

EXAMINING MORPHOLOGICAL SENSITIVITY AND INDIVIDUAL DIFFERENCES  
IN THE ACQUISITION OF AGREEMENT IN A SECOND LANGUAGE

BY

Kristi Bond

Submitted to the graduate degree program in Linguistics  
and the Graduate Faculty of the University of Kansas  
in partial fulfillment of the requirements for the degree of  
Master's of Arts.

Committee:

---

Dr. Alison Gabriele  
Co-Chairperson

---

Dr. Robert Fiorentino  
Co-Chairperson

---

Dr. Harold Torrence  
Member

Date defended: February 18, 2011

The Thesis Committee for Kristi Bond certifies  
that this is the approved Version of the following thesis:

EXAMINING MORPHOLOGICAL SENSITIVITY AND INDIVIDUAL DIFFERENCES  
IN THE ACQUISITION OF AGREEMENT IN A SECOND LANGUAGE

Committee:

---

Dr. Alison Gabriele  
Co-Chairperson

---

Dr. Robert Fiorentino  
Co-Chairperson

---

Dr. Harold Torrence  
Member

Date approved: March 8, 2011

## ABSTRACT

This experiment investigates learner sensitivity to agreement violations involving uninterpretable number and gender features in Spanish. Twenty-two low-proficiency, English-speaking learners and a control group of twelve native Spanish-speakers gave grammaticality judgments after viewing sentences presented rapidly in word-by-word format. The stimuli targeted agreement between nouns in subject positions and adjectives in the predicate. Responses to violations of number agreement, which is instantiated on verbs in English but not adjectives, were compared to violations of gender agreement, which is unique to the L2 Spanish. In line with the Full Transfer/Full Access theory (Schwartz and Sprouse, 1994; 1996), it was hypothesized that learners would demonstrate the same pattern of sensitivity to agreement violations as native speakers, but that they would be less sensitive to gender versus number violations. The question of whether or not individual differences impact grammaticality judgments was also explored, due to the claim of the Fundamental Difference Hypothesis (Bley-Vroman, 1989; 1990) that successful adult L2 acquisition relies on exceptional abilities outside the realm of Universal Grammar (UG). Learners completed the Short Form of the Modern Languages Aptitude Test (Carroll and Sapon, 1959) and the Raven Advanced Progressive Matrices (Raven, 1965) as measures of verbal and nonverbal aptitude, respectively. Results support Full Transfer/Full Access, suggesting that uninterpretable features can be acquired by adult L2 learners, even in the case of gender features that are not present in the L1. Unique L2 features do seem to present a challenge at this stage of acquisition, as the learners performed significantly worse for gender versus number agreement. In line with FTFA, however, performance on gender agreement correlated most strongly with proficiency. Grammaticality judgments were not found to correlate with either verbal or nonverbal aptitude.

## ACKNOWLEDGEMENTS

My path to this degree has been directed by God, whose sense of humor continues to surprise me. “In all your ways acknowledge Him and He will direct your paths” (Proverbs 3:6). I am also grateful that He blessed me many years ago with a husband who is not afraid to support my goals even when their outcome is not clear. Thank you, Eric, for protecting me and my time, for loving our kids as much as I do, and for making many sacrifices so that I could finish this project. I also owe many thanks to my parents, Les and Vicki Jones, who gave me my earliest and best education, without a doubt contributing to all the rest, and who on at least one occasion provided extended child care so that I could write. My in-laws, Dean and Judy Bond, also deserve thanks for their support and for the child care they provided.

I am thankful for my committee members, Alison Gabriele, Robert Fiorentino, and Harold Torrence, each of whom provided excellent feedback throughout this research project. Alison, thanks for walking with me through the whole experience of graduate school. Your guidance and instruction have been invaluable, both throughout this project and in other endeavors. Rob, many thanks for your help in research design, set up and data analysis, as well as for making me feel I had contributed to the lab in many ways.

Many others contributed to the successful completion of this project. I express deepest thanks to José Alemán Bañon, who made much of this journey with me and whose help was instrumental at every stage from stimuli design to data collection and analysis. I also thank other researchers in the Neurolinguistics and Language Processing Lab who stepped in to help with data collection on numerous occasions: Jamie Bost, Ella Fund-Reznicek, Yuka Naito-Billen, Lamar Hunt, Stephen Politzer-Ahles, and Jeff Shouse. For help in recruiting participants, I owe

thanks to Rosalea Postma-Carttar and Leigh Rysko of the KU Department of Spanish and Portuguese, Wendy Herd, and Alonso Canales (who also aided in stimuli design).

Finally, I want to thank my Christian sisters whose support and prayers were critical during the toughest moments of balancing graduate school and parenting: Kristy McKinney, Marti Elford, Sonya Schinkel, Kristy Teska, MaryBeth Petr, and Elaine McCullough.

## DEDICATION

This thesis is dedicated to my daughter Hannah, whose grin instantly dissolves tension during the most tedious of tasks, and to my son Micah for the numerous recreational breaks from writing.

Yes, son, I can play now.

