anterior negativities which are more bilateral and extend to central electrode sites may occur later in processing (600+ ms), though whether these are functionally distinct from the earlier LANs has yet to be resolved. Finally, LANs have been found in L2 processing, though only at higher levels of L2 proficiency and more often for phrase structure or verb agreement violations than for violations in the nominal domain (Gillon Dowens, et al., 2009; Ojima, et al., 2005; Rossi, et al., 2006). Thus, L1-like morphosyntactic processing is possible in L2, though the attainment of such processing appears to be a function of proficiency as well as linguistic target.

**P600s**

The P600 is a monophasic positive-going waveform with a posterior scalp distribution which occurs 500-900ms post-stimulus onset (Kaan, Harris, Gibson, & Holcomb, 2000; Osterhout & Mobley, 1995). Though P600s are often found over posterior electrodes, they have also been observed as demonstrating a more anterior distribution (Friederici, Hahne, & Saddy, 2002; Hagoort, Brown, & Osterhout, 1999; Kaan & Swaab, 2003). This positivity is, like the LAN described above, elicited in response to morphosyntactic violations, including phrase structure (Friederici & Mecklinger, 1996; Neville, et al., 1991), subject-verb agreement (Coulson, King, & Kutas, 1998; Kaan, 2002; Osterhout & Mobley, 1995; Roehm, Bornkessel, Haider, & Schlesewsky, 2005; Silva-Pereyra & Carreiras, 2007), and nominal (gender/number) agreement (Barber & Carreiras, 2005; Hagoort, 2003; Martin-Loeches, et al., 2006).

P600s are also elicited in response to garden-path sentences, that is, sentences in which the target word is unexpected given the preferred reading of a sentence (“The broker persuaded to sell the stock was sent to jail”, from Osterhout & Holcomb, 1992). In morphosyntactic
processing, P600s have been found co-occur with LANs and this biphasic LAN-P600 response may be a signature of L1-like morphosyntactic processing (Friederici, et al., 1993; Osterhout & Mobley, 1995; Steinhauer & Connolly, 2008). The amplitude of P600s has been found to vary according to sentence complexity (Gunter, Stowe, & Mulder, 1997; Kaan, et al., 2000; Munte, Szentkuti, Wieringa, Matzke, & Johannes, 1997) and may also be modulated by task conditions. For example, one study found that P600 amplitudes were larger when participants performed grammaticality judgments than when they passively read the sentences (Osterhout, Holcomb, & Swinney, 1994).

The neural sources underlying the generation of the P600 response are, like the LAN, not well understood. One study on patients whose brain lesion included the basal ganglia (n = 7) and those whose lesion site did not (n = 7) found that only those with intact basal ganglia structures showed P600s in response to morphosyntactic violations in German (Frisch, et al., 2003; see also Friederici & Kotz, 2003). In an fMRI study on patients with Parkinson’s disease (for whom the neural degeneration involves the substantia nigra, which has connections to and is often associated with the basal ganglia; Felten & Shetty, 2010; Kleiner-Fisman, et al., 2006; MacDonald, et al., 2011), the left superior temporal gyrus and deep frontal operculum were implicated in syntactic processing (specifically local phrase structure building). In one MEG study on healthy subjects (Kwon, et al., 2005), the middle temporal gyrus was identified as neural generator of the P600 effect. Thus, the evidence suggests that the regions of the temporal lobe are likely sources of P600 effects, with a potential role for the basal ganglia as well.

Given an interpretation that LANs reflect early automatic structural processing, some researchers have proposed that P600s reflect structural reanalysis. In Angela Friederici’s three-stage model of sentence processing (Friederici & Weissenborn, 2007; Friederici, 2002), P600s
are part of the third stage, during which semantic and syntactic streams of information are integrated. If this integration is disrupted, additional controlled syntactic reanalysis and repair processes are needed, resulting in a P600. Another proposal has been that P600s may not index reanalysis, but rather the cost of reprocessing a syntactic violation. Osterhout et al. (1994) suggested this in light of their findings that P600s have the largest amplitude for grammatical violations, smaller amplitudes for grammatically well-formed but less preferred structures, then smallest for preferred structures. Edith Kaan and colleagues (Kaan, et al., 2000), however, propose that P600s are not indexing reanalysis or reprocessing cost per se, but instead index syntactic integration difficulty in general. This theory is supported by studies which have found P600s in response non-language contexts which also involved integration difficulties, for example in processing violations of musical structure (Patel, Gibson, Ratner, Besson, & Holcomb, 1997) and mathematical rules (Lelekov, Dominey, & Garcia-Larrea, 2000).

Recently, several studies have also reported P600s in sentences that were neither structurally illicit or less preferred, but rather semantically anomalous. Kim & Osterhout (2005), for example, found a P600 in response to the sentence “The hearty meal was devouring…” (see also Hoeks, Stowe, & Doedens, 2004; Kuperberg, et al., 2003). These “semantic P600 effects” are still not well understood in the traditional context of P600s as syntax-related components. Gina Kuperberg has suggested that these semantic P600s provide evidence against syntax-driven accounts of sentence processing (Kuperberg, 2007). Specifically, she proposes that the language processor is composed of two processing streams: a semantic memory-based stream and a combinatorial stream which is engaged in both morphosyntactic and semantic-thematic processing. When the output of these two streams conflicts, P600s are found. However, Bornkessel-Schlesewsky & Schlesewsky (2008) argue that Kuperberg’s account does not
consider all the contexts in which semantic P600 effects have been found and propose an alternative account. Within their eADM (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2008) semantic P600s are posited to stem from problems in the final stages of processing, namely during Generalised Mapping which requires successful integration of Compute Linking and Plausibility steps. Thus, for a sentence-beginning such as “The hearty meal was devouring…”, plausibility is not satisfied and cannot be successfully integrated with linking at the mapping stage, leading to a P600 effect. The two theories, however, have yet to be teased apart regarding semantic P600 effects and how their elicitation may impact the syntax-centered view of P600s in general.

P600s have been found in L2 learners in response to morphosyntactic violations, though emergence of this effect, and the timing of it, may depend amount of experience in the L2. For example, Weber-Fox and Neville (1996) found P600s in response to phrase structure violations to vary as a function of age of acquisition in the Chinese L1-English L2 learners. Specifically, they report L1-like P600s in L2 learners with early ages of acquisition (1-10 years old), delayed P600s in learners with slightly later ages of acquisition (11-13 years old), and no P600s in learners who acquired the L2 after the age of 16.

Subsequent studies have found that late L2 learners in fact often show P600s, though they may vary as a function of L2 proficiency. Rossi et al. (2006), for example, reported P600s in their late high proficiency L2 learners for both subject-verb agreement and phrase structure. Their low proficiency L2 learners also showed P600 effects, though they were delayed in both conditions. At intermediate L2 proficiency, Hahne (2001) also found delayed P600s for phrase structure (Russian L1-German L2). Tanner et al., (2012) found P600s in response to subject-verb agreement violations in German by native English speakers with three years of classroom
experience but did not find the same effect in learners with just one year of classroom experience, and found that wave amplitude positively correlated with ability to discriminate correct from violation structures. White, Genesee, and Steinhauser (2012), alternatively, found that after just nine weeks of intensive English as a Foreign Language instruction, both Korean and Chinese speakers showed P600s in response to English past-tense violations; in this study wave amplitude was also found to positively correlate with performance.

Ojima et al. (2005), in contrast, found no P600s in response to subject-verb agreement in their late L2 English learners, at either high or low proficiency, whereas Tokowicz and MacWhinney (2005) reported P600s for their low proficiency English L1-Spanish L2 learners; for verb tense marking and also nominal gender agreement (but not nominal number agreement). In low proficiency English L1-French L2 learners P600s were also missing in the nominal number agreement condition (Osterhout, McLaughlin, Pitkänen, Frenck-Mestre, & Molinaro, 2006) whereas another study on high proficiency L2 learners of German (Russian L1) showed P600s in response to such violations (Hahne, et al., 2006). Gillon-Dowens et al. (2009) reported P600s in high proficiency L2 learners (English L1-Spanish L2) for number and gender agreement, though the P600s were larger (in amplitude) for number than gender agreement, a difference that was not present in the native-speaker comparison group. Gillon-Dowens et al. (2011) also found P600s for number and gender agreement in late Chinese L1-Spanish L2 high proficiency learners, this time with no difference in P600 amplitude between the two agreement conditions.

Finally, in a study using a miniature language, Mueller et al. (2005) reported P600s for phrase structure as well as case violations in their high proficiency L2 learners and in another set of studies using an artificial language, Morgan-Short, Steinhauser, et al. (2012) and Morgan-Short
et al. (2010) found P600s for phrase structure in addition to noun-article gender agreement at high (but not low) proficiency. Thus, it appears that P600s, though perhaps more commonly found at higher levels of L2 proficiency, can be found at lower levels of L2 proficiency as well, though they may be delayed compared to L1. Recall that, as mentioned above, a biphasic (L)AN-P600 response for morphosyntactic violations may be a neural signature of L1-like processing. This biphasic response has been found in the L2, though in fewer studies than either (L)ANs or P600s alone, and only at high proficiency (Gillon Dowens, et al., 2009; Morgan-Short, Steinhauer, et al., 2012; Rossi, et al., 2006).

**P600 Summary**

To summarize, the P600 is a positivity with a predominantly posterior distribution elicited in response to morphosyntactic violations. Given its relatively late timing (500-900ms), it has been considered a marker of controlled processing during sentence comprehension, perhaps due to structural reanalysis or general syntactic integration difficulty. P600s may depend on regions of the temporal lobe, and also sub-cortical structures such as the basal ganglia. In L2, P600s have been reported, though more often at high L2 proficiency than low. In contrast, L1-like biphasic LAN-P600 responses have been reported less often for L2, and only in learners with high L2 proficiency. In sum, L2 research on P600 effects suggests that controlled aspects of processing of morphosyntax in L2, and L1-like processing signatures more generally (i.e., LAN-P600), can be attained. However, the emergence of this L1-like processing may be related to higher levels of L2 proficiency experience.
Conclusions: ERPs in late-learned language processing

The extant research on late-learned language processing as measured using ERPs suggests the following. First, insofar as N400s index lexical/semantic processing during sentence comprehension it appears as though such processing is not different between L1 and late L2 learners. However, the exact characteristics of the N400 effect may differ. Second, automatic morphosyntactic processing as indexed by (L)ANs has been evidenced in late L2 learners, though only at high L2 proficiency and more often for certain structures (phrase structure) than others (nominal agreement). Like the N400 effect, these ANs may be characteristically different from native-speakers’ LANs, in terms of scalp distribution and potentially also timing. Finally, more controlled aspects of morphosyntactic processing as measured by P600s have been found in L2, though again such native-like processing appears more often at higher than lower levels of L2 proficiency or experience. Additionally, the characteristics of the P600 may differ from native speakers. Thus, it appears as though L1-like processing of late-learned language is indeed attainable. There exist, however, differences in L1 and L2 processing signatures in terms of timing, amplitude, and topography that may be related to levels of proficiency or experience in the late-learned language.

Rationale for the current study

The research which has compared bilingual L3A to monolingual L2A suggests that bilinguals may have an advantage over monolinguals at learning an additional language in adulthood, at least for performance-based outcomes. However, limitations of the existing work in terms of experimental design and the measures used to assess L3A preclude making strong conclusions regarding the existence of such an advantage. Additionally, the work has left unanswered several important questions, specifically regarding the mechanisms underlying
bilingual L3A compared to monolingual L2A and also to a certain extent the effects of instructional context on the emergence of bilingual advantages at additional language learning.

The neurocognitive theories reviewed above differ in their accounts of whether and to what extent L1-like processing can be attained by late language learners. However, these neurocognitive approaches are framed in terms of monolingual L2A and it would be worthwhile to explore their claims within bilingual L3A research. This is especially relevant given the supposition that bilinguals may be better than monolinguals at late language learning and therefore research on bilingual L3A from a neurocognitive perspective could significantly inform both research and theory on late-learned language in adults.

The neurocognition of bilingual L3A compared to monolingual L2A has yet to be investigated. Using ERPs one can reliably study potential differences in L3A and L2A regarding the neural processing signatures of language, specifically with respect to three well-studied activation patterns. Work in L2 processing suggests that L1-like processing of lexical/semantics and morphosyntax is possible, but is likely a function of language proficiency (especially for the latter). However, even at higher levels of proficiency or experience, ERP effects in late L2 learners may be characteristically different from L1 speakers. Thus, increased L1-like processing in bilinguals as compared to monolinguals could be a neurocognitive manifestation of bilingual advantages at additional language learning.

In brief, the study sought to investigate bilingual L3A under two different instructional conditions and at two points during additional language learning (low and high experience). In addition to collecting behavioral measures of additional language performance (comprehension and production accuracy and discrimination performance on grammaticality judgment tasks), the study measured ERP responses to correct and violation target language structures. The study
compared the acquired data to existing monolingual L2A data that was collected under the same conditions as those in the current project. The acquisition of both behavioral and electrophysiological data and the existence of a comparable monolingual control group enabled the study to better characterize the breadth and depth of potential bilingual advantages at additional language learning and it aimed to answer the following research questions.

**Research Questions**

**Research Question 1**
When exposed to additional language in a traditional classroom (Instructed) context, is performance in judging correct and violation structures different in bilinguals compared to monolinguals at low and/or high experience?

**Research Question 2**
When exposed to additional language in a traditional classroom (Instructed) context, are there differences in neurocognitive processing, as indexed by ERPs, in bilinguals compared to monolinguals at low and/or high experience?

**Research Question 3**
When exposed to additional language in a communicative classroom or more naturalistic (Uninstructed) context, is performance in judging correct and violation structures different in bilinguals compared to monolinguals at low and/or high experience?

**Research Question 4**
When exposed to additional language in a communicative classroom or more naturalistic (Uninstructed) context, are there differences in neurocognitive processing, as indexed by ERPs, in bilinguals compared to monolinguals at low and/or high experience?
Predictions

Behavioral

Regarding Research Questions 1 and 3, the majority of bilingual L3A research has found that bilinguals do indeed outperform monolinguals on behavioral measures of learning (e.g., (Cenoz & Valencia, 1994; Kaushanskaya & Marian, 2009; Safont Jorda, 2003; Sanz, 2000; Thomas, 1988). Thus, it was predicted that bilinguals would outperform monolinguals in judging correct and violation target language structures. It was predicted that this performance advantage would be evident at high experience, and possibly even low. No predictions were made regarding the extent to which bilingual/monolingual differences would depend on instructional context for the emergence of these behavioral advantages, since previous studies on bilingual L3A compared to monolingual L2A under different instructional contexts have produced inconsistent findings (Lado, 2008; Lin, 2009; Nation & McLaughlin, 1986; Nayak, et al., 1990).

Electrophysiological

Though there are no models of bilingual L3A which make neurocognitive predictions, the findings in the monolingual L2 processing domain were used to formulate predictions for bilingual L3A. As briefly mentioned above, a potential manifestation of bilingual advantages at additional language learning might be greater language-like processing signatures compared to monolinguals. Thus, regarding Research Questions 2 and 4 it was predicted that bilinguals would show more language-related ERP effects than monolinguals in response to syntactic and morphosyntactic violations, possibly as reflected by more left-lateralized and less centrally distributed ANs as well as more robust P600s. Regarding the instructional context under which such advantages would manifest themselves, language-related processing was predicted to be
more apparent at high compared to low experience and perhaps also more for the Uninstructed than Instructed condition. These final two predictions were motivated by the findings of Morgan-Short and colleagues’ work in monolingual L2A under the same experimental context used here (Morgan-Short, Finger, et al., 2012; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012).
Chapter 3 – Research methods and design

Introduction

This chapter is dedicated to detailing the design and procedure of the study on bilingual L3A and is divided into the following sections. The first section provides a description of the study, including the bilingual L3 learners and monolingual L2 learners of Brocanto2 and the materials used to train and test these groups. The next sections review the design and procedure for training and testing the bilingual L3 learners and discusses how data was coded and analyzed. Note that the design and procedure of this proposed study are identical to those used in Morgan-Short and colleagues’ study on monolinguals (Morgan-Short, 2007; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012) and it is a subset of these monolinguals to which bilingual L3A is compared, thus providing a highly reliable basis on which to compare monolingual L2A and bilingual L3A, under the same exposure conditions and within the same artificial language approach.

Participants

Twenty-four Mandarin-English bilinguals were tested in this study. Mandarin and English were chosen because both languages differ from Brocanto2 in the syntactic and morphosyntactic properties that are the targets of the study. However, of these 24, three participants did not reach low proficiency by criterion (2 Instructed, 1 Uninstructed; see below in Procedure), one participant was fully tested but began learning English after the age of 6 (at 7 years old, Uninstructed), one was fully tested but scored below the pre-established IQ cutoff for participation in the study (Instructed), one participant (Instructed) could not complete the final session of the study because of mechanical failure with the EEG system and one could not re-
schedule their final session within the time limits of the study (Uninstructed, 1-5 days from Session 2). Finally, one participant’s data was discarded due to a large number of eye blinks in their EEG (Uninstructed).

Thus, the participants whose full set of data were included in the present study were 16 early, balanced Mandarin-English bilinguals; 12 female, all right-handed (Oldfield, 1971). As discussed in Chapter 1, these bilinguals began learning both English and Mandarin prior to the age of 6 (Chee, Caplan, et al., 1999; Chee, Tan, & Thiel, 1999; Fabbro, 1999; Paradis, 2004, 2009; Stafford, 2011), were literate in both (Bialystok, 2001; Sanz, 2007), and reported similar ratings of proficiency in the two languages (Marian, et al., 2007). Participants were pseudo-randomly assigned (to enable matching on other factors) to either an Instructed or Uninstructed learning condition (referred to by Morgan-Short and colleagues as Explicit and Implicit). Nine bilinguals (7 female) were trained in the Instructed condition and 7 in the Uninstructed condition (5 female). Morgan-Short and colleagues (Morgan-Short, 2007; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012) analyzed the data of 30 subjects (16 Instructed and 14 Uninstructed), but because only a limited number of Mandarin-English bilinguals who fit the study’s criteria were available and completed the entire study, a subset of the original monolingual subject groups was analyzed in comparison to the present bilinguals. Ten Instructed (7 female) and 10 Uninstructed (7 female) monolinguals were pseudo-randomly (to ensure matching in age, IQ, years of education, etc.) selected from the original pool of 30 subjects.

Aside from the bilingualism of the Mandarin-English subjects, they were matched as best as possible to the monolingual control group to limit differences in other factors between the monolinguals and bilinguals and across the Instructed and Uninstructed training conditions. All subjects had normal or corrected vision and no neurological or developmental disorders. Subjects
were screened for their experience in learning additional languages (aside from English and Mandarin for the bilinguals) and met the following criteria: no more than 3 years of classes and no more than 2 weeks of immersion, and none at all of either in the last 3 years, especially for Romance languages or other languages with nominal gender agreement (i.e., German).

In the bilingual group, Instructed and Uninstructed subjects were not different in terms of age, years of education, IQ, or number of foreign languages studied (all \( p > .05 \) on independent samples \( t \)-tests, for descriptive statistics see Table 2). They were also not different in age of Chinese exposure, age of English exposure, or their proficiency ratings in any modality for either language (all \( p < .05 \), see Table 1).

Participants were also given Dunn and Fox Tree’s Bilingual Dominance Scale (Dunn & Fox Tree, 2009), modified for Mandarin and English (see Appendix A). This tool has been confirmed to correlate with independent measures of lexical and sentence translation, which attests to its reliability as a measure of balance and bilingual skill in their two languages. In this scale, a composite score close to zero indicates balance in the two languages (-5 to +5 is an acceptable range to confirm balance), or the absolute value of each language’s score can be compared. The Instructed bilinguals had an average composite score of -1.4 and the Uninstructed bilinguals had an average composite score of -.85 – with no statistical difference between the two groups, \( t(14) = -.119, p = .907 \). This indicates that not only were both groups balanced in Mandarin and English, they were also not different from each other in their balance. This was also true for the absolute values of Mandarin and English scores on this questionnaire: English, \( t(14) = .299, p = .770 \); Mandarin, \( t(14) = .587, p = .567 \).
Table 1

Descriptive information of language status in bilinguals

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Mandarin</th>
</tr>
</thead>
<tbody>
<tr>
<td>AoEng</td>
<td>2.8 (2.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>AoMdn</td>
<td>8.6 (1.5)</td>
<td>9.1 (1.0)</td>
</tr>
<tr>
<td>S</td>
<td>9.2 (0.9)</td>
<td>9.3 (1.0)</td>
</tr>
<tr>
<td>U</td>
<td>9.4 (0.8)</td>
<td>9.4 (0.7)</td>
</tr>
<tr>
<td>R</td>
<td>7.7 (2.7)</td>
<td>8.1 (2.3)</td>
</tr>
</tbody>
</table>

Instructed

<table>
<thead>
<tr>
<th></th>
<th>3.5 (2.6)</th>
<th>0 (0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AoEng</td>
<td>8.7 (1.5)</td>
<td>9.1 (1.4)</td>
</tr>
<tr>
<td>AoMdn</td>
<td>9.0 (1.0)</td>
<td>9.4 (0.7)</td>
</tr>
<tr>
<td>S</td>
<td>9.4 (0.7)</td>
<td>9.4 (0.7)</td>
</tr>
<tr>
<td>U</td>
<td>9.2 (0.9)</td>
<td>9.3 (1.0)</td>
</tr>
<tr>
<td>R</td>
<td>7.7 (2.7)</td>
<td>8.1 (2.3)</td>
</tr>
</tbody>
</table>

Uninstructed

Notes. Information presented represents group means with standard deviations in parentheses. AoEng = Age of first exposure to English in years; AoMdn = Age of first exposure to Mandarin in years; S= Speaking; U= Understanding; R=Reading; rating scale of 0-10 where 10= perfect and 0= none (Marian, et al., 2007).

Table 2

Descriptive characteristics of Instructed and Uninstructed bilinguals and monolinguals

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Age</th>
<th>Years Ed.</th>
<th>KBIT</th>
<th>S.1</th>
<th>S.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>9</td>
<td>20.6 (3.2)</td>
<td>14.7 (3.0)</td>
<td>110 (9.8)</td>
<td>2.0 (1.0)</td>
<td>2.2 (1.0)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>10</td>
<td>23.6 (4.2)</td>
<td>15.5 (2.4)</td>
<td>121 (13.5)</td>
<td>1.7 (1.4)</td>
<td>1.9 (1.2)</td>
</tr>
<tr>
<td>Uninstructed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>7</td>
<td>21.8 (2.1)</td>
<td>15.7 (1.8)</td>
<td>114 (15.5)</td>
<td>2.0 (1.7)</td>
<td>2.1 (1.3)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>10</td>
<td>25.9 (6.0)</td>
<td>16.5 (2.2)</td>
<td>120 (15.9)</td>
<td>1.3 (1.1)</td>
<td>2.5 (1.6)</td>
</tr>
</tbody>
</table>

Notes. Information presented represents group means with standard deviations in parentheses. Years Ed. = years of education; KBIT= the Kaufman Brief Intelligence Test to measure general IQ; S.1= days between sessions 1 and 2 of the study; S.2= days between sessions 2 and 3 of the study.

In the monolingual group, Instructed and Uninstructed subjects were not different in age, years of education, IQ, or number of foreign languages studied (all p > .05, Tables 2 and 3). There were no differences between monolinguals and bilinguals across Instructed and Uninstructed conditions in age, years of education, or IQ, but there was a significant difference between monolinguals and bilinguals in number of foreign languages studied in the Instructed group, t(17) = -2.27, p = .032, where the monolinguals reported knowing or having studied more languages than the bilinguals (Table 3). This same measure approached significance in
comparisons between bilinguals and monolinguals in the Uninstructed group, $r(15) = -1.79, p = .093$.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Information on additional language experience in Instructed and Uninstructed bilinguals and monolinguals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Additional languages studied</td>
</tr>
<tr>
<td>Instructed</td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>1.1 (.60)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>2.2 (1.3)</td>
</tr>
<tr>
<td>Uninstructed</td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>1.4 (.97)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>2.4 (1.1)</td>
</tr>
</tbody>
</table>

Notes. Group means are presented with standard deviations in parenthesis, as well as ranges. ‘Samples’ refers to a sampling of the languages participants listed as having studied.

Materials

The artificial language: Brocanto2

The target language used in this study was Brocanto2. This is an artificial language that was designed by Morgan-Short and colleagues in order to provide control over a number of factors that are difficult to control for in natural language research, including age of first exposure to the target language as well as language transfer effects and amount, timing, and type of language input. Additionally, because Brocanto2 is composed of a small lexicon and only a few grammar rules, it can be learned to high levels of performance over the course of several days, and therefore allows for an in-depth examination of the trajectory of learning, from low to higher performance or experience. The design of Brocanto2 was based off of another artificial language used by Angela Friederici and colleagues (Friederici, et al., 2002; Opitz & Friederici,
That artificial language (Brocanto) and the one used in the proposed study have elicited neural patterns similar to those found in natural language processing (Friederici, et al., 2002; Morgan-Short, Finger, et al., 2012; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012; Opitz & Friederici, 2002), which suggests that the learning of Brocanto2 in a laboratory setting has the potential to be highly informative for our understanding of the learning of full, natural languages in real-word settings.

Brocanto2 has lexicon of 13 pronounceable nonce words that can be combined to form 1,404 possible sentences which refer to game moves in a computerized chess-like game. There are 4 nouns in Brocanto2, each of which refers to a game token (pleck, neep, vode, blom) - 2 are feminine and 2 are masculine (gender is not overtly marked on the nouns). There is one article which is marked for linguistic gender (li, masculine; lu, feminine) and there are 2 adjectives, also marked for linguistic gender (trois-, neim-; troise/neime – masculine; troiso/neimo – feminine). The language also has 2 adverbs (noyka, zayma) and 4 verbs (klin, praz, yab, nim) that refer to the actual game moves (move, switch, release, capture, respectively) and differ with respect to transitivity (Table 4 below). Brocanto2 has fixed subject-object-verb word order at the sentence level and noun-(adjective)-article word at the noun-phrase level – making it a head-final language. A sentence such as Pleck troise li vode troiso lu praz means that “the round pleck switches with the round vode”. The targets of learning in the present study were word order and gender agreement. Note that participants did not see Brocanto2 until the end of the study; during training, practice, and ERP assessment Brocanto2 was presented solely auditorily.
Table 4

Descriptive characteristics of Brocanto2

<table>
<thead>
<tr>
<th>Grammatical category</th>
<th>Brocanto2</th>
<th>Token</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns (n = 4)</td>
<td>pleck\textsubscript{mas}</td>
<td>neep\textsubscript{mas}</td>
</tr>
<tr>
<td>Article (n = 1)</td>
<td>li\textsubscript{mas} - u\textsubscript{fem}</td>
<td>the</td>
</tr>
<tr>
<td>Adjectives (n = 2)</td>
<td>trois \textsubscript{mas} - o\textsubscript{fem}</td>
<td>round</td>
</tr>
<tr>
<td></td>
<td>neim \textsubscript{mas} - o\textsubscript{fem}</td>
<td>square</td>
</tr>
<tr>
<td>Adverbs (n = 2)</td>
<td>noyka</td>
<td>vertically</td>
</tr>
<tr>
<td></td>
<td>zayma</td>
<td>horizontally</td>
</tr>
<tr>
<td>Verbs (n = 4)</td>
<td>praz\textsubscript{transitive}</td>
<td>swap</td>
</tr>
<tr>
<td></td>
<td>klin\textsubscript{intransitive}</td>
<td>move</td>
</tr>
<tr>
<td></td>
<td>yab\textsubscript{transitive/intransitive}</td>
<td>release</td>
</tr>
<tr>
<td></td>
<td>nim\textsubscript{transitive/intransitive}</td>
<td>capture</td>
</tr>
</tbody>
</table>

Note: Subscript \textsubscript{mas} = masculine linguistic gender; subscript \textsubscript{fem} = feminine linguistic gender.

The lexicon of Brocanto2 follows English phonotactics, and none of the words exist in English. In Mandarin, \textit{li} can either be proper noun or a reference to a unit of distance (falling-rising tone means the distance is approximately 500 meters) and \textit{lu} can have several meanings, depending on whether the tone is rising, falling, or rising-falling and also on the sentence context. Mandarin is a tonal language (Zhou, Ye, Cheung, & Chen, 2009) and thus meaning is
encoded in the rising and falling tones on lexical items. Given this, the words “li” and “lu” as they are presented in Brocanto2 (always with even tone, that is, never rising or falling) are unlikely to be confused with Mandarin possibilities (for example, cottage, mushroom, natural salt). Although this was a slight limitation of targeting Mandarin as one of the already-learned languages of the bilinguals, no other bilingual groups were accessible who do not have more severe limitations in their language characteristics. Japanese, for example, has the same subject-object-verb phrase structure as Brocanto2, as does German. Romance languages, such as Spanish or French, have nominal gender agreement on the article and adjective (as does Brocanto2). Thus, Mandarin was the best candidate for controlling language transfer confounds in terms of the target language characteristics that were the focus of the study: word order and nominal gender agreement.

**Design**

**Instructional conditions**

Morgan-Short trained monolinguals under two instructional conditions: explicit or implicit. The bilinguals in the present study were trained under these exact same conditions, but the conditions are instead referred to as Instructed (instead of explicit) and Uninstructed (instead of implicit).

In the Instructed condition participants underwent approximately 13 minutes⁶ of training in Brocanto2. During these 13 minutes, they were presented with 33 meaningful examples of Brocanto2 (constellations or game moves presented along with the corresponding Brocanto2 sentence). Also during this time, participants received metalinguistic explanations about the ___________________

---

⁶ Morgan-Short (2007) reports that the Explicit condition was 13 minutes, 8 seconds long and the Implicit condition was 13 minutes, 29 seconds long.
grammar of Brocanto2, for example, “There is only one article in Brocanto2, which has two forms: a masculine form and a feminine form.”

In the Uninstructed condition participants also received the same 33 examples presented in the Instructed condition. However, in this condition an additional 94 meaningful examples were presented in order to equate the two conditions in terms of amount of time spent in training. No metalinguistic explanations were provided.

Morgan-Short and colleagues controlled for several factors across these training conditions. Both conditions began with the presentation of simple phrases and moved to complex sentences, the frequency of words in both instructional conditions was the same across grammatical categories, and the timing between items in each condition as well as total training time were held constant. Thus, the only difference between the two conditions was designed to be the type of input learners received during the 13 minutes of training: meaningful examples + metalinguistic explanations in the Instructed condition versus an abundance of meaningful examples – metalinguistic explanations in the Uninstructed condition. These conditions were designed to approximate real-world language learning situations in traditional classroom (Instructed) and communicative classroom or naturalistic (Uninstructed) settings. Appendices B and C provide the full scripts for the two training conditions.

**Comprehension and production practice in Brocanto2**

Over the course of the study, participants engaged with a total of 44 blocks of comprehension and production practice in Brocanto2. In comprehension practice, participants heard a Brocanto2 sentence and made the corresponding game move on the computer screen using a mouse. In production practice, participants viewed a game move on the computer screen and said aloud the Brocanto2 sentence that described that move. Participants’ score in the game
increased or decreased depending on the accuracy of their response; the scoring was controlled by the researcher during production practice and automatically by the computer program during comprehension practice. The practice blocks were completed in alternating blocks of two (two comprehension, then two production, etc.) and each block consisted of 20 items, or game moves.

Behavioral/ERP assessment

At two points in the study, participants underwent behavioral/ERP assessments in Brocanto2. Specifically, participants completed an untimed grammaticality judgment task (GJT) while EEG data were recorded. The GJTs (of which there were 4 versions, counter-balanced across subjects and test sessions) contained 240 items, 120 of these contained grammatical violations in Brocanto2, 40 of which were phrase structure violations and 48 of which were gender agreement violations. The following presentation sequence occurred for each sentence: A fixation cross appeared and a Brocanto2 sentence was presented auditorily through earphones. A prompt (“Good?”) was presented 500 ms after the last word of each sentence. Subjects had up to 5 seconds to make a judgment with a mouse and the next sentence was presented right after the response.

EEG acquisition followed the same parameters as those used by Morgan-Short and colleagues (Morgan-Short, 2007; Morgan-Short, Finger, et al., 2012; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012). Scalp EEG was recorded in DC mode at a sampling rate of 500 Hz from 64 tin electrodes (10-20 system, (Jasper, 1958) mounted in an elastic cap (Electro-Cap International, Inc., Eaton, OH). Data were analyzed using EEProbe software (Advanced Neuro Technology, ANT, Enschede, the Netherlands). Scalp electrodes were referenced online to the left mastoid and impedances were kept below 5kΩ. Electrooculagram (EOG) activity was measured using free electrodes placed above and below the right eye.
(vertical EOG) and on the right and left canthi (horizontal EOG). EEG was amplified by Neuroscan SynAmps². Data was filtered online with a band-pass filter (DC to 100 Hz, 24-dB/octave attenuation) and offline using a 0.16-30 Hz band-pass filter. Participants’ data from target words free of artifacts greater than 30 μV in the EOG and greater than 75 μV in the EEG were included in subsequent analyses.

**Procedure**

Following subject screening and Informed Consent, subjects were pseudo-randomly assigned to either the Instructed or Uninstructed learning conditions. After being assigned to one of these conditions during Session 1 of Brocanto2 learning, participants were taught the rules of the chess-like computer game of Brocanto2. This game provided a meaningful context in which the language could be learned; the rules of the game are independent of the grammar rules of Brocanto2 and playing the game consisted of comprehension and production-based game moves. Increased accuracy during comprehension and production increased the participant’s score, as mentioned above. Participants read the rules of the game at their own pace and practiced making the 4 moves (capture, release, move, switch) with a computer mouse. This portion of training was presented in English; no Brocanto2 or English-equivalent words (“capture”, for example) were used.

Following this introduction to the rules of Brocanto2 the game, participants learned the names of the 4 game tokens in Brocanto2 the language. During this vocabulary training subjects saw a token on the computer screen and listened to the name of it by clicking the image with a computer mouse. Participants moved through the training at their own pace and were allowed to hear the names of the tokens as many times as they wanted. Once they signaled to the researcher
that they had learned the names of the tokens, they took a vocabulary quiz on which they had to score 100% accuracy for naming before continuing to the next step of the study.

Immediately after vocabulary training, participants engaged with either the Instructed or Uninstructed exposure condition. Following this, participants completed comprehension and production practice (described below) until they reached above-chance accuracy on two consecutive comprehension practice blocks. This is considered low proficiency/experience in Brocanto2. Once this criterion was reached, participants continued to their first behavioral/ERP assessment.

Session 2 of the study took place one to five days following Session 1 (depending on subjects’ schedules). During this session participants again engaged with Instructed or Uninstructed training in Brocanto2 (per subject this was the same training condition they were exposed to in Session 1). After this additional training, subjects again completed comprehension and production practice, up to block 36 of the 44 total practice blocks in the study.

Session 3 took place one to five days following Session 2. During this session participants completed the remaining 8 blocks of Brocanto2 practice (4 comprehension, 4 production) and then continued to the second behavioral/ERP assessment. After this behavioral/ERP assessment, participants filled out a detailed questionnaire that probed for their awareness and metalinguistic knowledge of Brocanto2 and also a brief questionnaire that asked participants whether they had searched for any information about the artificial language in between sessions. This questionnaire was not included in Morgan-Short and colleagues’ original study with the monolinguals, but was used in the present study as a precaution against subjects searching for and potentially finding information from the already-published papers on
Brocanto2. Following completion of this questionnaire, subjects’ participation in the study was complete.

**Data analysis**

The behavioral data were analyzed following Morgan-Short and colleagues (Morgan-Short, 2007; Morgan-Short, Finger, et al., 2012; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012). Accuracy data for Brocanto2 comprehension and production as well as reaction time data for comprehension blocks was automatically recorded by the computer program, and it is this accuracy and reaction time data that were submitted to statistical analyses using SPSS, version 21 (IBM Corp.). Participants’ responses to the grammaticality judgment task completed during EEG acquisition were transformed to $d'$ (d-prime) scores, which provide an unbiased measure of participants’ ability to discriminate between correct and violation items (Macmillan & Creelman, 2005; Stanislaw & Todorov, 1999; Wickens, 2002). These scores were calculated based on performance on the GJT at low and high experience using the formula $d' = z(\text{hit rate}) - z(\text{false alarm rate})$. These values were the index of performance used in statistical analyses, though accuracy is also reported below.

In order to answer Research Questions 1 and 3, the $d'$ scores were entered into separate Analysis of Variance (ANOVA) for the linguistic targets of word order and gender agreement (SPSS, version 21; IBM Corp.) in each Instructional Condition. Each ANOVA contained Bilingualism (Bilingual/Monolingual) as the between-subjects factor and Time (low/high experience) as the within-subjects factor.

For Research Questions 2 and 4, the analysis of the electrophysiological data also followed that of Morgan-Short and colleagues. The EEG data, which were time-locked to the onsets of correct and violation items, were averaged for each subject for both lateral and midline
electrodes using a 200ms pre-stimulus baseline. An array of 42 electrodes was included in the lateral analysis (Morgan-Short, Finger, et al., 2012). These electrodes comprised seven levels of anterior/posterior distribution: Anterior 0- FP3, FF3, FF1, FF2, FF4, FP4; Anterior 1 – F7, F5, F3, F4, F6, F8; Anterior 2 – FC7, FC5, FC3, FC5, FC6, FC8; Central 1 – T3, C5, C3, C4, C6, T4; Central 2 – CT7, CT5, CP3, CP4, CT6, CT8; Posterior 1 – T5, P5, P3, P4, P6, T6; and Posterior 2 – OL, PO3, O1, O2, PO4, OR. Within these levels the electrodes also covered Laterality (3 levels) and Hemisphere (2 levels, left and right). Three midline electrodes were also analyzed (Fz, Cz, POz). Analysis of midline electrodes are included here only if they provided additional information not found in the lateral analysis.

Subject averages were entered into grand averages for Instructed and Uninstructed groups for each linguistic target. To answer Research Questions 2 and 4, mean ERP amplitudes were submitted to ANOVAs for each target and Instructional Condition using SAS (version 4.3, SAS Institute Inc.). Within each target, five time-windows (TWs) were analyzed; these are representative of time windows corresponding to the neural patterns of interest in the present study (N400s, ANs, and P600s) and replicate the time windows used in Morgan-Short, Finger, et al. (2012): TW1, 150-300 ms; TW2, 300-500 ms; TW3, 500-700 ms; TW4, 700-900 ms, TW5, 900-1200 ms. The ANOVAs contained Bilingualism (Bilingual/Monolingual) as the between-subjects factor and Time (low/high), Violation (correct/violation) and scalp distribution factors of Hemisphere (left/right), Anterior-Posterior (7 levels), and Laterality (3 levels) as within-subjects factors. Any significant interactions (at p < .05) were further explored with step-down ANOVAs.
Chapter 4 – Results

Introduction

This chapter is dedicated to reviewing the results of behavioral performance on practice blocks and grammaticality judgment tasks as well as the electrophysiological measurements of sensitivity to violations in word order and gender agreement, for Instructed bilinguals and monolinguals and Uninstructed bilinguals and monolinguals. The results are presented as follows. First, performance on practice will be presented for comprehension and production across the four experimental groups. Then, behavioral and electrophysiological results for processing of the word order and gender agreement targets in Brocanto2 will be presented for the Instructed bilinguals and monolinguals followed by behavioral and electrophysiological results for the Uninstructed bilinguals and monolinguals.

Practice performance

There are 44 blocks of Brocanto2 practice. All participants that reached low experience criterion (45% accuracy twice in a row on comprehension blocks) by block 30 in Session 1 participated in the full study and thus completed all 44 blocks by the end of the study (Session 3). Any participant that did not reach this performance criterion was discontinued from the study and their data is not included here. In the original monolingual study, three participants (Instructed n =1; Uninstructed n = 2) were discontinued from the study for failing to reach criterion. In the present study, three bilinguals (Instructed n = 2; Uninstructed n = 1) were discontinued.

Practice performance was operationalized in several ways. Specifically, performance was characterized by how many blocks participants took to reach low experience criterion, average
accuracy at low experience, number of blocks participants took to reach high experience (from low), overall comprehension accuracy, reaction time per block on comprehension, overall production accuracy, and performance on the last 8 blocks of practice, completed in the final session of the study.

First, ANOVAs were conducted on participant data for each of the following, with Bilingualism and Instructional Condition as between-subjects factors: blocks to low experience (Bl_Lo), mean accuracy at low experience (Avg_LP), blocks to high experience\(^7\) (Bl_Hi), mean overall comprehension accuracy (Comp), mean overall production accuracy (Prod), and mean accuracy on the final 8 blocks (Last8). Descriptive statistics for these measures across the four groups can be seen in Table 5. For Bl_LP there was a significant main effect for Instructional Condition, \(F(1,32) = 12.29, p = .001, \eta^2_p = .278^8\), observed power = .925, whereby the Instructed learners reached low experience criterion in fewer blocks than Uninstructed learners, and an Instructional Condition x Bilingualism interaction, \(F(1,32) = 4.66, p = .038, \eta^2_p = .127\), observed power = .554. For this interaction, follow-up analyses using the Scheffé post hoc criterion for significance showed that the Instructional Condition x Bilingualism interaction was driven by the Uninstructed bilinguals going through significantly more practice blocks than the Instructed bilinguals (\(p = .003\)) and monolinguals (\(p = .021\)) to reach low experience. (Note that the Uninstructed monolinguals were not different from any group and the Instructed bilinguals and monolinguals were not different from each other, all \(p > .05\)).

\(^7\) Note that not all participants reached what Morgan-Short (2007) previously identified as “high proficiency:” 95% accuracy twice in a row on comprehension or production blocks (when using all the words of the language in speaking). In the monolingual group, three Uninstructed learners did not reach high proficiency by the end of practice. In the bilingual group, five Uninstructed learners did not reach high proficiency. All Instructed participants reached high proficiency. For ease of exposition and clarity, practice and GJT results are framed within the term high experience. In order to calculate Bl_Hi for learners who did not reach that pre-specified criterion for high proficiency, the final block of practice was used as the reference value.

\(^8\) Following Green and Salkind (2005), values of \(\eta^2_p\) in the .01–.05 range signal small effect sizes, in the .06–.13 range signal medium effect sizes, and values greater than or equal to .14 indicate large effect sizes.
Table 5

Practice performance in Instructed and Uninstructed bilinguals and monolinguals

<table>
<thead>
<tr>
<th></th>
<th>Bl_Lo</th>
<th>Avg_LP</th>
<th>Bl_Hi</th>
<th>Comp</th>
<th>Prod</th>
<th>Last8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instructed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>3 (2.7)</td>
<td>56.7 (11.6)</td>
<td>9 (6.3)</td>
<td>87.6 (8.7)</td>
<td>90.7 (6.8)</td>
<td>95.7 (6.8)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>6 (7.2)</td>
<td>60.4 (20.5)</td>
<td>7.9 (6.4)</td>
<td>85.5 (13.9)</td>
<td>84.3 (15.2)</td>
<td>94.1 (5.5)</td>
</tr>
<tr>
<td><strong>Uninstructed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>23 (5.6)</td>
<td>39.8 (6.7)</td>
<td>23 (5.6)</td>
<td>62.9 (12.8)</td>
<td>28.8 (25.3)</td>
<td>67.4 (17.9)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>9 (8.6)</td>
<td>58.8 (27.2)</td>
<td>17 (16.5)</td>
<td>75.9 (18.0)</td>
<td>55.5 (33.4)</td>
<td>79.0 (23.8)</td>
</tr>
</tbody>
</table>

Notes. Information presented represents group means, with standard deviations in parentheses. Bl_LP=number of blocks to reach low experience criterion; Avg_Lo=average accuracy at low experience criterion; Bl_Hi=number of blocks to reach high experience/proficiency; Comp=comprehension accuracy; Prod=production accuracy; Last8=accuracy on the final 8 blocks of practice.

For Avg_LP there were no main effects or interactions: Instructional Condition, $F(1,32) = 2.01, p = .166$; Bilingualism, $F(1,32) = 2.86, p = .100$; Instructional Condition x Bilingualism, $F(1,32) = 1.41, p = .244$. For Bl_Hi, there was a main effect for Instructional Condition, $F(1,32) = 11.65, p = .002, \eta^2_p = .267$, observed power = .912, whereby the Instructed learners reached high experience in fewer blocks from low than the Uninstructed learners. No other main effects or interactions were found. For Comp performance, a main effect for Instructional Condition was again found, $F(1,32) = 13.17, p = .001, \eta^2_p = .292$, observed power = .940, wherein Instructed learners outperformed Uninstructed learners on comprehension accuracy. A main effect for Instructional Condition was also found for Prod, $F(1,32) = 36.01, p = .000, \eta^2_p = .529$, observed power = 1.0 and Last8 performance, $F(1,32) = 18.61, p = .000, \eta^2_p = .368$, observed power = .987, in which the Instructed learners outperformed Uninstructed learners on production accuracy overall and also for performance on the final 8 blocks of practice (no other significant main effects or interactions).
In sum, for practice performance across these measures, the instructional condition containing metalinguistic information led to superior performance across both monolinguals and bilinguals as compared to a condition where no such metalinguistic information was provided. For bilinguals, the training condition containing metalinguistic information seemed particularly important as compared to the one without this information in terms of reaching low experience criterion.

A second set of analyses was performed across the practice blocks to investigate changes over time across the four groups. A repeated-measures ANOVA with Instructional Condition and Bilingualism as between-subjects factors and Time (4 series) as the within-subjects factor revealed a main effect for Time, $F(3, 96) = 44.96, p = .000, \eta^2_p = .584$, indicating that all learners improved as practice progressed (Table 6). Between-subjects, there was a main effect for Instructional Condition, $F(1,32) = 28.37, p = .000, \eta^2_p = .470$, observed power = .999, wherein the Instructed learners outperformed Uninstructed learners. There was also an Instructional Condition x Bilingualism interaction that approached significance, $F(1,32) = 4.03, p = .053, \eta^2_p = .055$, observed power = .263. In sum, all learners improved over the course of practice in Brocanto2, and Instructed learners routinely outperformed Uninstructed learners (see Figure 1 which shows performance across the full 44 blocks).

---

9 Note that for these time series the 44 blocks were divided into 4 series of 11 practice blocks (grouped consecutively), categorized as Time1, Time2, Time3, Time4. The full 44 blocks could not be used as separate time series in a repeated-measures design because the ANOVA would offer no interpretative value for any potential interactions among variables (Rusan Chen, p.c.) so a more statistically manageable time series was calculated from participants’ practice block performance data.
### Table 6

Practice accuracy over the 44 blocks of practice (time series) in Instructed and Uninstructed bilinguals and monolinguals

<table>
<thead>
<tr>
<th></th>
<th>Time1</th>
<th>Time2</th>
<th>Time3</th>
<th>Time4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instructed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>72.8 (16.2)</td>
<td>92.8 (7.1)</td>
<td>95.1 (4.7)</td>
<td>96.0 (3.7)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>65.9 (26.3)</td>
<td>86.5 (24.1)</td>
<td>93.8 (7.6)</td>
<td>94.5 (4.4)</td>
</tr>
<tr>
<td><strong>Uninstructed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>23.1 (28.23)</td>
<td>44.2 (22.2)</td>
<td>53.7 (21.9)</td>
<td>81.6 (21.3)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>56.8 (29.0)</td>
<td>64.1 (33.5)</td>
<td>73.8 (25.1)</td>
<td>86.7 (19.4)</td>
</tr>
</tbody>
</table>

**Notes.** Information presented represents group means, with standard deviations in parentheses. Time 1 = Blocks 1-11, Time 2 = Blocks 12-22, Time 3 = Blocks 23-33, Time 4 = Blocks 34-44.

![Figure 1](image)

**Figure 1.** Accuracy across all 44 practice blocks.
Similar analyses were performed on comprehension and production performance separately. To examine comprehension performance over time, a repeated-measures ANOVA with Instructional Condition and Bilingualism as between-subjects factors and Time (4 series\textsuperscript{10}) as the within-subjects factor was performed. The descriptive statistics for these 4 time series in comprehension can be seen in Table 7. The analysis revealed a significant main effect for Time, $F(3,96) = 73.27, p = .000, \eta^2_p = .696$, observed power = 1.0, and a significant Time x Instructional Condition interaction, $F(3,96) = 4.13, p = .008, \eta^2_p = .114$, observed power = .837 within subjects, indicating that all learners improved and the Instructed learners performed better across time than the Uninstructed learners. Between-subjects a main effect for Instructional Condition was again found, $F(1,32) = 13.07, p = .001, \eta^2_p = .290$, observed power = .939, with the Instructed learners outperforming Uninstructed learners overall. Performance data for the four groups across all 22 blocks can be seen in Figure 2.

\textsuperscript{10} For comprehension, the 22 blocks were divided into the following time series: Time1 = blocks 1-4, Time2 = 5-11, Time3 = 12-18, Time4 = 19-22. This was done so that Time 1 and Time 4 represent the first and last 4 comprehension blocks, corresponding conceptually to the measures of performance at the earliest and latest stages of additional language learning in this study. The same procedure was followed for the production blocks.
Table 7

Comprehension accuracy over the 22 blocks of practice (time series) in Instructed and Uninstructed bilinguals and monolinguals

<table>
<thead>
<tr>
<th></th>
<th>Time1</th>
<th>Time2</th>
<th>Time3</th>
<th>Time4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instructed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>66.1 (13.3)</td>
<td>90.4 (10.7)</td>
<td>92.7 (7.9)</td>
<td>95.4 (5.0)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>65.5 (25.4)</td>
<td>85.2 (23.5)</td>
<td>92.2 (7.1)</td>
<td>96.7 (3.2)</td>
</tr>
<tr>
<td><strong>Uninstructed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>35.0 (6.6)</td>
<td>53.5 (19.3)</td>
<td>76.3 (13.7)</td>
<td>84.4 (9.0)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>57.7 (28.3)</td>
<td>71.2 (25.1)</td>
<td>84.3 (15.6)</td>
<td>90.2 (11.2)</td>
</tr>
</tbody>
</table>

*Notes. Information presented represents group means, with standard deviations in parentheses. Time 1 = Blocks 1-4, Time 2 = Blocks 5-11, Time 3 = Blocks 12-18, Time 4 = Blocks 19-22.*

![Figure 2](image.png)

*Figure 2. Accuracy for comprehension performance.*
For production performance over time, the repeated-measures ANOVA with Instructional Condition and Bilingualism as between-subjects factors and Time (4 series) as the within-subjects factor revealed a main effect for Time, $F(3,96) = 38.05, p = .000, \eta_p^2 = .543$, observed power = 1.0, indicating that all learners improved (see Table 8 below for performance information). Between-subjects, there was a main effect for Instructional Condition, $F(1,32) = 35.64, p = .000, \eta_p^2 = .527$, observed power = 1.0 and a significant Instructional Condition x Bilingualism interaction, $F(1,32) = 4.52, p = .041, \eta_p^2 = .124$, observed power = .541, indicating that Instructed groups performed better overall and the Uninstructed Bilinguals performed differently from the other three groups, who do not appear to be different from each other.

Performance data for the four groups across all 22 blocks can be seen in Figure 3.

Table 8  
*Production accuracy over the 22 blocks of practice (time series) in Instructed and Uninstructed bilinguals and monolinguals*

<table>
<thead>
<tr>
<th></th>
<th>Time1</th>
<th>Time2</th>
<th>Time3</th>
<th>Time4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>67.2 (24.0)</td>
<td>94.5 (4.8)</td>
<td>97.3 (2.1)</td>
<td>96.2 (3.3)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>53.3 (31.1)</td>
<td>86.3 (26.5)</td>
<td>96.1 (5.4)</td>
<td>94.1 (7.8)</td>
</tr>
<tr>
<td>Uninstructed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>2.8 (3.3)</td>
<td>23.6 (29.4)</td>
<td>37.3 (35.1)</td>
<td>50.5 (35.3)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>28.1 (36.1)</td>
<td>53.0 (40.5)</td>
<td>66.0 (34.7)</td>
<td>68.5 (37.0)</td>
</tr>
</tbody>
</table>

*Notes.* Information presented represents group means, with standard deviations in parentheses. Time 1 = Blocks 1-4, Time 2 = Blocks 5-11, Time 3 = Blocks 12-18, Time 4 = Blocks 19-22.
Finally, reaction time (RT) data from comprehension performance was considered. Figure 4 shows RTs across the 22 blocks of comprehension practice. A repeated-measures ANOVA with Instructional Condition and Bilingualism as between-subjects factors and Time (4 series) as the within-subjects factor revealed a significant main effect for Time, $F(3, 96) = 47.98$, $p = .000$, $\eta^2_p = .600$, observed power = 1.0, which indicates that learners’ RTs decreased as practice progressed (Table 9). Additionally, there was a significant Time x Instructional Condition x Bilingualism interaction, $F(3, 96) = 32.11$, $p = .006$, $\eta^2_p = .120$, observed power = .858, and a Time x Bilingualism interaction which approached significance, $F(3, 96) = 2.58$, $p = .058$, $\eta^2_p = .075$, observed power = .619. These interactions appear to be driven by the Uninstructed bilinguals being slower in the RTs across time (especially in the beginning of...
practice) than the other three groups, who do not appear to be different from each other (Figure 4). This is echoed in the between-subjects analysis in which there was a significant Bilingualism x Instructional Condition interaction, $F(1,32) = 4.32, p = .046, \eta^2_p = .119$, observed power = .523, with no other significant main effects.

<table>
<thead>
<tr>
<th>Instructed</th>
<th>Time1</th>
<th>Time2</th>
<th>Time3</th>
<th>Time4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilinguals</td>
<td>9.5 (3.3)</td>
<td>4.4 (2.4)</td>
<td>2.9 (1.1)</td>
<td>2.7 (1.5)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>10.6 (3.8)</td>
<td>5.0 (3.1)</td>
<td>3.6 (2.1)</td>
<td>2.5 (1.6)</td>
</tr>
<tr>
<td>Uninstructed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>14.7 (12.6)</td>
<td>10.8 (7.1)</td>
<td>6.1 (4.0)</td>
<td>4.1 (2.2)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>6.5 (3.1)</td>
<td>5.1 (3.2)</td>
<td>4.3 (2.9)</td>
<td>3.2 (2.4)</td>
</tr>
</tbody>
</table>

Notes. Reaction time is in seconds, the information presented represents group means with standard deviations in parentheses. Time 1 = Blocks 1-4, Time 2 = Blocks 5-11, Time 3 = Blocks 12-18, Time 4 = Blocks 19-22.
Figure 4. Reaction time performance on comprehension blocks.

**Practice performance summary**

When considering both bilingualism and also instructional context as important variables in the ability to accurately produce and understand an additional language, the results showed that both bilinguals and monolinguals benefited from receiving metalinguistic information about the language. The results also suggest that, especially for bilinguals, the metalinguistic information was particularly useful in developing the ability to accurately produce language. Overall, however, metalinguistic information was not strictly necessary for production and comprehension skills to develop, as learners in both instructional conditions significantly
improved in their performance, especially for comprehension accuracy and reduced reaction times, as practice progressed.

**Instructed learners: Grammaticality judgment performance**

Accuracy information and $d'$ scores on the GJT for Instructed bilinguals and monolinguals for the word order and gender agreement targets are presented in Tables 10 and 11. For word order performance, the ANOVA revealed a main effect for Time, $F(1,17) = 17.25, p = .001, \eta^2_p = .504$, observed power = .974, indicating that both bilinguals and monolinguals improved in their ability to discriminate correct and violation word order from low to high experience$^{11}$. No other significant main effects or interactions were found. For gender agreement, the ANOVA again revealed a main effect for Time, $F(1,17) = 6.21, p = .023, \eta^2_p = .268$, observed power = .652 and no other main effects or interactions. Thus, both bilinguals and monolinguals improved in their ability to discriminate language violations, at both the syntactic (word order) and morphosyntactic (gender agreement) level, with no statistically significant differences in their ability.

---

$^{11}$ Note that there were not group differences at either low or high experience for word order or gender agreement on GJT performance as assessed by independent samples $t$-tests between bilinguals and monolinguals at each time point and for each structure (all $p > .05$).
Table 10
Performance on word order judgment in Instructed bilinguals and monolinguals

<table>
<thead>
<tr>
<th></th>
<th>Low experience</th>
<th>High experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>d’ score</td>
</tr>
<tr>
<td>Instructed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>65.2 (22.8)</td>
<td>0.96 (1.50)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>82.3 (10.0)</td>
<td>1.93 (0.75)</td>
</tr>
</tbody>
</table>

Notes. Information presented represents groups means with standard deviations in parentheses.

Table 11
Performance on gender agreement judgment in Instructed bilinguals and monolinguals

<table>
<thead>
<tr>
<th></th>
<th>Low experience</th>
<th>High experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>d’ score</td>
</tr>
<tr>
<td>Instructed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>62.1 (16.8)</td>
<td>0.76 (1.08)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>74.8 (14.9)</td>
<td>1.57 (0.99)</td>
</tr>
</tbody>
</table>

Notes. Information presented represents group means with standard deviations in parentheses.

In light of the above results for grammaticality judgment performance in the Instructed learners, the answer to Research Question 1, *When exposed to additional language in a traditional classroom (Instructed) context, is performance in judging correct and violation structures different in bilinguals compared to monolinguals at low and/or high proficiency*, was No, there were no bilingual/monolingual differences on this behavioral measure of additional language performance.

**Instructed learners: Electrophysiological patterns**

**Word Order**

Grand average waveforms and topographical voltage maps for the Instructed bilinguals’ and monolinguals’ processing of word order violations in Brocanto2 are presented in Figures 5
and 6. Visual inspection of these waveforms and maps suggested the following. At both low and high experience, both bilinguals and monolinguals appeared to show an early positivity. At low experience, both groups exhibited a central-parietal negativity (N400), followed by a later anterior negativity that extended into the latest time-window of interest and appeared to have a more focal frontal distribution in bilinguals than monolinguals. Finally, at high experience, both groups exhibited a posterior positivity (P600), though it appeared to have an earlier onset (500-700 ms) in bilinguals than monolinguals. Summaries of the ANOVA F-values from the analysis of the word order target can be found in Table 12.
Figure 5. ERP data. Processing of word order targets in Instructed bilinguals and monolinguals.

Averaged ERPs are for correct (blue) and violation (red) targets. Range -3 – 3 uV; 200 ms pre-stimulus baseline through 1200 ms.
Figure 6. *Topographical Maps*. Voltage maps for processing of word order targets in Instructed bilinguals and monolinguals. Significant effects are indicated by letters (a) – (d). (a) Early anterior-central positivity (150-300 ms) over both groups at both low and high experience. (b) N400 at low experience over both groups (500-700 ms). (c) Late anterior negativity over both groups at low experience beginning the 700-900 ms time-window and continuing. (d) P600 over both groups at high experience beginning in the 700-900 ms time-window and continuing.
Table 12

*Summary of ANOVA F-values for the comparison between Instructed bilinguals and monolinguals
processing of the word order*

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>TW1 150-300 ms</th>
<th>TW2 300-500 ms</th>
<th>TW3 500-700 ms</th>
<th>TW4 700-900 ms</th>
<th>TW5 900-1200ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viol</td>
<td>1,17</td>
<td>4.81*</td>
<td>1.25</td>
<td>&lt;1</td>
<td>1.19</td>
<td>1.87</td>
</tr>
<tr>
<td>T x V</td>
<td>1,17</td>
<td>&lt;1</td>
<td>1.70</td>
<td>7.66*</td>
<td>9.16**</td>
<td>9.78**</td>
</tr>
<tr>
<td>V x Bil</td>
<td>1,17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2.12</td>
<td>1.68</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Bil</td>
<td>1,17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>3.49*</td>
<td>1.05</td>
<td>1.76</td>
</tr>
<tr>
<td>V x Lat</td>
<td>2,34</td>
<td>3.67*</td>
<td>2.20</td>
<td>2.43</td>
<td>11.70**</td>
<td>18.23***</td>
</tr>
<tr>
<td>V x Lat x Bil</td>
<td>2,34</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP</td>
<td>6,102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>4.38*</td>
<td>4.99*</td>
</tr>
<tr>
<td>V x AP x Bil</td>
<td>6,102</td>
<td>1.42</td>
<td>2.91*</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x H</td>
<td>1,17</td>
<td>4.04*</td>
<td>1.80</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x H x Bil</td>
<td>1,17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.89</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat</td>
<td>2,34</td>
<td>2.47</td>
<td>1.40</td>
<td>&lt;1</td>
<td>3.01*</td>
<td>1.50</td>
</tr>
<tr>
<td>T x V x Lat x Bil</td>
<td>2,34</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP</td>
<td>6,102</td>
<td>3.98*</td>
<td>1.86</td>
<td>&lt;1</td>
<td>1.39</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP x Bil</td>
<td>6,102</td>
<td>1.76</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x AP</td>
<td>12,204</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.09</td>
<td>2.22*</td>
<td>1.98*</td>
</tr>
<tr>
<td>V x Lat x AP x Bil</td>
<td>12,204</td>
<td>1.98</td>
<td>2.19*</td>
<td>&lt;1</td>
<td>1.28</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x H</td>
<td>1,17</td>
<td>1.14</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x H x Bil</td>
<td>1,17</td>
<td>&lt;1</td>
<td>1.09</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H</td>
<td>2,34</td>
<td>2.57</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H x Bil</td>
<td>2,34</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.69</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H</td>
<td>6,102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2.04</td>
<td>2.30</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H x Bil</td>
<td>6,102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.18</td>
</tr>
<tr>
<td>T x V x Lat x AP</td>
<td>12,204</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.52</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x AP x Bil</td>
<td>12,204</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x H</td>
<td>1,17</td>
<td>1.14</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x H x Bil</td>
<td>1,17</td>
<td>&lt;1</td>
<td>1.09</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H</td>
<td>2,34</td>
<td>2.57</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H x Bil</td>
<td>2,34</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.69</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H</td>
<td>6,102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2.04</td>
<td>2.30</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H x Bil</td>
<td>6,102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.18</td>
</tr>
<tr>
<td>T x V x Lat x AP</td>
<td>12,204</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.52</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x AP x Bil</td>
<td>12,204</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x H</td>
<td>2,34</td>
<td>1.88</td>
<td>1.73</td>
<td>2.61</td>
<td>1.75</td>
<td>2.81*</td>
</tr>
<tr>
<td>T x V x Lat x H x Bil</td>
<td>2,34</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP x H</td>
<td>6,102</td>
<td>2.48*</td>
<td>3.72*</td>
<td>3.81*</td>
<td>1.97</td>
<td>1.79</td>
</tr>
<tr>
<td>T x V x AP x H x Bil</td>
<td>6,102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x AP x H</td>
<td>12,204</td>
<td>&lt;1</td>
<td>1.33</td>
<td>&lt;1</td>
<td>1.28</td>
<td>1.08</td>
</tr>
</tbody>
</table>
Analysis of the ERP data at TW1 (150-300 ms) revealed a significant main effect of Violation, $F(1,17) = 4.61, p = .0464$. This main effect reflects the anterior-central positivity over both bilinguals and monolinguals at both low and high experience (Figures 5 and 6). There were also four interactions that approached statistical significance: Violation x Laterality ($p = .0662$); Violation x Hemisphere ($p = .0605$); Time x Violation x Anterior-posterior ($p = .0529$); and Time x Violation x Anterior-posterior x Hemisphere ($p = .0861$).

Analyses of the ERP data for TW2 (300-500 ms) revealed a significant Time x Violation x Anterior-posterior x Hemisphere interaction, $F(6,102) = 3.72, p = .0281$. There were also two additional interactions that approached significance: Violation x Anterior-posterior x Bilingualism ($p = .0964$) and Violation x Laterality x Anterior-posterior x Bilingualism ($p = .0769$). Step-down analyses on the significant Time x Violation x Anterior-posterior x Hemisphere interaction did not produce any significant effects.

Analyses of the ERP data for TW3 (500-700 ms) revealed a Time x Violation interaction ($p = .0149$) which was qualified by a Time x Violation x Anterior-posterior x Hemisphere interaction, $F(6,102) = 3.81, p = .0254$. There was also a Time x Violation x Bilingualism interaction which approached significance ($p = .0771$). Step-down analyses on the Time x Violation x Anterior-posterior x Hemisphere interaction in TW3 revealed a main effect for
Violation at low experience over Posterior-2 electrodes in the right hemisphere, $F(1,17) = 5.40, p = .0328$. This effect was shared in both bilinguals and monolinguals and likely reflects an N400 effect at low experience for these learners.

Analysis of the ERP data in the fourth time-window (700-900 ms) revealed three significant interactions: Time x Violation, $F(1,17) = 9.16, p = .0076$; Violation x Laterality, $F(2,34) = 11.70, p = .0026$; and Violation x Anterior-posterior, $F(6,102) = 4.38, p = .0440$. There were also Time x Violation by Laterality ($p = .0944$) and Violation x Laterality x Anterior-posterior ($p = .0640$) interactions that approached significance.

Step-down analyses on the Time x Violation interaction revealed a main effect of violation at low experience, $F(1,17) = 4.93, p = .0403$, reflecting an anterior-central negativity over both bilinguals and monolinguals in this time-window (Figure 5). Analysis of the Violation x Laterality interaction revealed no additional effects with the exception of a main effect for Violation that approached significance at medial electrodes ($p = .0969$). Finally, step down analyses on the Violation x Anterior-posterior interaction revealed a main effect for violation at both Posterior-1, $F(1,17) = 7.00, p = .0170$, and Posterior-2 electrodes, $F(1,17) = 7.76, p = .0127$. This reflects the shared positivity (P600) over both bilinguals and monolinguals, which is particularly evident at high experience. Thus, in this time-window, both bilinguals and monolinguals showed an anterior-to-central negativity at low and a P600 at high (and also low) experience.

Analyses of the ERP data in the final time-window (TW5, 900-1200 ms) revealed significant interactions for Time x Violation, $F(1,17) = 9.78, p = .0061$; Violation x Laterality, $F(2,34) = 18.23, p = .0002$; and Violation x Anterior-posterior, $F(6,102) = 4.81, p = .0316$. There
was also a Time x Violation x Laterality x Hemisphere interaction that approached significance \((p = .0816)\), as well as a Violation x Laterality x Anterior-posterior interaction \((p = .0604)\).

Step-down analyses on the Time x Violation interaction in TW5 revealed, similarly to the results for TW4, a main effect for Violation at low experience, \(F(1,17) = 7.44, p = .0143\), reflecting the continuation of the anterior-central negativity from the previous time-window. Analysis of the Violation x Laterality interaction revealed no significant results except that the effect of Violation at medial electrodes again approached significance \((p = .0511)\). Finally, follow up analyses on the Violation x Anterior-posterior interaction revealed a main effect of Violation at three electrode sites: Central-2, \(F(1,17) = 7.23, p = .0155\); Posterior-1, \(F(1,17) = 14.36, p = .0015\); Posterior-2, \(F(1,17) = 9.90, p = .0059\). This result reflects a posterior positivity over both groups and across low and high experience, though it is more obvious at high (Figures 5 and 6).

To summarize, both bilinguals and monolinguals exhibited an anterior-central positivity at the earliest time-window (150-300 ms) following a violation in Brocanto2 word order, and this was found for processing at both low and high experience. Second, at low experience, both groups of learners exhibited an apparent N400 effect in the 500-700 ms time-window; this negativity had a predominantly posterior, right hemisphere distribution. In later time-windows at low experience both groups continued to show a negativity (700-900 ms and 900-1200 ms), though visual examination of the voltage maps and waveforms suggests that the negativity becomes less posterior and more anterior-central as time goes on (up to 900-1200 ms). Finally, both bilinguals and monolinguals showed a P600 in response to Brocanto2 word order violations. This effect is most obvious at high experience and low power due to small group sizes may have
concealed an interaction with Time (note the interaction just above that included this factor and approached significance).

**Gender Agreement**

Grand average waveforms and topographical voltage maps for the Instructed bilinguals’ and monolinguals’ processing of gender agreement violations in Brocanto2 are presented in Figures 7 and 8. Visual inspection of these waveforms and maps suggested the following. At low experience, the Instructed bilinguals appeared to show an early, centrally-distributed negativity that continued into later time-windows. The monolinguals instead showed a late positivity. At high proficiency, the bilinguals showed a late anterior-central negativity and a weak posterior positivity. The monolinguals also showed a late anterior negativity, together with a posterior positivity (P600). Summaries of the ANOVA F-values from the analysis of the gender agreement target can be found in Table 13.
Figure 7. ERP data. Processing of gender agreement targets in Instructed bilinguals and monolinguals. Averaged ERPs are for correct (blue) and violation (red) targets. Range -3 – 3 uV; 200 ms pre-stimulus baseline through 1200 ms.
Figure 8. Topographical Maps. Voltage maps for processing of gender agreement targets in Instructed bilinguals and monolinguals. Significant effects are indicated by letters (a) – (d). (a) Late anterior negativity at low and high experience beginning in the 700-900 ms time-window. (b) P600 at low and high in the 700-900 ms time-window, apparent in the monolinguals. (c) P600 at low in the monolinguals, contrasts with (d) the late anterior negativity at low in the bilinguals.
### Table 13

*Summary of ANOVA F-values for the comparison between Instructed bilinguals and monolinguals’ processing of gender agreement*

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>TW1 150-300 ms</th>
<th>TW2 300-500 ms</th>
<th>TW3 500-700 ms</th>
<th>TW4 700-900 ms</th>
<th>TW5 900-1200 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viol</td>
<td>1,17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.23</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V</td>
<td>1,17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.69</td>
<td>9.21**</td>
</tr>
<tr>
<td>V x Bil</td>
<td>1,17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>3.02*</td>
<td>5.23*</td>
</tr>
<tr>
<td>T x V x Bil</td>
<td>1,17</td>
<td>2.57</td>
<td>1.15</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.74</td>
</tr>
<tr>
<td>V x Lat</td>
<td>2.34</td>
<td>5.33*</td>
<td>7.07**</td>
<td>2.12</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x Bil</td>
<td>2.34</td>
<td>1.02</td>
<td>&lt;1</td>
<td>1.46</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP</td>
<td>6.102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2.24</td>
<td>11.89***</td>
<td>13.76***</td>
</tr>
<tr>
<td>V x AP x Bil</td>
<td>6.102</td>
<td>1.40</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x H</td>
<td>1.17</td>
<td>&lt;1</td>
<td>3.23*</td>
<td>1.97</td>
<td>&lt;1</td>
<td>1.30</td>
</tr>
<tr>
<td>V x H x Bil</td>
<td>1.17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2.27</td>
<td>1.83</td>
</tr>
<tr>
<td>T x V x Lat</td>
<td>2.34</td>
<td>&lt;1</td>
<td>1.27</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x Bil</td>
<td>2.34</td>
<td>1.19</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.37</td>
<td>1.04</td>
</tr>
<tr>
<td>T x V x AP</td>
<td>6.102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>3.04*</td>
<td>1.31</td>
</tr>
<tr>
<td>T x V x AP x Bil</td>
<td>6.102</td>
<td>1.46</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x AP</td>
<td>12.204</td>
<td>1.15</td>
<td>1.23</td>
<td>&lt;1</td>
<td>1.25</td>
<td>2.52*</td>
</tr>
<tr>
<td>V x Lat x AP x Bil</td>
<td>12.204</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x H</td>
<td>1.17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>3.92</td>
<td>2.86</td>
</tr>
<tr>
<td>T x V x H x Bil</td>
<td>1.17</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H</td>
<td>2.34</td>
<td>1.43</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H x Bil</td>
<td>2.34</td>
<td>2.59</td>
<td>4.16*</td>
<td>2.12</td>
<td>1.07</td>
<td>1.17</td>
</tr>
<tr>
<td>V x AP x H</td>
<td>6.102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.39</td>
<td>&lt;1</td>
<td>2.15</td>
</tr>
<tr>
<td>V x AP x H x Bil</td>
<td>6.102</td>
<td>&lt;1</td>
<td>1.25</td>
<td>1.23</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x AP</td>
<td>12.204</td>
<td>1.68</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x AP x Bil</td>
<td>12.204</td>
<td>1.26</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x H</td>
<td>2.34</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x H x H</td>
<td>2.34</td>
<td>&lt;1</td>
<td>1.05</td>
<td>1.93</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP x H x H</td>
<td>6.102</td>
<td>3.48*</td>
<td>1.30</td>
<td>&lt;1</td>
<td>1.75</td>
<td>1.79</td>
</tr>
<tr>
<td>T x V x AP x H x x H</td>
<td>6.102</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.00</td>
<td>1.82</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>
Notes. T = Time (Low/High); V/Viol = Violation (Correct/Violation structure); AP = Anterior-posterior; Lat = Laterality; H = Hemisphere; Bil = Bilingualism (Bilingual/Monolingual). +p ≤ .10, *p ≤ .05, **p ≤ .01, ***p ≤ .001.

Analysis of the ERP data for TW1 (150-300 ms) revealed a significant Violation x Laterality interaction, $F(2,34) = 5.33, p = .0291$. There was also a significant Time x Violation x Anterior-posterior x Hemisphere interaction, $F(6,102) = 3.48, p = .0455$. However, step-down analyses on these interactions revealed no further effects.

Analysis of ERP data for TW2 (300-500 ms) revealed a Violation x Laterality interaction ($p = .0236$) which was qualified by a Violation x Laterality x Hemisphere x Bilingualism interaction, $F(2,34) = 6.40, p = .0096$. However, follow up analyses revealed no further effects.

Analysis of ERP data for TW3 (500-700 ms) revealed no significant interactions.

In TW4 (700-900 ms) the ANOVA revealed a significant Violation x Anterior-posterior interaction, $F(6,102) = 12.07, p = .0007$. There was also a Violation x Bilingualism interaction that approached significance ($p = .0906$) as well as a Time x Violation x Anterior-posterior interaction ($p = .0859$).

Step-down analysis of the Violation x Anterior-posterior interaction in TW4 showed a main effect of Violation at the following electrodes sites: Anterior-0, $F(1,17) = 7.01, p = .0169$; Anterior-1, $F(1,17) = 5.02, p = .0386$; Anterior-2, $F(1,17) = 5.67, p = .0292$; and Posterior-2,
$F(1,17) = 6.14, p = .0240$. These results reflect an anterior negativity and posterior positivity (P600) over both groups, which is more visually salient at high experience as compared to low. (Note again that an interaction with Time may be obscured by low power.)

Finally, analysis of the last time-window of interest (TW5, 900-1200 ms) revealed interactions for Time x Violation, $F(1,17) = 9.21, p = .0075$; Violation x Anterior-posterior, $F(6,102) = 13.76, p = .0004$; and Violation x Bilingualism, $F(1,17) = 5.23, p = .0353$. There was also a Violation x Laterality x Anterior-posterior interaction that trended towards significance ($p = .0602$). Follow up analysis of the Time x Violation interaction showed a main effect of Violation at low experience, $F(1,17) = 8.45, p = .0098$ which is likely indexing different effects in monolinguals and bilinguals in this time-window: an anterior negativity in the bilinguals but a posterior-central positivity in the monolinguals (Figures 7 and 8).

Follow up analysis of the Violation x Anterior-posterior interaction showed a main effect of Violation at the following electrode sites: Anterior-0, $F(1,17) = 5.37, p = .0333$; Posterior-1, $F(1,17) = 6.32, p = .0205$; Posterior-2, $F(1,17) = 8.90, p = .0084$. These results, like those from TW4, are taken to reflect an anterior negativity and posterior positivity (P600) in both bilinguals and monolinguals which is especially salient at high experience. Finally, the Violation x Bilingualism interaction revealed a main effect of Violation that approached significance in the bilinguals ($p = .0558$).

In sum, neither group exhibited early electrophysiological responses to violations in Brocanto2 gender agreement at either low or high proficiency, but in later time-windows both bilinguals and monolinguals exhibited a late anterior negativity (beginning 700-900 ms) together with a P600, the latter of which is more obvious in the monolinguals. This pattern is visually salient at high experience in both groups of learners. At low experience in the final time-window
(900-1200 ms) is when the bilinguals and monolinguals appear to be least similar – with the bilinguals showing a late anterior negativity and the monolinguals exhibiting a posterior positivity.

**Instructed learners: Summary of electrophysiological results**

The electrophysiological data from Instructed bilinguals and monolinguals suggests that the two groups of learners are similar with regards to their neurocognitive patterns in response to violations in the additional language, especially at high experience. However, both groups appeared to respond somewhat differently to word order as compared to gender agreement violations. For word order, both bilinguals and monolinguals showed an early anterior positivity at low and high experience and an N400 at low. Also at low experience, both bilinguals and monolinguals showed a continuing negativity up to 900-1200 ms following the violation; this negativity, at 700-900 and 900-1200 ms, appears to be less posterior and more anterior-central than that of the 500-700 ms time-window. Finally, both groups of learners exhibited a P600 in these later time-windows (700-900 ms and 900-1200 ms), which is most obvious at high experience.

For gender agreement, neither bilinguals nor monolinguals exhibited early electrophysiological responses to violations. Not until the fourth time-window (700-900 ms) did a main effect of Violation reach statistical significance. In this and the subsequent time-window (900-1200 ms), both bilinguals and monolinguals showed a late anterior negativity together with a P600, and this response appeared most obvious at high experience. At low experience in the final time-window (900-1200 ms) is perhaps the only time when bilinguals and monolinguals clearly diverged in their neural patterns, with the bilinguals exhibiting a late anterior negativity while the monolinguals showed instead a posterior positivity consistent with a P600.
To conclude, Research Question 2 asked *When exposed to additional language in a traditional classroom (Instructed) context, are there differences in neurocognitive processing, as indexed by ERPs, in bilinguals compared to monolinguals at low and/or high proficiency?* The answer to this question was both Yes and No. Within each linguistic target of word order and gender agreement, there were not significant differences between bilinguals and monolinguals with respect to their electrophysiological responses to language violations. At high experience the two groups neural patterns appear more similar than different, though perhaps stronger in bilinguals than monolinguals. At low experience in the final time-window of interest for gender agreement there is evidence for a difference in processing between the two groups.

**Uninstructed learners: Grammaticality judgment performance**

Accuracy information and *d’* scores on the GJT for Uninstructed bilinguals and monolinguals for the word order and gender agreement targets are presented in Tables 14 and 15. For word order performance, the ANOVA revealed a main effect for Time, $F(1,15) = 59.12, p = .000, \eta_p^2 = .798$, observed power = 1.0, whereby all learners improved in their ability to discriminate correct and violation word order structures. Between-subjects, there was also a main effect for Bilingualism, $F(1,15) = 5.75, p = .030, \eta_p^2 = .277$, observed power = .612, but in the unexpected direction: the monolinguals were better than the bilinguals in judging this target.

For gender agreement, there was also a significant main effect for Time, $F(1,15) = 12.43, p = .003, \eta_p^2 = .453$, observed power = .909, indicating, as with word order, that learners improved in their ability to judge illicit items from low to high experience. There was also a Time x Bilingualism interaction which approached significance, $F(1,15) = 3.20, p = .094, \eta_p^2 =$.
.176, observed power = .388. This trending interaction\textsuperscript{12} is reinforced by the main effect of Bilingualism between-subjects, $F(1,15) = 5.45$, $p = .034$, $\eta^2_p = .267$, observed power = .589, wherein the monolinguals outperform the bilinguals in judging gender agreement violations.

In sum, both bilinguals and monolinguals in the Uninstructed condition improved in their ability to discriminate language violations, at both the syntactic and morphosyntactic level. However, monolinguals outperformed bilinguals in judging both structures. Thus, in response to Research Question 3, \textit{When exposed to additional language in a communicative classroom or more naturalistic (Uninstructed) context, is performance in judging correct and violation structures different in bilinguals compared to monolinguals at low and/or high proficiency}, it seems as though, Yes, there are differences between monolinguals and bilinguals, though not in the direction expected under the premise of bilingual advantages at additional language learning.

\begin{table}[h]
\centering
\caption{Performance on word order judgment in Uninstructed bilinguals and monolinguals}
\begin{tabular}{lccccc}
\hline
 & \multicolumn{2}{c}{Low experience} & \multicolumn{2}{c}{High experience} \\
 & Accuracy & $d'$ score & Accuracy & $d'$ score \\
\hline
Uninstructed & & & & \\
Bilinguals & 55.0 (13.7) & 0.31 (0.83) & 78.7 (9.7) & 1.72 (0.79) \\
Monolinguals & 71.8 (15.0) & 1.34 (0.97) & 88.9 (14.5) & 2.88 (1.23) \\
\hline
\end{tabular}
\textit{Notes.} Information presented represents groups means with standard deviations in parentheses.
\end{table}

\textsuperscript{12} Independent-samples $t$-tests between the groups show that though there are no differences at low experience [$t(15) = 1.43, p = .172$], monolinguals outperform bilinguals at high, $t(15) = 2.49, p = .025$. 

126
Table 15

*Performance on gender agreement judgment in Uninstructed bilinguals and monolinguals*

<table>
<thead>
<tr>
<th></th>
<th>Low experience</th>
<th></th>
<th>High experience</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>$d'$ score</td>
<td>Accuracy</td>
<td>$d'$ score</td>
</tr>
<tr>
<td><strong>Uninstructed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilinguals</td>
<td>49.7 (3.5)</td>
<td>-0.02 (0.16)</td>
<td>58.3 (10.9)</td>
<td>0.40 (0.74)</td>
</tr>
<tr>
<td>Monolinguals</td>
<td>59.8 (16.8)</td>
<td>0.54 (1.04)</td>
<td>75.1 (18.4)</td>
<td>1.86 (1.41)</td>
</tr>
</tbody>
</table>

*Notes.* Information presented represents group means with standard deviations in parentheses.

**Uninstructed learners: Electrophysiological patterns**

**Word Order**

Grand average waveforms and topographical voltage maps for the Uninstructed bilinguals’ and monolinguals’ processing of word order violations in Brocanto2 are presented in Figures 9 and 10. At low experience, the bilinguals appeared to show minimal effects with the exception of a late positivity with an anterior distribution. The monolinguals, in contrast, appeared to show a centrally-distributed negativity (N400) which continued into later time-windows. At high experience, the bilinguals exhibited an early anterior positivity and a potential (though weak) N400 effect. The monolinguals at high on the other hand showed an AN followed by a later anterior negativity together with a P600 in later time-windows. Summaries of the ANOVA F-values from the analysis of the word order target can be found in Table 16.
Figure 9. ERP data. Processing of word order targets in Uninstructed bilinguals and monolinguals. Averaged ERPs are for correct (blue) and violation (red) targets. Range -3 – 3 uV; 200 ms pre-stimulus baseline through 1200 ms.
Figure 10. Topographical Maps. Voltage maps for processing of word order targets in Uninstructed bilinguals and monolinguals. Significant effects are indicated by letters (a) – (e). (a) N400 at low in the 300-500 ms time-window for the monolinguals. (b) P600 at low for the bilinguals, contrasts with (c) the late anterior negativity in the monolinguals in this same time-window. (d) Early anterior negativity in the monolinguals at low experience, contrasts with (e) the N400 at high experience in this same time-window for the bilinguals.
Table 16
Summary of ANOVA F-values for the comparison between Uninstructed bilinguals and monolinguals’ processing of word order

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>TW1 150-300 ms</th>
<th>TW2 300-500 ms</th>
<th>TW3 500-700 ms</th>
<th>TW4 700-900 ms</th>
<th>TW5 900-1200 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viol</td>
<td>1,15</td>
<td>&lt;1</td>
<td>4.85*</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V</td>
<td>1,15</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Bil</td>
<td>1,15</td>
<td>4.01*</td>
<td>2.83</td>
<td>1.97</td>
<td>1.49</td>
<td>2.30</td>
</tr>
<tr>
<td>T x V x Bil</td>
<td>1,15</td>
<td>1.90</td>
<td>&lt;1</td>
<td>1.90</td>
<td>1.07</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat</td>
<td>2,30</td>
<td>&lt;1</td>
<td>2.31</td>
<td>1.38</td>
<td>&lt;1</td>
<td>1.69</td>
</tr>
<tr>
<td>V x Lat x Bil</td>
<td>2,30</td>
<td>3.00*</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP</td>
<td>6,90</td>
<td>&lt;1</td>
<td>1.21</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>4.13*</td>
</tr>
<tr>
<td>V x AP x Bil</td>
<td>6,90</td>
<td>3.00*</td>
<td>1.20</td>
<td>2.51</td>
<td>2.45</td>
<td>2.08</td>
</tr>
<tr>
<td>V x H</td>
<td>1,15</td>
<td>&lt;1</td>
<td>1.33</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.35</td>
</tr>
<tr>
<td>V x H x Bil</td>
<td>1,15</td>
<td>1.00</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat</td>
<td>2,30</td>
<td>1.66</td>
<td>&lt;1</td>
<td>1.52</td>
<td>5.78*</td>
<td>3.40*</td>
</tr>
<tr>
<td>T x V x Lat x Bil</td>
<td>2,30</td>
<td>1.92</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.22</td>
</tr>
<tr>
<td>T x V x AP</td>
<td>6,90</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>4.02*</td>
<td>2.25</td>
</tr>
<tr>
<td>T x V x AP x Bil</td>
<td>6,90</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.66</td>
<td>2.12</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x AP</td>
<td>12,180</td>
<td>1.83</td>
<td>3.02*</td>
<td>3.71*</td>
<td>2.08</td>
<td>1.81</td>
</tr>
<tr>
<td>V x Lat x AP x Bil</td>
<td>12,180</td>
<td>1.96</td>
<td>2.34*</td>
<td>2.03</td>
<td>1.89</td>
<td>1.90</td>
</tr>
<tr>
<td>T x V x H</td>
<td>1,15</td>
<td>&lt;1</td>
<td>4.24*</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x H x Bil</td>
<td>1,15</td>
<td>1.70</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H</td>
<td>2,30</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H x Bil</td>
<td>2,30</td>
<td>&lt;1</td>
<td>1.26</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H</td>
<td>6,90</td>
<td>&lt;1</td>
<td>1.02</td>
<td>&lt;1</td>
<td>1.91</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H x Bil</td>
<td>6,90</td>
<td>1.85</td>
<td>1.39</td>
<td>1.98</td>
<td>1.06</td>
<td>1.90</td>
</tr>
<tr>
<td>T x V x Lat x AP</td>
<td>12,180</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.07</td>
<td>1.36</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x AP x Bil</td>
<td>12,180</td>
<td>1.15</td>
<td>1.15</td>
<td>1.36</td>
<td>1.24</td>
<td>1.43</td>
</tr>
<tr>
<td>T x V x Lat x H</td>
<td>2,30</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.39</td>
</tr>
<tr>
<td>T x V x Lat x H x Bil</td>
<td>2,30</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.40</td>
</tr>
<tr>
<td>T x V x AP x H</td>
<td>6,90</td>
<td>2.08</td>
<td>&lt;1</td>
<td>5.21**</td>
<td>6.27**</td>
<td>4.48*</td>
</tr>
<tr>
<td>T x V x AP x H x Bil</td>
<td>6,90</td>
<td>1.22</td>
<td>2.36*</td>
<td>4.37**</td>
<td>3.60*</td>
<td>2.93*</td>
</tr>
</tbody>
</table>
Analysis of the ERP data in TW1 (150-300 ms) revealed a significant 6-way interaction, for Time x Violation x Laterality x Anterior-posterior x Hemisphere x Bilingualism, $F(12,180) = 3.78, p = .0163$. There were four interactions which approached significance: Violation x Bilingualism ($p = .0636$), Violation x Laterality x Bilingualism ($p = .0979$), Violation x Anterior-posterior x Bilingualism ($p = .0923$), and Time x Violation x Laterality x Anterior-posterior x Hemisphere ($p = .0800$).

Follow-up analyses on the significant 6-way interaction did not produce any further significant results, though a main effect for Violation did approach significance in bilinguals for Anterior-0 and Lateral-2 electrode sites in the right hemisphere at high experience, $F(1,6) = 4.44, p = .0797$. This would reflect the anterior positivity visible in the bilinguals at high experience (Figure 9).

Analysis of the ERP data in TW2 (300-500 ms) revealed a main effect of Violation, $F(1,15) = 4.85, p = .0438$ which is likely driven by the negativity in monolinguals at both low and high experience, and the weak negativity visible most clearly at high in the bilinguals. There was also a significant interaction for Violation x Laterality x Anterior-posterior ($p = .0435$).
which was qualified by a Time x Violation x Laterality x Anterior-posterior x Hemisphere x Bilingualism, $F(12,180) = 3.19, p = .0427$. Additionally, there were several interactions that approached significance: Violation x Laterality x Anterior-posterior x Bilingualism ($p = .0909$), Time x Violation x Hemisphere ($p = .0573$), and Time x Violation x Anterior-posterior x Bilingualism ($p = .0990$).

Follow-up analyses on the significant 6-way interaction, like those for TW1, did not produce any significant results. However, a main effect for Violation did approach significance in the bilingual learners at high experience, also like TW1, but instead was found at Anterior-2 and Lateral-2 electrode sites in the right hemisphere, $F(1,6) = 5.07, p = .0654$. Again, this would be driven by an anterior positivity in the bilinguals (Figure 9).

Analysis of ERP data in TW3 (500-700 ms) revealed the following effects. There was a Violation x Laterality x Anterior-posterior interaction, $F(12,180) = 3.71, p = .0150$ as well as a Time x Violation x Anterior-posterior x Hemisphere interaction, $F(6,90) = 5.21, p = .0048$, which was qualified by a Time x Violation x Anterior-posterior x Hemisphere x Bilingualism interaction, $F(6,90) = 4.37, p = .0111$. There was also a 6-way interaction for Time x Violation x Laterality x Anterior-posterior x Hemisphere x Bilingualism that approached significance at $p = .0542$. Follow-up analyses, however, revealed no subsequent significant results.

In the fourth time-window (TW4, 700-900 ms) there were four significant interactions: Time x Violation x Laterality, $F(2,30) = 5.78, p = .0249$, Time x Violation x Anterior-posterior, $F(6,90) = 4.01, p = .0455$ and Time x Violation x Anterior-posterior x Hemisphere, $F(6,90) = 6.27, p = .0016$, which were both qualified by a 5-way Time x Violation x Anterior-posterior x Hemisphere x Bilingualism interaction, $F(6,90) = 3.60, p = .0235$. There was an additional 6-way interaction that approached significance (Time x Violation x Laterality x Anterior-posterior...
x Hemisphere x Bilingualism, \( p = .0690 \). Unfortunately, none of the significant interactions reported above produced significant results in the step-down analyses.

Finally, analyses on TW5 (900-1200 ms) revealed a significant Violation x Anterior-posterior interaction (\( p = .0443 \)) which was qualified by a Time x Violation x Anterior-posterior x Hemisphere interaction, \( F(6,90) = 4.48, p = .0139 \). There were two interactions that approached significance: Time x Violation x Anterior-posterior x Hemisphere x Bilingualism (\( p = .0577 \)) and Time x Violation x Laterality (\( p = .0771 \)).

Step-down analyses on the Time x Violation x Anterior-posterior x Hemisphere interaction revealed a main effect for Violation at Anterior-1 electrode sites for bilinguals and monolinguals at low experience in the right hemisphere, \( F(1,15) = 5.34, p = .0355 \). A similar main effect of Violation (again over both groups, at low proficiency) approached significance at Anterior-0, \( F(1,15) = 4.04, p = .0628 \), as well as Posterior-2 sites, \( F(1,15) = 3.25, p = .0913 \). Though the effect was found over both groups in this time-window, it is clearly indexing different patterns – a positivity in the bilinguals, but a negativity in the monolinguals.

Note that that the visually salient AN-P600 and later anterior negativity patterns in the monolinguals were not statistically borne out in this analysis. This may stem from low power in comparing across both of the groups for low and high experience. Note also that in the full set of subjects, this group of monolinguals did indeed show the patterns that are visible in the waveforms and voltage maps (Morgan-Short, Steinhauer, et al., 2012).

Though many of the significant interactions revealed in the ANOVA did not step down to significant results, there are indications of differences between bilinguals and monolinguals in their processing of word order violations. First, the trend towards significance in the first and second time-windows at high experience for bilinguals points to the presence of an anterior
positivity in these learners that is not exhibited in the monolinguals. Second, in the final time-window of interest at low experience, bilinguals exhibited a positivity while the monolinguals instead exhibited a negativity. Additionally, in TW2 (300-500 ms) there is a negativity that is apparent only in the monolinguals, and consistent with an N400 effect at low. At high experience, the negativity is more frontally distributed in the monolinguals and would be consistent with an AN for these learners, but the distribution in the bilinguals is better aligned with an N400 effect. However, note given the lack of significance with Bilingualism and other distributional factors, it is difficult to make robust claims in this small group of subjects.

**Gender Agreement**

Grand average waveforms and topographical voltage maps for the Uninstructed bilinguals’ and monolinguals’ processing of gender agreement violations in Brocanto2 are presented in Figures 11 and 12. At low experience, the bilinguals appeared to exhibit an AN with a left-lateralized distribution in the 300-500 ms time-window, followed by a (weak) posterior positivity and a later anterior-central negativity (900-1200 ms). The monolinguals also appeared to show this late anterior-central negativity in the 900-1200 ms time-window, but showed no positivity in any time-window. Additionally, though they did exhibit a negativity in the 300-500 ms time-window, it appeared to be more centrally-distributed than that of the bilinguals - consistent with an N400. At high experience, it appeared that the bilinguals showed an early, posterior positivity, a later (500-700) negativity with an anterior-central distribution and a positivity (P600) beginning in the 700-900 ms time-window and continuing. The monolinguals appeared to show few effects early in processing, with a P600 beginning around 700-900 ms and continuing into the final time-window. Summaries of the ANOVA F-values from the analysis of the gender agreement target can be found in Table 17.
Figure 11. ERP data. Processing of gender agreement targets in Uninstructed bilinguals and monolinguals. Averaged ERPs are for correct (blue) and violation (red) targets. Range -3 – 3 uV; 200 ms pre-stimulus baseline through 1200 ms.
Figure 12. Topographical Maps. Voltage maps for processing of gender agreement targets in Uninstructed bilinguals and monolinguals. Significant effects are indicated by the letter (a). (a) P600 at high experience in the 900-1200 ms time-window for both groups.
Table 17

Summary of ANOVA F-values for the comparison between Uninstructed bilinguals and monolinguals’ processing of gender agreement

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>TW1</th>
<th>TW2</th>
<th>TW3</th>
<th>TW4</th>
<th>TW5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>150-300 ms</td>
<td>300-500 ms</td>
<td>500-700 ms</td>
<td>700-900 ms</td>
<td>900-1200 ms</td>
</tr>
<tr>
<td>Viol</td>
<td>1,15</td>
<td>&lt;1</td>
<td>1.59</td>
<td>1.14</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V</td>
<td>1,15</td>
<td>&lt;1</td>
<td>6.44*</td>
<td>&lt;1</td>
<td>4.17*</td>
<td>4.80*</td>
</tr>
<tr>
<td>V x Bil</td>
<td>1,15</td>
<td>2.45</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Bil</td>
<td>1,15</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>3.20*</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat</td>
<td>2,30</td>
<td>&lt;1</td>
<td>7.36*</td>
<td>2.52</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x Bil</td>
<td>2,30</td>
<td>6.75*</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.56</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP</td>
<td>6,90</td>
<td>1.30</td>
<td>4.13*</td>
<td>3.52*</td>
<td>1.40</td>
<td>6.49*</td>
</tr>
<tr>
<td>V x AP x Bil</td>
<td>6,90</td>
<td>3.65*</td>
<td>4.95*</td>
<td>5.26*</td>
<td>1.77</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x H</td>
<td>1,15</td>
<td>2.74</td>
<td>4.80*</td>
<td>1.96</td>
<td>2.61</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x H x Bil</td>
<td>1,15</td>
<td>1.43</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat</td>
<td>2,30</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2.44</td>
<td>2.83</td>
</tr>
<tr>
<td>T x V x Lat x Bil</td>
<td>2,30</td>
<td>&lt;1</td>
<td>1.37</td>
<td>3.19*</td>
<td>3.51*</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP</td>
<td>6,90</td>
<td>&lt;1</td>
<td>1.55</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP x Bil</td>
<td>6,90</td>
<td>3.37*</td>
<td>2.70</td>
<td>2.00</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x AP</td>
<td>12,180</td>
<td>1.22</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.34</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x AP x Bil</td>
<td>12,180</td>
<td>1.84</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2.05</td>
<td>2.98*</td>
</tr>
<tr>
<td>T x V x H</td>
<td>1,15</td>
<td>&lt;1</td>
<td>2.32</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x H x Bil</td>
<td>1,15</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H</td>
<td>2,30</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x Lat x H x Bil</td>
<td>2,30</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H</td>
<td>6,90</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.16</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>V x AP x H x Bil</td>
<td>6,90</td>
<td>2.38</td>
<td>1.43</td>
<td>1.52</td>
<td>3.16*</td>
<td>2.25</td>
</tr>
<tr>
<td>T x V x Lat x AP</td>
<td>12,180</td>
<td>1.92</td>
<td>&lt;1</td>
<td>1.85</td>
<td>1.90</td>
<td>1.79</td>
</tr>
<tr>
<td>T x V x Lat x AP x Bil</td>
<td>12,180</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.14</td>
<td>1.99</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x H</td>
<td>2,30</td>
<td>&lt;1</td>
<td>4.38*</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x Lat x H x Bil</td>
<td>2,30</td>
<td>&lt;1</td>
<td>2.38</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP x H</td>
<td>6,90</td>
<td>&lt;1</td>
<td>1.47</td>
<td>2.52</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>T x V x AP x H x Bil</td>
<td>6,90</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>
Analysis of the ERP data in TW1 (150-300 ms) revealed a significant interaction for Violation x Laterality x Bilingualism, $F(2,30) = 6.75, p = .0165$ and two interactions which approached significance: Violation x Anterior-posterior x Bilingualism ($p = .0659$) and Time x Violation x Anterior-posterior x Bilingualism ($p = .0786$). Step-down ANOVAs on the Violation x Laterality x Bilingualism interaction revealed no subsequent effects.

Analysis of the ERP data in TW2 (300-500 ms) revealed an interaction for Time x Violation ($p = .0228$), Violation x Laterality ($p = .0111$), and Violation x Hemisphere ($p = .0448$) which were all qualified by a Time x Violation x Laterality x Hemisphere interaction, $F(2,30) = 4.38, p = .0382$. There was also a significant Violation x Anterior-posterior x Bilingualism interaction, $F(6,90) = 4.95, p = .0357$. Step-down analyses for the Violation x Anterior-posterior x Bilingualism interaction revealed no further significant effects, nor did step-down analyses of the Time x Violation x Laterality x Hemisphere interaction.

Analyses of the ERP data in TW3 (500-700 ms) revealed a single significant interaction, for Violation x Anterior-posterior x Bilingualism, $F(6,90) = 5.26, p = .0199$. Three additional interactions approached significance: Time x Violation x Bilingualism ($p = .0936$), Violation x

<table>
<thead>
<tr>
<th>Bil</th>
<th>V x Lat x AP x H</th>
<th>12,180</th>
<th>1.35</th>
<th>1.25</th>
<th>1.63</th>
<th>1.89</th>
<th>1.64</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V x Lat x AP x H</td>
<td>12,180</td>
<td>1.52</td>
<td>1.15</td>
<td>1.86</td>
<td>2.16*</td>
<td>1.25</td>
</tr>
<tr>
<td>x Bil</td>
<td>T x V x Lat x AP x H</td>
<td>12,180</td>
<td>&lt;1</td>
<td>1.02</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>T x V x Lat x AP x H</td>
<td>12,180</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>1.32</td>
<td>1.05</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

Notes. T = Time (Low/High); V/Viol = Violation (Correct/Violation structure); AP = Anterior-posterior; Lat = Laterality; H = Hemisphere; Bil = Bilingualism (Bilingual/Monolingual). +p≤.10, *p≤.05, **p≤.01, ***p≤.001.
Anterior-posterior ($p = .0582$), and Time x Violation x Laterality x Bilingualism ($p = .0877$). Unfortunately, follow-up analyses of the Violation x Anterior-posterior x Bilingualism interaction revealed no further effects.

Analyses of the ERP data in TW4 (700-900 ms) did not produce any significant effects, though there were three interactions that trended towards statistical significance: Time x Violation ($p = .0591$), Time x Violation x Laterality x Bilingualism ($p = .0664$), Violation x Anterior-posterior x Hemisphere x Bilingualism ($p = .0561$), and Violation x Laterality x Anterior-posterior x Hemisphere x Bilingualism ($p = .0773$).

In the final time-window (TW5, 900-1200 ms) there was a Time x Violation interaction, $F(1,15) = 4.80, p = .0447$ and also a Violation x Anterior-posterior interaction ($p = .0136$). The latter of these interactions was qualified by a Violation x Laterality x Anterior-posterior x Bilingualism interaction, $F(12,180) = 2.98, p = .0491$. Step-down ANOVAs on these interactions revealed a main effect of Violation at high experience in the Time x Violation interaction, $F(1,15) = 5.66, p = .0311$. This effect indexes the P600 in both groups for this time-window. No other follow-up analyses revealed significant effects.

The set of results for this target and instructional condition revealed the lowest magnitude of effects in the study, with the only statistically significant effect being the P600 in the final time-window of interest at high experience for both bilinguals and monolinguals. However, given the intriguing sets of interactions that were revealed in the ANOVA and also the visually apparent differences for bilinguals and monolinguals in the processing of this target, it is probable that many of the effects failed to step down to significance due to the small number of subjects involved in this analysis.
Uninstructed learners: Summary of electrophysiological results

The electrophysiological data from Uninstructed bilinguals and monolinguals suggests that, like the Instructed groups, there are differences in the neurocognitive profiles for processing word order and gender agreement violations. However, the results also suggest that there may be important differences between Uninstructed bilinguals and monolinguals’ neural patterns in response to violations in the additional language. For word order both groups showed a negativity in the 300-500 ms time-window, though the saliency and distribution of this effect were different at low and high experience and between the bilinguals and monolinguals. The effect at low is an apparent N400 in the monolinguals. At high experience, the negativity in the bilinguals was more posterior than that of the monolinguals (which is more anterior-central than the bilinguals’ and also than the negativity at low proficiency) and thus likely reflects an N400 in the bilinguals but an AN in the monolinguals. Additionally, later in processing (the final time-window) the bilinguals exhibited a positivity with an anterior distribution whereas the monolinguals showed a negativity. Finally, at high experience, the bilinguals showed an early anterior positivity that approached significance and was not evident in the monolinguals.

For gender agreement, the only effect to reach statistical significance was the P600 at high experience for both bilinguals and monolinguals. However, it is clear from the visualizations of the ERP data for these learners (Figures 11 and 12) that the neural responses are (a) somewhat distinct from those for word order and (b) dissimilar between bilinguals and monolinguals. Unfortunately, the low number of subjects may be inhibiting other meaningful effects from being realized in the present set of data.

To conclude, Research Question 4 asked When exposed to additional language in a communicative classroom or more naturalistic (Uninstructed) context, are there differences in
neurocognitive processing, as indexed by ERPs, in bilinguals compared to monolinguals at low and/or high proficiency? The answer to this question was tentatively Yes. Within the word order target, the neural profiles of bilinguals and monolinguals were different at both low and high proficiency. However, within the gender agreement target, Research Question 4 is more difficult to conclusively answer. Though there are apparent differences between bilinguals and monolinguals’ ERP data, the shared P600 effect was the only significant result.
Chapter 5 – Discussion and Conclusion

This chapter reviews the predictions put forth in Chapter 3 and situates the findings of this study within the broader context of theory and experimentation on bilingual as compared to monolingual additional language learning. The results will first be addressed, followed by limitations and areas for future research, and general conclusions drawn from the present study.

Discussion

Behavioral performance

In Chapter 3, it was predicted that bilinguals would outperform monolinguals in judging correct and violation structures, especially at high experience and possibly even at low. This prediction was made based on previous research which has indeed found bilingual advantages on behavioral measures of additional language learning as compared to monolinguals (e.g., Cenoz & Valencia, 1994; Kaushanskaya & Marian, 2009a; Safont Jorda, 2003; Sanz, 2000; Thomas, 1988). No predictions were made regarding whether or how instructional context would interact with the predicted behavioral advantage.

In an instructional condition that was designed to mimic traditional foreign language classroom exposure by providing metalinguistic grammar explanations together with meaningful examples, no significant differences were found between bilinguals and monolinguals on any behavioral assessment used in the present study: the grammaticality judgment task across both word order and gender agreement targets and the various indices of practice performance.

One of the first studies to investigate the role of instructional context in bilingual L3A compared to monolingual L2A, Nation and McLaughlin (1986), also failed to find performance
differences between bilinguals and monolinguals - on AG learning under an explicit, rule-search condition. Differences in their study were only found in the implicit condition, and only for multilinguals as compared to the bilinguals and monolinguals, who were not different from each other. In the study by Nayak et al. (1990) differences were indeed found in the explicit condition whereby multilinguals outperformed monolinguals (but recall that no bilinguals were tested). Also, in the study by Lado (2008), there were similarities in performance among Native-like, High, and Intermediate levels of bilingualism when learners were provided with metalinguistic information. And finally, in Stafford, Sanz, and Bowden (2010) early and late bilinguals were also found to perform similarly when their instructional condition contained metalinguistic information, and it was posited that such a condition may level the playing field and obscure some individual differences, such as age of acquisition.

In the context of these laboratory studies, then, it is not surprising that the present study did not find significant behavioral differences between bilinguals and monolinguals in the Instructed condition. When providing language learners with essentially all of the information they need to complete the learning task, it may be the case that adult learners generally make sufficient use of such information - performing similarly and quite well – and some individual difference factors, such as bilingualism here, become muted by the environment. However, when considering the results from dual- and single-language environments on bilingual L3A as compared to monolingual L2A, the absence of behavioral differences in the Instructed condition are somewhat more difficult to reconcile.

In previous work, most of the research comparing bilingual L3A and monolingual L2A in dual- or single-language contexts reported bilingual advantages in performance (e.g., Cenoz & Valencia; Sanz, 2000; Keshavarz & Astanheh, 2004; Klein, 1995), and these studies were all
conducted on bilinguals and monolinguals who were learning their additional language in a classroom setting. Such a setting would presumably be similar to either the Instructed (traditional classroom) or Uninstructed (communicative classroom, or naturalistic setting) condition employed here. Thus, it may not seem intuitively obvious why the present study failed to find similar bilingual advantages when comparing Instructed bilinguals and monolinguals. However, the very fact that bilinguals and monolinguals in this previous research were already studying the additional language in a classroom setting at the time they were tested may account for the differences between their pattern of results and those of the current study.

Background information gathered from the bilinguals and monolinguals who were trained and tested on Brocanto2 revealed, as discussed in Chapter 3, a significant difference between the two groups in amount of foreign language learning experience: the Instructed monolinguals reported knowing or having studied more foreign languages than the bilinguals. The bilinguals learned Mandarin and English by and large via naturalistic exposure at home and once they began school in the United States, and reported little experience in learning other additional languages. The monolinguals then, despite their broad monolingual status, were likely “more experienced” than the bilinguals at learning additional languages in classroom environments similar to the Instructed condition. This difference is a parallel explanation for why bilingual advantages did not emerge in this study – had the bilinguals been equated with the monolinguals on classroom language learning experience they may have in fact been more ready to activate any one of the posited sources of a bilingual advantage at language learning, such as enhanced metalinguistic awareness (e.g., Jessner, 2008), better strategy use (e.g., Kemp, 2007), or their broader linguistic repertoire (e.g., Cenoz, 2011). As an aside, recall that in the studies conducted in dual- and single-language environments there were important uncontrolled variables such as
amount of exposure, age of exposure, and language similarity, and it is impossible to know how the constellation of these and other factors involved in adult additional language learning may have interacted with the outcomes in those studies to produce the reported bilingual advantages. Additionally, those studies largely used broad measures of performance and were thus measuring quite different abilities from the targeted grammaticality judgment tasks used here.

Therefore, one possible explanation for the absence of a discernible bilingual advantage in the Instructed condition is the greater amount of foreign language learning experience in the monolinguals, coupled with the provision of metalinguistic information. This allowed the monolinguals to perform at their highest potential, and evenly with the bilinguals who, despite their lack of foreign language experience, were also able to capitalize on the provided metalinguistic information. The evidence that Instructed learners outperformed the Uninstructed learners on practice measures lends credence to the idea that the provision of metalinguistic information buttresses learning outcomes and so it is not surprising that both bilinguals and monolinguals could successfully use such information.

The superior performance by the Instructed bilinguals and monolinguals as compared to Uninstructed bilinguals on practice measures also aligns itself nicely with the previous meta-analyses that explored the effectiveness of more and less explicit instructional contexts in monolingual L2A (Norris & Ortega, 2000; Spada & Tomita, 2010). All four groups evidenced learning in the present study, but the rate was faster and outcomes were higher for practice in the Instructed than Uninstructed groups. This provides evidence that both monolingual L2A and bilingual L3A benefit noticeably from the provision metalinguistic information, at least in terms of language comprehension and production. The results also, however, underscore the fact that such information is not strictly necessary for language learning and development to take place.
(Grey, Williams, & Rebuschat, Forthcoming 2014; Leung & Williams, 2011; Rebuschat & Williams, 2012; Rosa & O'Neill, 1999; VanPatten & Oikkenon, 1996; Williams & Lovatt, 2003).

Despite the findings that both bilinguals and monolinguals in this study evidenced learning of Brocanto2 in the absence of metalinguistic information about the language, the results also showed that Uninstructed bilinguals routinely performed worse than their monolingual counterparts. In practice, Uninstructed bilinguals were different from Uninstructed monolinguals and Instructed and Uninstructed bilinguals (who were not different from each other) in terms of reaching low proficiency and accurately producing the language. On the grammaticality judgment task, the Uninstructed monolinguals outperformed the bilinguals on discriminating correct and violation structures for both word order and gender agreement.

This result is perhaps the most surprising of the present study in terms of behavioral assessments, especially considering that while some studies have not found performance differences between bilinguals and monolinguals on aspects of additional language learning (Bartolotti, et al., 2011; Gibson, et al., 2001; Klein, 1995; Lin, 2009), only one has found a monolingual advantage (Okita & Jun Hai, 2001), and only for the learning of an additional language writing system. The results are also surprising in the context of the laboratory-based research on bilingual L3A compared to monolingual L2A which found that the less explicit condition that did not contain metalinguistic information revealed bilingual advantages in performance (e.g., Lado, 2008, but recall null effects in Lin, 2009; Nayak, et al., 1990).

There is, however, a tenable explanation for the results found here. This explanation relates, as above regarding the Instructed groups, to the amount of foreign language learning experience between bilinguals and monolinguals in the present study. Though the Uninstructed bilinguals and monolinguals were not statistically different in terms of this variable, the
difference did trend to a significant difference. Thus, it is at least the case that the Uninstructed bilinguals and monolinguals were somewhat different in terms of their experience in learning additional languages in classroom environments. The Uninstructed condition did not contain metalinguistic information and was instead characterized by an abundance of meaningful input, and in this way may have been more similar to a communicative language classroom (Lee & VanPatten, 1995).

The task of the learners in the Uninstructed condition was largely to ‘figure the system out’ and only by doing that would they be successful at comprehending, producing, and judging Brocanto2. Though the bilinguals did improve in these abilities, especially for comprehension – their language production never reached high levels. The monolinguals had higher accuracy on this variable (though still not very high), but they also had more general experience in practicing additional languages in classroom learning contexts.

According to The Output Hypothesis (Swain, 1985, 1995, 2005) at least two of the central functions of producing an additional language are hypothesis formulation and testing and syntactic processing. In hypothesis formulation and testing, learners produce language as a way of testing new language forms (hypotheses) in conveying their message. The hypothesis, according to (Schachter, 1993), is a prediction about how certain aspects of the language are organized and involves finding regularities in the data and generalizing. Learners gather information about the accuracy of their hypothesis via external resources and can accordingly modify it depending on the feedback from such external resources. In The Output Hypothesis it is also suggested that producing language should lead learners to move from relying on semantic processing to using syntactic, structure-based processing. This is an important difference from
simply comprehending the language, in which a learner can circumvent structural information
and use semantic information to understand the message.

Considering these two potential functions of production in language learning contexts, it
could arguably be the case that the monolinguals in the Uninstructed condition were more
accustomed to testing *and modifying* their hypotheses about how an additional language structure
might work, and thus moved more easily from relying on semantic information to using syntactic
processes – reaching higher levels of productive ability in Brocanto2, and also greater structural
sensitivity in their ability to discriminate correct and violation language exemplars on the
grammaticality judgment tasks. The bilinguals, on the other hand, owing to their comparatively
smaller amount of experience producing language in a learning context, may have formed
hypotheses but not *modified* them as practice progressed.

During Brocanto2 practice, the learners received feedback on the correctness of their
response, which would be the external resource upon which learners were encouraged to either
modify their output or continue with a successful hypothesis. The Uninstructed monolinguals
appear to have better integrated this feedback and made some successful modifications in their
hypotheses/production – and such modifications may have led to their superior ability to identify
structures that were incongruent with their confirmed hypotheses on the grammaticality
judgment task (syntactic processing). The bilinguals never reached high levels of production
(50% accuracy at time series 4, compared to nearly 70% for the monolinguals), but did
eventually reach more than 80% accuracy on comprehension (time series 4). The relationship
between production and the development of syntactic processing skills on the one hand and
comprehension and reliance on semantic processing on the other may then explain why
bilinguals performed markedly lower than monolinguals on the grammaticality judgment tasks –
they may have moved to a much lesser extent than the monolinguals from semantic to syntactic processing.

Another function of the practice in Brocanto2 is also relevant to point out when considering the behavioral differences between the Uninstructed bilinguals and monolinguals. Independent of the extent of reliance on either semantic or syntactic processing for comprehending and producing the language as conceptualized by The Output Hypothesis, the practice in this study was designed to be task-essential (Loschky & Bley-Vroman, 1993), especially with respect to word order processing and production practice overall. This means that the learners were required to attend to Brocanto2 word order in order to accurately make a game move; and attend both word order and gender agreement in order to accurately describe a move. Previous studies on the provision of metalinguistic information during language learning have found evidence that such information may not be necessary when learners engage with practice that is task-essential (Sanz & Morgan-Short, 2004; Stafford, Bowden, & Sanz, 2012). Thus, in the absence of metalinguistic information, successful engagement with task-essential material should arguably be enough to promote positive learning outcomes.

By time series 4 in the present study (the last 4 blocks), the Uninstructed bilinguals were comprehending Brocanto2 at 84.4% accuracy and producing it at 50.5% accuracy. The Uninstructed monolinguals were comprehending Brocanto2 at 90.2% accuracy and producing it at 68.5% accuracy. Raw accuracy alone indicates that the bilinguals were less successful at the task-essential practice than the monolinguals. If it is the case that task-essential practice is enough to promote learning in the absence of metalinguistic information, but the bilinguals were not engaging as well as the monolinguals with such practice then this may explain why their learning as indexed by grammaticality judgment was significantly lower than that of the
monolinguals. They were not provided with metalinguistic information the way the Instructed bilinguals were and were not reaping the benefits of task-essential practice the way the Uninstructed monolinguals were. It is not unambiguously clear why the bilinguals were not engaging as successfully as the monolinguals with the task-essential practice, but at least one possible reason likely harkens back to the bilinguals’ general, and comparative, lack of foreign language learning experience and these or similar types of language pedagogy tools.

In sum, though the present study did not find evidence for bilingual advantages at additional language learning, at least as measured by behavioral assessments, there are important insights into the interactions between instructional context and the manifestation of such advantages. In an Instructed condition, where learners are provided with all of the language grammar they need, both bilinguals and monolinguals appeared to make successful use of this information. This may be related to a leveling of the playing field (Stafford, et al., 2010) in terms of the effects that early bilingualism would be predicted to have in additional language learning, though previous research on bilingual language learning advantages in classroom contexts like the Instructed one blur that explanation slightly. As such, it is reasonable that the greater amount of classroom foreign language learning experience in the monolinguals interacted in a non-negligible way with the traditional classroom learning context used here and resulted in similar performance between monolinguals, who were experienced in learning in such contexts, and bilinguals, who took appropriate advantage of the provided metalinguistic explanations.

In terms of the Uninstructed bilinguals and monolinguals – again no bilingual advantage was found and the monolinguals in fact outperformed the bilinguals. Though this result could be counter-intuitive, the experience of the monolinguals in learning foreign languages in classroom contexts may again explain this outcome. In the case that language production is important for
the development, testing, and modification of hypotheses about language function and leads to
the development of syntactic processing (Swain, 1985, 1995, 2005), the monolinguals – who
were more accurate in producing Brocanto2 - may have simply been more experienced in
making such modifications as they received feedback and thus more readily moved to reliance on
syntactic processing, which could be used to judge language exemplars. In the case that task-
essential practice should be enough to promote language learning in the absence of
metalinguistic information (Sanz & Morgan-Short, 2004; Stafford, et al., 2010), the bilinguals
were less experienced and less successful than the monolinguals at such practice in a language
learning context and concomitantly would not have equivalently realized the benefits of such
practice for their performance on grammaticality judgment tasks.

Overall, the results of the present study reinforce the notion that bilingual advantages at
language learning are unlikely to be absolute. And as with nearly all other research on
bilingualism and adult additional language learning, behavioral outcomes interact in complicated
ways with other variables, such as instructional context and perhaps even amount of classroom
language learning experience.

**Electrophysiological outcomes**

In Chapter 3 it was predicted that, on the presupposition of a bilingual advantage at
language learning, bilinguals would show more language-related ERP effects than monolinguals.
These effects were posited to take the form of more left-lateralized ANs and more robust P600s.
Additionally, it was predicted that bilingual advantages would manifest themselves more at high
as opposed to low experience and it was suggested that this would be more apparent in the
Uninstructed than Instructed condition based on previous work done with the experimental
design employed in the present study (Morgan-Short, Finger, et al., 2012; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012).

In the Instructed condition, processing of Brocanto2 word order violations at low experience led to an N400 in both bilinguals and monolinguals. The N400 found in both groups at low experience is least consistent with neurocognitive models of language which suggest that first and second language processing profiles should be similar in the absence of cross-language competition, even at low levels of proficiency (Hernandez, et al., 2005; Tokowicz & MacWhinney, 2005). The N400 found here is instead consistent with neurocognitive theories of adult language learning which posit that learners may rely on lexical/semantic processing for aspects of grammar (Clahsen & Felser, 2006a, 2006b, 2006c; Paradis, 2004, 2009), at the very least at low proficiency or experience in the additional language (Ullman, 2000, 2001b, 2005, 2013). To the author’s knowledge, this study provides the first set of evidence that such reliance on lexical/semantic processing for aspects of grammar might be involved in bilingual L3A in addition to monolingual L2A.

The results also revealed later anterior negativities in Instructed bilinguals and monolinguals at low proficiency. These later anterior negativities (here beginning 700-900 ms post-stimulus) have been reported in first and second language processing, though for higher levels of L2 proficiency (Friederici, et al., 1993; Martin-Loeches, et al., 2005; Morgan-Short, Steinhauer, et al., 2012; Mueller, et al., 2005), and have not been reported for learners at low levels of proficiency or experience. Though it has been suggested that such negativities may be continuations of the early, automatic processing indexed by earlier ANs (Steinhauer & Drury, 2011), the manifestation of these anterior negativities at low experience are potentially more consistent with suggestions that this neural pattern is related to increased working memory.
demands during sentence processing (Fiebach, et al., 2002; Martin-Loeches, et al., 2005).

Martin-Loeches et al. (2005), for example, manipulated their stimuli according to both presence of a grammatical in violation in short, simple sentences and also working memory load involved in processing subject-relative and object-relative embedded clauses. The authors reported that the anterior negativities elicited for grammatical violations were more left-lateralized and the negativities related to working memory load demonstrated a wider distribution. Additionally, they found that the negativities elicited under the working memory conditions displayed longer durations. The authors suggested that maintaining an active syntactic structure and the operations required for detecting a grammatical violation may pertain to the same system of language processing capacity (MacDonald & Christiansen, 2002), but qualitatively different subsystems. However, owing that later anterior negativities are still not theoretically or empirically well understood, the root of this effect throughout this discussion is speculative and future research will determine how later anterior negativities fit into the landscape of late-learned language processing.

A third effect found in the Instructed learners’ processing of word order violations was the P600 response at both low and high experience – though visual inspection of the waveforms and voltage maps indicates that this effect is more robust at high proficiency and begins earlier (500-700 ms post-stimulus) in the bilinguals than in the monolinguals. Even though these learners appear to have been employing lexical/semantic processing mechanisms earlier in processing (500-700 ms), later in the processing stream they were also engaging structure-based mechanisms to process the language violations, perhaps through structural reanalysis (Friederici & Weissenborn, 2007; Friederici, 2002), syntactic reprocessing (Osterhout, et al., 1994), or
syntactic integration processes (Kaan, et al., 2000), and were clearly employing these mechanisms in the absence of lexical/semantic processing (N400s) at high experience.

That the P600 in these Instructed learners appears more robust at high experience fits well with previous research in L2A which has found higher experience or proficiency to lead to P600s (e.g., (Osterhout, et al., 2008; Rossi, et al., 2006; Steinhauer, et al., 2009; Tanner, et al., 2012). However, considering that P600s have also been found at low proficiency or experience (e.g., (Tokowicz & MacWhinney, 2005; White, et al., 2012) it is not entirely striking that they are apparently also present (perhaps more so for the bilinguals than monolinguals) at low experience in Brocanto2. Additionally, the waveforms and voltage maps suggest that at high experience the bilinguals’ response to word order violations are more robust than that of the monolinguals, which may reflect superior, or more active, controlled structural processing mechanisms in this group.

At a theoretical level, the manifestation of P600s in the Instructed learners fits well with neurocognitive theories that posit that additional language processing of grammar can converge with that of first language processing (Bates & MacWhinney, 1989; Green, 2003, 2005; MacWhinney, 2001), especially at higher levels of proficiency or experience (e.g., Ullman, 2001b, 2005). This result additionally suggests that adult additional language learners may not be confined to shallow processing mechanisms (Clahsen & Felser, 2006a, 2006b, 2006c). That no statistical differences were found between bilinguals and monolinguals for the ERPs indicates that the neurocognitive mechanisms employed by the two groups are similar – but visual analysis of the effects suggests additionally that, at least for the P600 at high experience, the structural processing mechanisms may be more strongly activated in bilinguals as compared to
monolinguals. This, in turn, could be one dimension of a bilingual advantage in the neurocognition of late-learned language.

Finally, at both at low and high experience violation of word order led to an early anterior-central positivity (150-300 ms post-stimulus) in Instructed bilinguals and monolinguals. The early positivity, while not traditionally found in language processing studies (but see (Morgan-Short, Finger, et al., 2012; Morgan-Short, Steinhauer, et al., 2012; Mueller, Oberecker, & Friederici, 2009), has been posited to index the execution of attentional mechanisms, and perhaps specifically those that are related to context-updating as information comes in (Donchin, 1981; Polich, 2007, 2012). According to the Context Updating Theory (Donchin, 1981) of this positivity – termed the P300, attention-driven mechanisms evaluate the sensory input in comparison to previous information/events that are stored in working memory. If there is no change, or difference, the context is maintained and no P300 emerges. However, if a new stimulus characteristic is detected, the information context is updated, resulting in a P300.

Given that the Instructed learners in this study were provided with the very grammar information that was violated during the grammaticality judgment task, their context-updating attentional mechanisms could have been more sensitive to the violations – which would have been new stimulus characteristics. In this way, online mechanistic comparisons between the explicitly provided information and incoming illicit/different information may have been enhanced. The emergence of this positivity in both bilinguals and monolinguals may, therefore, be at least partly a result of the instructional context under which they learned Brocanto2- which likely directed their attention to the grammatical forms (Leow & Bowles, 2005; Sanz & Morgan-Short, 2005). However, an early anterior positivity was also evidenced in the Uninstructed bilinguals at high experience for word order, so the effect is unlikely to be exclusively related to
the instructional context (see below for further discussion on this effect for the Uninstructed bilinguals).

Instructed bilinguals and monolinguals evidenced somewhat distinct patterns of processing for word order than for gender agreement. For the processing of gender agreement, neither group showed early neural responses to violations, though this may be due to a lack of power given the small subject groups as opposed to a true lack of ERP effects for the processing of this structure. Nonetheless, there was no evidence of the early anterior-central positivity found for word order which indicates that attentional mechanisms were not as heavily recruited for gender agreement processing. This could be function of differences in complexity or saliency (DeKeyser, 2005; Doughty & Williams, 1998; Hulstijn & De Graaff, 1994), wherein word order may be less complex and more salient than gender agreement and therefore more readily detected by attention-based resources.

At low experience in the final time-window (900-1200 ms) the bilinguals and monolinguals exhibited distinct electrophysiological responses following violations in gender agreement. The monolinguals showed a P600 whereas the bilinguals showed a late anterior negativity, which, again, is not routinely found at low levels of proficiency or experience, but was also found for word order above. This may indicate that the bilinguals’ working memory was more engaged for this structure (Fiebach, et al., 2002; Martin-Loeches, et al., 2005) whereas the monolinguals may have been relying on controlled, structural-based processes (Friederici & Weissenborn, 2007; Friederici, 2002; Kaan, et al., 2000; Osterhout, et al., 1994).

The results indicated additionally that bilinguals and monolinguals showed later anterior negativities beginning in the 700-900 ms time-window together with a P600, both of which continued to the final time-window of interest. Though these effects were ‘found’ at both low
and high experience, an effect of Time and perhaps even Bilingualism may be obscured by low power – visualization of the ERP effects indicates that the anterior negativity is stronger at high experience and in the bilinguals, and the P600 is more salient in the monolinguals. (Note however that the extension of the late anterior negativity in the bilinguals may be over-powering the positivity somewhat). Previous work on additional language processing (monolingual L2A) of nominal gender agreement has found P600s at low (Tokowicz & MacWhinney, 2005) and also high proficiency or experience (Gillon Dowens, et al., 2011; Gillon Dowens, et al., 2009). And like the results for word order, these findings are consistent with neurocognitive theories of late-learned language that predict similarities between first and additional language for processing of some aspects of grammar, perhaps as a function of proficiency or experience (Green, 2003, 2005; Ullman, 2001b, 2005). The later anterior negativities found in Instructed bilinguals and monolinguals for the processing of gender agreement may, like word order, reflect increased working memory load involved in sentence processing demands (Fiebach, et al., 2002; Martin-Loeches, et al., 2005), which appear to be more active at high as compared to low experience in the language. However, again this explanation is speculative.

In terms of differences between bilinguals and monolinguals in processing of gender agreement violations – the two groups are most obviously different at low experience, with the bilinguals showing more negative-going neural activity in general and the monolinguals showing more positive-going activity. This undoubtedly reflects the recruitment of different neural mechanisms in bilinguals and monolinguals at low experience. Given the ambiguity of what later anterior negativities represent in the neurocognition of language and the absence of early ERP effects in the groups, perhaps the only conclusion that can be made is that the monolinguals recruited structure-based processing more so than the bilinguals whereas the bilinguals recruited,
potentially, more working memory processes. At high experience, the two groups are more similar to each other at these later stages of processing though it appears that the bilinguals’ late anterior negativity is more robust than the monolinguals, especially in the final time-window.

In sum, for the neurocognition of both word order and gender agreement, bilingual L3A and monolingual L2A do not appear to be vastly different from each other when learners are trained under a condition for which metalinguistic information is provided, especially for word order processing and at high experience. However, the study does provide tentative evidence that bilinguals recruited processing mechanisms more strongly at high experience than the monolinguals – the P600 at high for word order and the late anterior negativity at high for gender agreement. And at low experience for gender agreement the bilinguals appeared to have recruited quite distinct processing mechanisms, perhaps related to processing within a higher working memory load. Therefore, the prediction that bilinguals would show more robust ERP effects than monolinguals at high as compared to low experience has been speculatively borne out in the present study and is suggestive of a bilingual advantage in the neurocognition of late-learned language.

The electrophysiological data for the Uninstructed bilinguals is somewhat more difficult to reliably interpret in the context of the paucity of effects that reached significance in the neural outcomes for these learners. This was likely due to low power - with just 7 bilinguals, as compared to the Instructed bilinguals who numbered 9 subjects. Nonetheless, at least three important findings emerged from the processing of word order for these Uninstructed learners.

First, like the Instructed learners, the Uninstructed monolinguals evidenced an N400 at low experience in response to word order violations. This effect, however, appeared earlier in Uninstructed monolinguals (300-500 ms post stimulus) than both groups of Instructed learners.
(500-700 ms post-stimulus). Also like the findings in the Instructed groups, the emergence of an N400 for the processing of word order at low experience is consistent with reliance on lexical/semantic processing for some aspects of grammar (Clahsen & Felser, 2006a, 2006b, 2006c; Paradis, 2004, 2009), particularly at low experience (Ullman, 2000, 2001b, 2005, Ullman, 2013). That the N400 appeared in an earlier time-window for the Uninstructed monolinguals indicates that the neural mechanisms involved in the processing of these violations may become active sooner in a condition for which metalinguistic information was not provided. In the context of the N400, then, the Uninstructed monolinguals in particular may be activating lexical/semantic processes to address the detection of a language violation more quickly than Instructed learners - though since these groups were not directly compared this conclusion should be evaluated cautiously.

Second, that a negativity was found for both Uninstructed bilinguals and monolinguals at high experience is taken to be indexing different processes in the two groups. For the monolinguals, the negativity is consistent with an AN, reported for this group in the full set of subjects and demonstrating a more frontal distribution than their negativity at low (Morgan-Short, Finger, et al., 2012; Morgan-Short, Steinhauer, et al., 2012). In the bilinguals, on the other hand, the distribution of the negativity suggests it is more consistent with an N400 for these learners. Given the distributional differences, the Uninstructed monolinguals at high experience appear to be activating early automatic language processing mechanisms (Friederici & Weissenborn, 2007; Friederici, 2002, 2004; Friederici & Kotz, 2003; Molinaro, et al., 2011). This is consistent with predictions made by the DP model of language which suggests that, insofar as N400s index lexical/semantic processing in declarative memory and ANs index grammatical processing in procedural memory - learners may undergo a declarative-to-
procedural memory shift in the processing of some aspects of grammar with enough experience or proficiency in the language (e.g., Ullman, 2000, 2001b, 2005).

The N400 evidenced at high in the bilinguals indicates that, even at high experience, they are relying on lexical/semantic processes (Clahsen & Felser, 2006a, 2006b, 2006c; Paradis, 2004, 2009). This difference between Uninstructed bilinguals and monolinguals could relate to the shift from semantic to syntactic processing which, at least in the behavioral outcomes, appeared to be more evident in monolinguals than bilinguals (see above Swain, 1985, 1995, 2005). Also, note that though the Instructed bilinguals were at ‘high experience’ in terms of having completed all practice blocks, they were still arguably not highly proficient in Brocanto2 – and definitely less so than the monolinguals. The Uninstructed bilinguals and monolinguals had exactly the same amount of experience, but were significantly different from each other in their performance. This tentatively suggests that level of proficiency/performance carries more weight than sheer amount of experience in leading learners to move from lexical/semantic based processing to more automatic, structure-based processing of an additional language, at least under a condition characterized by an abundance of meaningful exemplars and no metalinguistic information.

Finally, at low experience in the final time-window, the Uninstructed learners exhibited distinct electrophysiological responses for word order. Recall that Instructed bilinguals and monolinguals were also different from each for low experience at this later stage of processing, though for gender agreement, not word order. Additionally, the pattern of activity is reversed from that of the Instructed groups’ differences – the Uninstructed bilinguals exhibited more positive-going activity later in word order processing whereas the monolinguals exhibited more negative-going activity. This finding points not only to a difference in the effects of instructional
context on neurocognition of late-learned language for word order and gender agreement (Morgan-Short, Finger, et al., 2012; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012), but reveals that neural mechanisms involved in processing are differentially recruited in bilinguals and monolinguals, at least at low experience and especially as a function of the grammatical target.

The late anterior positivity is considered to be indexing a P600 in the bilinguals. Though P600s generally exhibit a posterior scalp distribution, they have also been reported with anterior distributions (Friederici, et al., 2002; Hagoort, et al., 1999; Kaan & Swaab, 2003). The anterior negativity in the monolinguals at this stage of processing, like those in the Instructed groups, may be reflecting increased working memory processing for these learners (Fiebach, et al., 2002; Martin-Loeches, et al., 2005), though it may alternatively be the case that these negativities reflect continual neural activity from early structure-based processing (Steinhauer & Drury, 2011).

The early anterior positivity in the Uninstructed bilinguals at high experience should also be briefly considered even though this effect did not reach statistical significance (but did approach significance). An early positivity was also found for word order in the Instructed bilinguals and monolinguals. On the premise that this effect is akin to the P300 and tied to attentional mechanisms (Donchin, 1981; Polich, 2007, 2012) it is notable that the Uninstructed bilinguals in this study appeared to be engaging these mechanisms similarly to the Instructed learners, whereas the monolinguals did not. As mentioned in Chapter 1, research on the non-linguistic benefits of bilingualism suggests that one dimension of an advantage in cognition may be tied to attention-related processes or processing (Bialystok, et al., 2004; Costa, et al., 2008; Feng, et al., 2009; Stafford, 2011, but note Paap & Greenberg, 2013). Viewing the early anterior
positivity in light of attentional mechanisms suggests that the Uninstructed bilinguals’ attentional processing of word order violations in the present study may have been enhanced similarly to that of the Instructed groups’, and even when not having been trained under a condition containing attention-favored metalinguistic information. It could be the case, then, that even in the absence of attention-directing instruction, bilinguals more so than monolinguals engaged neural mechanisms associated with increased attentional processing when encountering language violations, at least for word order (and perhaps even for gender agreement, see Figures 11 and 12). Though it is premature to posit that the Uninstructed bilinguals’ attentional processing of the language was enhanced and that this may relate to superior attention-related processes or processing more generally, future research should explore this possibility as it relates to domain-general cognitive effects of bilingualism on the neurocognition of bilingual L3A.

The ERP effects for processing of gender agreement in the Uninstructed learners produced the smallest amount of significant results. Though there were intriguing interactions at the highest level of the ANOVAs, nearly all of them failed to step down to significant effects. This lack of effects should not be interpreted as a true result for gender agreement processing in these groups, however. Rather, the lack of significant effects are almost certainly due to a lower number of subjects in the Uninstructed group, specifically the Uninstructed bilinguals. Therefore, the absence of effects here should be evaluated with considerable caution.

The only ERP effect to reach statistical significance in this group of learners was the P600 in the final time-window of interest, at high proficiency and for both bilinguals and monolinguals. The manifestation of a P600 echoes the above results in terms of neurocognitive theories of late-learned (e.g., Green, 2003, 2005; Ullman, 2001b, 2005). It also fits well with previous research that has found this result for the processing of gender agreement in late
language learners (e.g., Gillon Dowens, et al., 2011; Gillon Dowens, et al., 2009). Like the P600 in the Instructed bilinguals at high experience for processing of word order, the P600 in the Uninstructed bilinguals here for gender agreement appears to be more robust than that of the Uninstructed monolinguals. This would suggest, again, that the bilinguals may be engaging structural processing mechanisms more strongly than monolinguals and points to a possible instantiation of a bilingual advantage in the neurocognition of late-learned language that is deserving of future research.

To summarize, in an instructional context that is characterized by an abundance of meaningful input and no metalinguistic information, bilinguals and monolinguals appear to be less similar to each other than when provided with metalinguistic information. For word order processing, the Uninstructed monolinguals appeared to be engaging language-related lexical/semantic processing at low experience and early, automatic structural processing at high experience. The bilinguals, on the other hand, may have still been relying on lexical/semantic mechanisms at high experience. Additionally, during later stages of processing at low experience, the bilinguals appeared to be recruiting controlled structure-based processing as compared to the later anterior negativity in the monolinguals, which may reflect working memory demands. For the processing of gender agreement at high experience is when the bilinguals and monolinguals showed the most similar neural effect (P600), and this seems to be more robust in bilinguals than monolinguals. Therefore, the prediction that bilinguals would show more robust ERP effects than monolinguals at high as compared to low experience has been suggestively realized for the processing of gender agreement, but is much less tenable for the processing of word order.
In Chapter 3, it was predicted that bilingual advantages would manifest themselves more in the Uninstructed than Instructed learners in terms of ERP effects. It is difficult to reliably evaluate the outcomes of this study in the context of that prediction. Full comparisons were not made across the four groups due to the low likelihood of finding any significant results once four (small) groups’ electrophysiological data was entered into a single ANOVA with an array of other factors (i.e., time, laterality, hemisphere, etc.). Therefore, comparisons between the ERP effects found for Instructed as compared to Uninstructed learners in this study may not be a valid undertaking for discussion. However, in terms of general observational differences between the neural effects found in this study under Instructed and Uninstructed contexts there are nonetheless potentially important points to note.

First, an Instructed condition containing metalinguistic information appeared to engage attentional mechanisms across both bilinguals and monolinguals more so than an Uninstructed condition, at least for a structurally salient target such as word order. This is not wholly surprising given that the Instructed learners’ attention was arguably directed to the very forms that were violated in the grammaticality judgment task (Leow & Bowles, 2005; Sanz & Morgan-Short, 2005). Notably, even in the absence of an attention-directing context the Uninstructed bilinguals also appeared to be engaging similar attentional mechanisms, which could be interpreted as being related to superior attention-based processing in bilinguals more generally (Bialystok, et al., 2004; Costa, et al., 2008; Feng, et al., 2009; Stafford, 2011). This interpretation should be viewed cautiously, however, since the effect only approached significance and no measures of attention are reported here.

Second, one possible interpretation of the late anterior negativities at low and high experience in Instructed learners for the processing of gender agreement and at low experience
for word order is that these learners were engaging working memory resources to a greater extent than Uninstructed learners. This may also be at least partly an artifact of the instructional context that provided explicit, metalinguistic information. Working memory has been found to be a variable that often interacts importantly with language learning and processing (see, for instance, Baddeley, 2003; Bowden, et al., 2005; Goo, 2010; Juffs & Harrington, 2011; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Tagarelli, Borges-Mota, & Rebuschat, 2011) and is needed for the initial encoding and ultimate recall of explicit knowledge (Frackowiak, et al., 2004; Kandel, Schwartz, & Jessel, 2000). Therefore, one potential important difference between the effects of Instructed and Uninstructed conditions on processing could be that working memory is more heavily recruited to access the explicit information that learners were provided with during training. Finding these negativities at low and high for gender agreement processing might, then, suggest that more complex/less salient structures persist in demanding working memory resources during sentence processing in late-learners and that less complex structures such as word order demand such resources less (or not at all in the bilinguals) as experience in the language increases. However, this explanation is only tenable if these later anterior negativities are indeed indexing working memory demands and not the continuation of an earlier, automatic structure-based process more specific to the detection of grammatical violations (Steinhauer & Drury, 2011). Future studies should explore this possibility in order to better elucidate not only the mechanisms underlying these negativities, but also how they interact with the effects of instruction on the neurocognition of late-learned language.

13 Note that, though the later anterior negativities in the Uninstructed monolinguals at high for word order were not statistically confirmed here (but see Morgan-Short, Finger, et al., 2012; Morgan-Short, Steinhauer, et al., 2012) the potential execution of working memory resources for that group may similarly indicate some interaction with recalling explicit knowledge, knowledge which the learners may have arrived at on their own during exposure or practice.
Finally, there is evidence to suggest that monolingual L2A processing of word order is benefited more by an instructional context for which ample meaningful examples are provided as opposed to meaningful examples in addition to metalinguistic information. This is true insofar as only the Uninstructed monolinguals evidenced early automatic neural responses to language violations (the AN for word order at high) and thereby may be exhibiting a declarative-to-procedural shift in their neurocognition of Brocanto2 for this target (Ullman, 2000, 2001b, 2005, 2013; see also, Opitz & Friederici, 2003). Neither the Uninstructed bilinguals nor the Instructed bilinguals and monolinguals demonstrated this effect. Both Instructed and Uninstructed conditions, however, seem to be more or less effective in promoting later, controlled processing as indexed by P600s despite differences in the timing, extent, and intensity of these effects across bilinguals and monolinguals and for the two grammatical targets\(^ {14} \).

From a purely observational perspective it does not appear to be the case that bilingual advantages at additional language learning are more obvious in an Uninstructed context. It in fact appears to be the case that such advantages are clearer in the Instructed context, where bilinguals appeared to be exhibiting earlier and stronger P600s at high for word order and stronger late anterior negativities for gender agreement. Nonetheless, the P600 in the final time-window for Uninstructed bilinguals’ processing of gender agreement also seems to be stronger than that of the Uninstructed monolinguals. It therefore may be that the neural mechanisms that lead to P600s in the neurocognition of language are more actively, or more strongly, recruited in bilinguals than monolinguals. However, future studies with larger numbers of subjects and that are perhaps exclusively focused on studying the P600 in bilingual L3A and monolingual L3A are needed in order to test the validity of this conclusion.

\(^ {14} \) An obvious caveat here is that no significant P600 was found in the Uninstructed learners for word order, though this is probably related to power issues and not a true absence of such an effect, especially in the monolinguals.
Limitations and Future Research

This study is the first, to the author’s knowledge, to compare bilingual L3A and monolingual L2A using both behavioral and neurocognitive measures, and across different instructional contexts and time points - and thus importantly contributes to the body of knowledge on bilingualism, the effects of instruction, and the neurocognition of late-learned language. However, there are obvious limitations that affect the study’s reliability and generalizability.

The first and clearest limitation of the present study is the low number of subjects involved in the analyses, particularly with respect to the ERP effects. Though 24 bilingual subjects were partially or fully tested in this study (see Chapter 3), only 16 subjects (9 Instructed, 7 Uninstructed) were included in the analysis. This small number of subjects may be evidencing effects that do not persist once a larger number of subjects are included in the analysis. As such, the results of the study should be interpreted more as preliminary information on the status of bilingual advantages in the neurocognition of late-learned language as opposed to conclusive evidence one way or the other. This is especially true considering that the full sets of Instructed and Uninstructed monolingual subjects have produced somewhat different findings than those reported here, especially for the Instructed monolinguals (Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012). Future research that includes larger subject groups is required in order to confirm and extend the results from the present study, particularly as they relate to bilingual L3A. Larger subject groups would additionally enable researchers to more reliably investigate interactions with instructional context and the realization of behavioral or neurocognitive differences in bilingual L3A and monolingual L2A, which was not possible here.
A second limitation of the present study is that independent measures of proficiency for the bilinguals’ two languages were not gathered. There is respectable evidence and research precedent to support the decision to use self-reports (Bachman & Palmer, 1985; Bartolotti & Marian, 2012; Bartolotti, et al., 2011; Bialystok, 2006; Bialystok, et al., 2006; Chincotta & Underwood, 1998; Dunn & Fox Tree, 2009; Fernandes, et al., 2007; Flege, et al., 2002; Flege, et al., 1999; Gillon Dowens, et al., 2009; Jia, et al., 2002; Kaushanskaya & Marian, 2007; Kaushanskaya & Marian, 2009a,b; MacIntyre, et al., 1997; Marian, et al., 2007; Ross, 1998; Shameem, 1998; Stafford, 2011; Stefani, 1994) and the bilinguals did indeed rate themselves as consistently high for both languages across several dimensions. However, future research on bilingual L3A should try, to the extent possible, to include not only self-reporting of proficiency in the languages, but also at least one independent measure of the proficiency. In this way, researchers will be able to extend empirical faith in the reliability and validity of self-reporting, and also address any peripheral concerns about how bilingual proficiency in the two languages might interact with late-learned language outcomes, particular as they relate to potential bilingual advantages therein.

A third limitation of the present study concerns its use of an artificial language. There is, like the use of self-reporting for proficiency, a considerable body of research that has used artificial or semi-artificial languages in order address pertinent questions about adult additional language acquisition (e.g., DeKeyser, 1996, 1997; Grey, et al., Forthcoming 2014; Hama & Leow, 2010; Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter, & Wong, 2013; Rebuschat & Williams, 2006, 2012; Tagarelli, et al., 2011; Williams, 2005; Williams & Kuribara, 2008, not to mention Morgan-Short, Finger, et al., 2012; Morgan-Short, et al., 2010; Morgan-Short, Steinhauer, et al., 2012). However, the highly controlled laboratory setting and
artificial linguistic system used here nonetheless provide only a limited view of how bilingual L3A and monolingual L2A might compare to each other in more complex, natural language learning settings. Therefore, though the use of the artificial language here undoubtedly provides important preliminary information about (a) bilingual L3A as compared to monolingual L2A, (b) influences of instructional context, and (c) the neurocognitive effects of these variables, the present study is but a stepping stone and future research that employs natural languages, or perhaps miniature languages (e.g., Lado, 2008; Lado, et al., 2013; Lin, 2009; Mueller, et al., 2005; Sanz, Lin, Lado, Bowden, & Stafford, 2009; Stafford, et al., 2012; Stafford, et al., 2010), will be required in order to more fully inform our understanding of the neurocognition of late-learned language in bilinguals and monolinguals.

Finally, the study is limited in that the bilingual and monolingual groups were not equated with each other across all variables of information gathered. The goal of the study was to isolate bilingualism and its effects on additional language learning and processing. However, independent of the early, lifelong bilingualism targeted here, there was an additional variable that may have influenced the results of the current study – particularly the behavioral outcomes. The Instructed (and also suggestively the Uninstructed) monolinguals in the present study had significantly more additional language learning experience in classroom settings than the bilinguals. Given that both of the two instructional conditions used in the present study approximated types of foreign language classrooms, and that bilinguals had less experience in these contexts, it is not improbable that the monolinguals’ added experience influenced the outcomes of the study to some extent, at the very least for the behavioral assessments of learning. However, this suggestion would need to be investigated by studies that target this variable more focally in research on bilingual L3A compared to monolingual L2A.
As a final note on directions for future research, perhaps the most salient ERP effect (visualization of waveforms and voltage maps) of the present study was the strength, and also timing for Instructed learners, of the P600 effect in bilinguals as compared to monolinguals. Insofar as the P600 is a neural signature of later, controlled structural processing of language violations (e.g., Friederici, 2002; Kaan, 2007; Kaan, et al., 2000; Kotz, 2009; Molinaro, et al., 2011; Osterhout & Mobley, 1995), it might be the case that bilinguals engage more heavily in such types of processing than monolinguals. Given that bilinguals have been found to have advantages in domain-general control-relevant cognition\(^{15}\) (e.g., Bartolotti & Marian, 2012; Bartolotti, et al., 2011; Bialystok, 1999; Bialystok, et al., 2008; Bialystok, et al., 2004; Bialystok, Craik, & Ryan, 2006; Bialystok & Feng, 2009), this could be an interesting language-related dimension of such control-related processing. Though this is a highly preliminary suggestion about the dimensions of a neurocognitive advantage in bilingual late-learned language processing, it warrants further empirical investigation\(^{16}\).

Conclusion

This study sought to investigate whether there are bilingual advantages at adult additional language learning. To achieve that, the study aimed to address limitations in previous research on this topic concerning, for example, onset, amount, and type of exposure, and typological similarity, so as to more reliably isolate the effects of bilingualism on language learning

\(^{15}\) The P600 may also be more generally an index of processing in declarative memory which may be enhanced in bilinguals (Michael Ullman, \textit{p.c}).

\(^{16}\) The later anterior negativities, which appeared to be stronger for bilingual processing of gender agreement under an Instructed condition, also warrant further investigation, not least as contributions towards determining what such negativities represent in language processing theory and research.
outcomes. As such, it employed a laboratory-based paradigm with an artificial language that was not similar to the bilinguals or monolinguals early known language(s) and that they would have had no prior experience in – thereby equating learners on these and other previously uncontrolled-for variables. The study additionally aimed to provide a multi-dimensional perspective on potential bilingual advantages by measuring both behavioral and neural outcomes, at different points in the learning trajectory, under different instructional contexts, and for different linguistic targets. The findings of the study broadly indicate the following.

First, the behavioral results indicate that bilinguals and monolinguals do not perform differently, either on practice or on judging correct and violation structures, when they are exposed to an additional language in a setting characterized by metalinguistic information and meaningful examples. In this condition no bilingual advantages were evidenced. This may be at least partly due to a leveling effect of the Instructed context (Stafford, et al., 2010). Nor were bilingual advantages evidenced in an instructional context characterized by an abundance of meaningful information. The behavioral results for bilinguals and monolingual trained under an Uninstructed condition showed that monolinguals in fact outperformed bilinguals. This may have been at least partly related to differences in the way the two groups capitalized on the opportunity to practice the language, and the function of such practice in promoting language development (Sanz & Morgan-Short, 2004; Stafford, et al., 2012; Swain, 1985, 1995, 2005). Importantly, these behavioral results reinforce the fact that instructional context is an active and important variable not only in monolingual L2A (Norris & Ortega, 2000; Spada & Tomita, 2010), but also bilingual L3A (Lado, 2008; Lado, et al., 2013; Lin, 2009).

The electrophysiological data in the current study also points to intriguing effects of instructional context on bilingual L3A as compared to monolingual L2A. Given the P600s
evidenced here, especially at high experience, the present study supports neurocognitive theories of language which posit that additional language mechanisms underlying processing of grammar can approximate those of a first language (Green, 2003, 2005; Ullman, 2000, 2001b, 2005; Ullman, 2013). The study provides perhaps the first set of evidence that bilingual L3A is also relevant for such theories and tentatively indicates that the P600 effect in particular may be at least one dimension of a bilingual advantage in the neurocognition of late-learned language. However, the study also highlights differences in the neurocognition of additional language as a function of experience, instructional context, and grammatical target. Specifically, Uninstructed monolinguals appeared to have undergone a neurocognitive shift in their processing of word order – consistent with the Declarative/procedural model of language (Ullman, 2000, 2001b, 2005, 2013) - whereas no other group evidenced such a shift and the Uninstructed bilinguals may have still been relying on lexical/semantic mechanisms at the same level of experience (Clahsen & Felser, 2006a, 2006b, 2006c; Paradis, 2004, 2009).

Additionally, at low experience in general is when bilinguals and monolinguals appeared to be most different from each other, at least for word order in the Uninstructed group and gender agreement in the Instructed group. This indicates that especially at earlier stages of language learning, bilinguals and monolinguals might recruit different neural mechanisms during processing. At high experience, on the other hand, more similar neural mechanisms appeared to be at play in bilinguals and monolinguals (Instructed, both targets; Uninstructed gender agreement) – though perhaps to different degrees of intensity. It has been suggested that the magnitude of the ERP effects in the bilinguals (P600 at high for Instructed word order and Uninstructed gender agreement, and late anterior negativity at high for Instructed gender agreement) may an instantiation of a bilingual advantage in the neurocognition of late-learned
language, though this suggestion will need to be empirically verified in replications and extensions of the current study with larger subject groups. Nonetheless, this is the first study to compare bilingual L3A and monolingual L2A on neurocognitive measures of processing at different points in their language experience and under different instructional contexts, and suggests interesting differences and also similarities as a function of those two variables. In this way, the study is an important first step in neurocognitive research on the multi-faceted and complex nature of bilingual L3A as it compares to monolingual L2A. Future research that also takes into account the trajectory of learning and effects of instructional context on the neurocognition of bilingual L3A will extend the study’s present results and broaden our knowledge on the extent and characterization of bilingual advantages at additional language learning.
Appendix A

Dunn and FoxTree’s (2009) Bilingual Dominance Scale and Scoring – adapted for Mandarin-English bilinguals

Questions 1 and 2:
At which age did you first learn Mandarin? _______
At which age did you first learn English? _______

Questions 3 and 4:
At what age did you feel comfortable speaking Mandarin? _______
At what age did you feel comfortable speaking English? _______

Question 5:
Which language do you predominantly use at home?
Mandarin _______
English _______
Both _______

Question 6:
When doing math in your head (such as multiplying 243 x 5), which language do you calculate numbers in?
__________________________

Question 7:
If you have a foreign accent, which language(s) is it in?
__________________________

Question 8:
If you had to choose which language to use for the rest of your life, which language would it be?
__________________________

Questions 9 & 10:
How many years of schooling (primary school through university) did you have in:
Mandarin ____________
Question 11:
Do you feel that you have lost any fluency in a particular language?

If yes, which one? At what age?

Question 12:
What country/region do you currently live in?
Scoring for the Bilingual Dominance Scale

Questions 1 and 2:
0-5 years = +5; 6-9 years = +3; 10-15 years = +1; 16 and up = +0

Questions 3 and 4:
0-5 years = +5; 6-9 years = +3; 10-15 years = +1; 16 and up = +0; “not yet” = +0

Question 5:
If one language is used at home, +5 for that language; if both used at home, +3 for each language

Question 6:
+3 for language used for math; +0 if both

Question 7:
If one language is listed, add +5 to the opposite language of the one listed; if both languages are listed, add +3 to both languages; if no language is listed, add nothing

Question 8:
+2 for language chosen for retention

Questions 9 and 10:
1-6 years = +1; 7 and more years = +2

Question 11:
-3 in language with fluency loss; -0 if neither has lost fluency

Question 12:
+4 for predominant language of country/region region of residence
Appendix B

Uninstructed Training Script

(Directions)

For the second part of the study you will receive training on an artificial language. Listen carefully to the information that is given. You will hear examples of words, phrases and sentences from this artificial language and will see both still and moving images on a game board that correspond to those words, phrases and sentences. Immediately after, you will be asked some questions about what you have seen.

Later, in the third part of this study, you will use the artificial language to play the game that was explained to you previously.

This second part of the study will last approximately 20 minutes and is timed by the computer. Please be aware that you will not be able to go back and review any of the information or examples. Click on the ‘Continue’ button when you are ready to start the second part of the study.

(Training)

Introduction to the artificial language

We will begin learning Brocanto2 by learning some of its vocabulary. Brocanto2 has a vocabulary of 16 words, which is comprised of nouns, articles, adjectives, verbs and adverbs. We will now learn the nouns, articles and adjectives.

Nouns

- Nouns. There are four nouns in Brocanto2. Each noun represents a different object. Each noun in Brocanto2 also has a gender. Thus, along with the meaning of the noun, you need to memorize whether each noun is ‘masculine’ or ‘feminine’. You will now hear the four nouns and see the objects they represent.
  - Pleck is a masculine noun
  - Neep is a masculine noun
  - Blom is a feminine noun
  - Vode is a feminine noun

- Let’s review this information again.
- **Pleck** is a masculine noun
- **Neep** is a masculine noun
- **Blom** is a feminine noun
- **Vode** is a feminine noun

### Articles

- Articles. Articles are used with nouns and specify that you are referring to one object in particular. There is only one article in Brocanto2, which has two forms: a masculine form and a feminine form.
  - **Li** is the masculine form
  - **Lu** is the feminine form

You should remember two points about articles in Brocanto2. First, articles always come after nouns. Second, articles must ‘agree’ with the gender of the noun. In other words, if a noun is masculine, the masculine form of the article must be used. Likewise, if a noun is feminine, the feminine form of the article must be used. Thus, whenever a noun is masculine, you will need to use the masculine article – **Li**, as in the following example:

  - **pleck li**

In this example, **Pleck** is a masculine noun, so you must use the masculine form of the article: **li**.

On the other hand, if a noun is feminine, you will need to use the feminine article – **Lu**, as in the following example:

  - **blom lu**

In this example, **Blom** is a feminine noun, so you must use the feminine form of the article: **lu**.

Here is another example. Think about whether the noun is masculine or feminine. Notice that the article is placed after the noun.

  - **vode lu**

### Adjectives

- Adjectives. Brocanto2 has two adjectives that describe the shape of the objects named by the nouns. Each adjective has two forms: masculine and feminine.
  - **Troise, troiso**
    - Both these adjectives represent the symbol you see on your screen. **Troise** is the masculine form and **troiso** is the feminine form.
  - **Neime, neimo**
Both of these adjectives represent the symbol you now see on your screen. **Neime** is the masculine form and **neimo** is the feminine form.

Note that the masculine form of both adjectives end with the sound –e, while the feminine form of both adjectives end with the sound –o.

Like with articles, you will need to remember two things about adjectives in Brocanto2. First, they are always placed immediately after the noun they are describing. This means that they will be placed between the noun and the article. Second, adjectives must also agree with the gender of the noun they are describing. Listen to the following example:

- **neep troise li**
  
  In this example, since **neep** is a masculine word, you hear the form of the adjective that ends in –e.

- **blom troiso lu**
  
  In this example, since **blom** is a feminine word, you hear the form of the adjective that ends in –o.

- **neep neime li**
  
  In this example, since **neep** is a masculine word, you hear the form of the adjective that ends in –e.

- **vode neimo lu**
  
  In this example, since **vode** is a feminine word, you hear the form of the adjective that ends in –o.

In sum, you need to remember that in Brocanto2, articles and adjectives are placed after nouns and must agree with the gender of the noun. If the noun is masculine, you must use the masculine form of the article and adjective. If the noun is feminine, you must use the feminine form of the article and adjective. Listen to two more examples of how nouns, adjectives and articles work together in Brocanto2:

- **blom neimo lu**
- **vode troiso lu**

We will now learn about Brocanto2 verbs and adverbs.

**Verbs**

- Verbs. There are four verbs in Brocanto2. You will now hear the four verbs and see the actions they represent.

  - **Klin**
Let’s review the verbs again.

- Klin
- Nim
- Yab
- Praz

Now that you know Brocanto2 verbs, we can start forming sentences. Some sentences in Brocanto2 only have a subject and a verb. These are called intransitive sentences. In these sentences the subject performs the action of the verb. Only three Brocanto2 verbs can be used in intransitive sentences. These verbs are *klin*, *yab*, and *nim*. Listen to some examples:

- Vode lu klin
- Neep li yab
- Neep li klin
- Blom lu yab
- Vode lu nim

Adjectives can be placed in these sentences to describe the noun. Remember, adjectives are placed immediately after the noun they describe and the article then follows the adjective. Here are some examples of sentence with adjectives:

- Neep troise li nim
- Blom neimo lu klin
- Blom neimo lu yab

Other sentences in Brocanto2 have a subject, an object and a verb. These are called transitive sentences. Only three Brocanto2 verbs can be used in transitive sentences. These verbs are *praz*, *yab*, and *nim*. Notice that *yab* and *nim* can be used in both intransitive and transitive sentences. *Klin* can be used only in intransitive sentences and *praz* can be used only in transitive sentences.

In transitive sentences, the subject performs the action of the verb and the object receives the action of the verb. In Brocanto2, both the subject and the object are placed before the verb. The subject occurs first and the object occurs second.
Thus, the word order for Brocanto2 sentences is subject-object-verb. Now listen to a few examples.

- **pleck li vode lu praz**

In this example, we first state the subject, *pleck li*. This noun is doing the action. Second, we state the object, *vode lu*. This noun is receiving the action. Finally, at the end of the sentence you find the verb, *klin*. Listen to the example again.

- **pleck li vode lu praz**

Here’s another example:

- **pleck li blom lu nim**

In this example, *pleck li* is the subject and comes at the beginning of the sentence. *Blom lu* is the object and is placed after the subject. Nim is the verb and is found after the object. Here are more examples for you to listen to:

- **vode lu neep li praz**
- **neep li pleck li yab**
- **blom lu pleck li praz**

  - Now let’s add in adjectives to describe the nouns. Listen to the following example:

- **neep neime li blom troiso lu praz**

In this sentence, *neep neime li* is the complete subject, and *blom troiso lu* is the complete object. Here are some more examples of transitive sentences with subject-object-verb word order for you to listen to:

- **blom neimo lu vode troiso lu nim**
- **pleck troise li vode neimo lu praz**
- **neep li pleck neime li yab**

**Adverbs**

- Adverbs. Finally, we will learn about Brocanto2 adverbs. Adverbs describe how the action of a verb is carried out. In Brocanto2 adverbs describe the direction of the actions. There are only two adverbs.

  - **Noyka**
o Zayma

Adverbs are placed immediately behind the verb. This means that if an adverb is used in a sentence, it will be the last word of the sentence. Let’s look at few examples of sentences with adverbs in them.

- Blom lu klin zayma
- Vode neimo lu nim noyka
- Neep troise li klin noyka
- Pleck neime li nim zayma
- Vode troiso lu pleck neime li yab noyka
- Blom troiso lu neep troise li praz zayma
- Pleck neime li vode troiso lu yab zayma
Appendix C

Uninstructed Training Script

(Directions)

For the second part of the study you will receive training on an artificial language. You will hear examples of words, phrases and sentences from this artificial language and will see both still and moving images on a game board that correspond to those words, phrases and sentences. Immediately after, you will be asked some questions about what you have seen.

Later, in the third part of this study, you will use the artificial language to play the game that was explained to you previously.

This second part of the study will last approximately 20 minutes and is timed by the computer. Please be aware that you will not be able to go back and review any of the examples. Click on the ‘Continue’ button when you are ready to start the second part of the study.

(Show Continue button)

(Training)

<table>
<thead>
<tr>
<th>pleck li</th>
<th>round pleck</th>
</tr>
</thead>
<tbody>
<tr>
<td>pleck troise li</td>
<td>round pleck</td>
</tr>
<tr>
<td>pleck li</td>
<td>square pleck</td>
</tr>
<tr>
<td>pleck neime li</td>
<td>square pleck</td>
</tr>
<tr>
<td>neep li</td>
<td>square neep</td>
</tr>
<tr>
<td>neep neime li</td>
<td>square neep</td>
</tr>
<tr>
<td>neep li</td>
<td>round neep</td>
</tr>
<tr>
<td>neep troise li</td>
<td>round neep</td>
</tr>
<tr>
<td>pleck troise li</td>
<td>round pleck</td>
</tr>
<tr>
<td>neep troise li</td>
<td>round neep</td>
</tr>
<tr>
<td>neep li</td>
<td>square neep</td>
</tr>
<tr>
<td>pleck li</td>
<td>round pleck</td>
</tr>
<tr>
<td>neep neime li</td>
<td>square neep</td>
</tr>
<tr>
<td>pleck neime li</td>
<td>square pleck</td>
</tr>
<tr>
<td>blom lu</td>
<td>round blom</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>blom troiso lu</td>
<td>round blom</td>
</tr>
<tr>
<td>blom lu</td>
<td>square blom</td>
</tr>
<tr>
<td>blom neimo lu</td>
<td>square blom</td>
</tr>
<tr>
<td>vode lu</td>
<td>square vode</td>
</tr>
<tr>
<td>vode neimo lu</td>
<td>square vode</td>
</tr>
<tr>
<td>vode troiso lu</td>
<td>round vode</td>
</tr>
<tr>
<td>vode lu</td>
<td>square vode</td>
</tr>
<tr>
<td>blom troiso lu</td>
<td>round blom</td>
</tr>
<tr>
<td>vode troiso lu</td>
<td>round vode</td>
</tr>
<tr>
<td>vode lu</td>
<td>round vode</td>
</tr>
<tr>
<td>blom lu</td>
<td>round blom</td>
</tr>
<tr>
<td>vode neimo lu</td>
<td>square vode</td>
</tr>
<tr>
<td>blom neimo lu</td>
<td>square blom</td>
</tr>
<tr>
<td>neep li</td>
<td>square neep</td>
</tr>
<tr>
<td>vode lu</td>
<td>round vode</td>
</tr>
<tr>
<td>neep troise li</td>
<td>round neep</td>
</tr>
<tr>
<td>vode troiso lu</td>
<td>round vode</td>
</tr>
<tr>
<td>neep neime li</td>
<td>square neep</td>
</tr>
<tr>
<td>vode neimo lu</td>
<td>square vode</td>
</tr>
<tr>
<td>pleck li</td>
<td>square pleck</td>
</tr>
<tr>
<td>blom lu</td>
<td>round blom</td>
</tr>
<tr>
<td>pleck troise li</td>
<td>round pleck</td>
</tr>
<tr>
<td>blom troiso lu</td>
<td>round blom</td>
</tr>
<tr>
<td>pleck neime li</td>
<td>square pleck</td>
</tr>
<tr>
<td>blom neimo lu</td>
<td>square blom</td>
</tr>
<tr>
<td>----------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>pleck troise li</td>
<td>round pleck</td>
</tr>
<tr>
<td>neep troise li</td>
<td>round neep</td>
</tr>
<tr>
<td>blom troiso lu</td>
<td>round blom</td>
</tr>
<tr>
<td>vode troiso lu</td>
<td>round vode</td>
</tr>
<tr>
<td>pleck neime li</td>
<td>square pleck</td>
</tr>
<tr>
<td>blom neimo lu</td>
<td>square blom</td>
</tr>
<tr>
<td>neep neime li</td>
<td>square neep</td>
</tr>
<tr>
<td>vode neimo lu</td>
<td>square vode</td>
</tr>
<tr>
<td>vode lu klin</td>
<td>round vode move horizontally</td>
</tr>
<tr>
<td>neep li klin</td>
<td>square neep move horizontally</td>
</tr>
<tr>
<td>pleck li klin</td>
<td>square pleck move vertically</td>
</tr>
<tr>
<td>blom lu klin</td>
<td>round blom move horizontally</td>
</tr>
<tr>
<td>neep troise li klin</td>
<td>round neep move vertically</td>
</tr>
<tr>
<td>blom troiso lu klin</td>
<td>round blom move horizontally</td>
</tr>
<tr>
<td>vode neimo lu klin</td>
<td>square vode move vertically</td>
</tr>
<tr>
<td>pleck neime li klin</td>
<td>square pleck move horizontally</td>
</tr>
<tr>
<td>neep neime li klin</td>
<td>square neep move vertically</td>
</tr>
<tr>
<td>blom neimo lu klin</td>
<td>square blom move horizontally</td>
</tr>
<tr>
<td>vode troiso lu klin</td>
<td>round vode move vertically</td>
</tr>
<tr>
<td>pleck troise li klin</td>
<td>round pleck move horizontally</td>
</tr>
<tr>
<td>vode lu nim</td>
<td>round vode capture horizontally</td>
</tr>
<tr>
<td>vode lu neep li nim</td>
<td>square vode square neep capture horizontally</td>
</tr>
<tr>
<td>pleck neime li nim</td>
<td>square pleck capture horizontally</td>
</tr>
<tr>
<td>pleck li blom lu nim</td>
<td>round pleck square blom capture vertically</td>
</tr>
<tr>
<td>neep neime li nim</td>
<td>square neep capture vertically</td>
</tr>
<tr>
<td>blom lu vode lu nim</td>
<td>round blom round vode capture horizontally</td>
</tr>
<tr>
<td>vode troiso lu nim</td>
<td>round vode capture vertically</td>
</tr>
<tr>
<td>pleck troise li blom neimo lu nim</td>
<td>round pleck square blom capture horizontally</td>
</tr>
<tr>
<td>neep troise li nim</td>
<td>round neep capture vertically</td>
</tr>
<tr>
<td>blom troiso lu pleck neime li nim</td>
<td>round blom square pleck capture horizontally</td>
</tr>
<tr>
<td>vode neimo lu nim</td>
<td>square vode capture horizontally</td>
</tr>
<tr>
<td>blom neimo lu vode troiso lu nim</td>
<td>square blom round vode capture vertically</td>
</tr>
<tr>
<td>neep li vode lu yab</td>
<td>square neep round vode release vertically</td>
</tr>
<tr>
<td>neep li yab</td>
<td>round neep release horizontally</td>
</tr>
<tr>
<td>blom lu pleck li yab</td>
<td>square blom square pleck release vertically</td>
</tr>
<tr>
<td>blom lu yab</td>
<td>square blom release horizontally</td>
</tr>
<tr>
<td>neep li pleck li yab</td>
<td>round neep round pleck release horizontally</td>
</tr>
<tr>
<td>blom neimo lu yab</td>
<td>square blom release vertically</td>
</tr>
<tr>
<td>neep troise li vode neimo lu yab</td>
<td>round neep square vode release horizontally</td>
</tr>
<tr>
<td>pleck troise li yab</td>
<td>round pleck release vertically</td>
</tr>
<tr>
<td>vode troiso lu neep neime li yab</td>
<td>round vode square neep release vertically</td>
</tr>
<tr>
<td>blom troiso lu yab</td>
<td>round blom release horizontally</td>
</tr>
<tr>
<td>neep li pleck neime li yab</td>
<td>round neep square pleck release vertically</td>
</tr>
<tr>
<td>pleck neime li yab</td>
<td>square pleck release vertically</td>
</tr>
<tr>
<td>vode lu neep li praz</td>
<td>round vode square neep swap horizontally</td>
</tr>
<tr>
<td>blom lu pleck li praz</td>
<td>square blom round neep swap vertically</td>
</tr>
<tr>
<td>neep li blom lu praz</td>
<td>square neep square blom swap horizontally</td>
</tr>
<tr>
<td>pleck li neep li praz</td>
<td>round pleck round neep swap horizontally</td>
</tr>
<tr>
<td>vode lu blom lu praz</td>
<td>square vode round blom swap horizontally</td>
</tr>
<tr>
<td>neep neime li pleck troise li praz</td>
<td>square neep round pleck swap horizontally</td>
</tr>
<tr>
<td>blom troiso lu vode neimo lu praz</td>
<td>round blom square vode swap vertically</td>
</tr>
<tr>
<td>Move Description</td>
<td>Resulting Move Description</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>blom troiso lu neep neime li praz</td>
<td>round blom square neep swap vertically</td>
</tr>
<tr>
<td>vode troiso lu pleck neime li praz</td>
<td>round vode square pleck swap horizontally</td>
</tr>
<tr>
<td>neep troise li blom neimo lu praz</td>
<td>round neep square blom swap vertically</td>
</tr>
<tr>
<td>pleck troise li vode neimo lu praz</td>
<td>round pleck square vode swap horizontally</td>
</tr>
<tr>
<td>neep li klin noyka</td>
<td>square neep move vertically</td>
</tr>
<tr>
<td>neep li klin zayma</td>
<td>square neep move horizontally</td>
</tr>
<tr>
<td>blom lu klin noyka</td>
<td>round blom move vertically</td>
</tr>
<tr>
<td>blom lu klin zayma</td>
<td>round blom move horizontally</td>
</tr>
<tr>
<td>neep troise li klin zayma</td>
<td>round neep move horizontally</td>
</tr>
<tr>
<td>blom troiso lu klin zayma</td>
<td>round blom move horizontally</td>
</tr>
<tr>
<td>neep troise li klin noyka</td>
<td>round neep move vertically</td>
</tr>
<tr>
<td>blom troiso lu klin noyka</td>
<td>round blom move vertically</td>
</tr>
<tr>
<td>pleck li nim zayma</td>
<td>round pleck capture horizontally</td>
</tr>
<tr>
<td>pleck li nim noyka</td>
<td>round pleck capture vertically</td>
</tr>
<tr>
<td>vode lu nim noyka</td>
<td>square vode capture vertically</td>
</tr>
<tr>
<td>vode lu nim zayma</td>
<td>square vode capture horizontally</td>
</tr>
<tr>
<td>neep li blom lu yab noyka</td>
<td>square neep square blom release vertically</td>
</tr>
<tr>
<td>vode lu pleck li yab noyka</td>
<td>round vode square pleck release vertically</td>
</tr>
<tr>
<td>pleck li vode lu yab zayma</td>
<td>square pleck round vode release horizontally</td>
</tr>
<tr>
<td>blom lu neep li yab zayma</td>
<td>round blom round neep release horizontally</td>
</tr>
<tr>
<td>pleck neime li nim noyka</td>
<td>square pleck capture vertically</td>
</tr>
<tr>
<td>vode neimo lu nim noyka</td>
<td>square vode capture vertically</td>
</tr>
<tr>
<td>pleck neime li nim zayma</td>
<td>square pleck capture horizontally</td>
</tr>
<tr>
<td>vode neimo lu nim zayma</td>
<td>square vode capture horizontally</td>
</tr>
<tr>
<td>pleck li vode lu praz noyka</td>
<td>round pleck round vode swap vertically</td>
</tr>
<tr>
<td>blom lu neep li praz noyka</td>
<td>square blom square neep swap vertically</td>
</tr>
<tr>
<td>neep li blom lu praz zayma</td>
<td>round neep  square blom  swap horizontally</td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>vode lu pleck li praz zayma</td>
<td>square vode  round pleck  swap horizontally</td>
</tr>
<tr>
<td>pleck troise li vode troiso lu praz noyka</td>
<td>round pleck round vode swap vertically</td>
</tr>
<tr>
<td>blom troiso lu neep troise li praz zayma</td>
<td>round blom round neep swap horizontally</td>
</tr>
<tr>
<td>neep neime li blom neimo lu praz noyka</td>
<td>square neep square blom swap vertically</td>
</tr>
<tr>
<td>vode neimo lu pleck neime li praz zayma</td>
<td>square vode square pleck swap horizontally</td>
</tr>
<tr>
<td>neep troise li blom neimo lu yab noyka</td>
<td>round neep square blom release vertically</td>
</tr>
<tr>
<td>vode troiso lu pleck neime li yab noyka</td>
<td>round vode square pleck release vertically</td>
</tr>
<tr>
<td>pleck neime li vode troiso lu yab zayma</td>
<td>square pleck round  vode release horizontally</td>
</tr>
<tr>
<td>blom troiso lu neep troise li yab zayma</td>
<td>round blom round neep release horizontally</td>
</tr>
</tbody>
</table>
References