## TABLE OF CONTENTS

1	Introduction.....	1
2	Gender and Number Agreement .....	3
2.1	Gender and Number Features in Spanish.....	3
2.1.1	Gender.....	3
2.1.2	Number .....	5
2.1.3	Agreement Marking in Spanish .....	6
2.2	Theoretical Accounts of Number and Gender Agreement.....	7
2.2.1	Number Agreement.....	7
2.2.2	Gender Agreement.....	8
3	Theories of Second Language Acquisition .....	9
3.1	Models of Access to Universal Grammar in Adult L2 Acquisition.....	10
3.2	Full Access: The Full Transfer/Full Access Theory .....	11
3.3	Partial Access: The Interpretability Hypothesis.....	14
3.4	No Access: The Fundamental Difference Hypothesis .....	17
3.5	Summary .....	19
4	Individual Differences in Second Language Acquisition .....	20
4.1	Verbal Aptitude.....	21
4.2	Nonverbal Aptitude.....	24
5	The Current Study.....	29
5.1	Research Questions .....	29
5.2	Predictions.....	30

5.2.1	Learner Sensitivity to Agreement Violations .....	30
5.2.2	Aptitude.....	32
5.2.2.1	Verbal Aptitude.....	32
5.2.2.2	Nonverbal Aptitude.....	33
6	Methods.....	33
6.1	Participants.....	33
6.2	Materials .....	36
6.2.1	Stimuli.....	36
6.2.2	Gender Assignment Task.....	38
6.2.3	Vocabulary Recognition Task.....	39
6.2.4	Verbal Aptitude: The Modern Languages Aptitude Test .....	39
6.2.5	Nonverbal Aptitude: Raven’s Advanced Progressive Matrices.....	40
6.2.6	Spanish Proficiency Test.....	41
6.3	Procedures.....	42
6.3.1	Learners.....	42
6.3.2	Native Speakers .....	45
7	Results.....	45
7.1	Learner Sensitivity to Agreement Violations .....	46
7.1.1	Descriptive Statistics.....	46
7.1.1.1	Grammaticality Judgments .....	46
7.1.1.2	Vocabulary Recognition .....	47
7.1.1.3	Gender Assignment.....	48
7.1.2	ANOVA Results .....	48



7.1.2.1	Analysis of Group and Grammaticality Effects .....	48
7.1.2.2	Analysis of Effects for Violation Type .....	50
7.2	Individual Differences .....	51
7.2.1	Independent Variables .....	51
7.2.1.1	Verbal Aptitude.....	51
7.2.1.2	Nonverbal Aptitude.....	52
7.2.1.3	Proficiency .....	52
7.2.1.4	Summary.....	52
7.2.2	Intercorrelations of Independent Variables.....	512
7.2.3	Correlations Between Independent and Dependent Variables.....	513
7.2.3.1	Verbal Aptitude and Sensitivity to Agreement Violations .....	54
7.2.3.2	Nonverbal Aptitude and Sensitivity to Agreement Violations .	56
7.2.3.3	Proficiency and Sensitivity to Agreement Violations.....	56
7.2.3.4	Summary.....	57
8	Discussion.....	57
9	Conclusion .....	67
	References.....	68
	APPENDIX 1: Target stimuli .....	76
	APPENDIX 2: Fillers .....	87
	APPENDIX 3: Mean Acceptance Rates by Condition for Native Speakers .....	98
	APPENDIX 4: Mean Acceptance Rates by Condition for Learners .....	99
	APPENDIX 5: Combined, Number, and Gender d' Scores for Native Speakers.....	999
	APPENDIX 6: Combined, Number, and Gender d' Scores for Learners.....	100

APPENDIX 7: Combined, Number, and Gender d' Scores Plotted Against MLAT Total,  
RAVEN, and Proficiency Scores .....101

## LIST OF TABLES

Table 1. Profile of L2 group characteristics.....	34
Table 2. Characteristics of native speakers, in years. ....	36
Table 3. Test order between groups of L2 participants. ....	42
Table 4. Results of repeated measures ANOVAs on acceptance rates for each violation type....	48
Table 5. Observed measures of aptitude and proficiency measures, along with the maximum possible scores for each measure .....	52
Table 6. Intercorrelations of aptitude and proficiency measures.....	53
Table 7. Correlations between d' scores and measures of verbal aptitude. ....	54
Table 8. Correlations between d' scores and measures of aptitude and proficiency. ....	56
Table 9. Mean acceptance rates by condition for native speakers.....	97
Table 10. Mean acceptance rates by condition for learners.....	98
Table 11. Native speaker d' scores for combined conditions, and for both number and gender agreement separately.....	99
Table 12. Learner group d' scores for combined conditions, and for both number and gender agreement separately.....	100

## LIST OF FIGURES

Figure 1. Mean acceptance rates for native speakers (NS) and learners (L2) across experimental conditions.....	46
Figure 2. The $d'$ scores for number agreement plotted against MLAT4 scores. ....	55
Figure 3. Combined $d'$ scores plotted against MLAT Total scores.....	101
Figure 4. Combined $d'$ scores plotted against RAVEN scores.....	101
Figure 5. Combined $d'$ scores plotted against proficiency scores. ....	102
Figure 6. The $d'$ scores for number agreement plotted against MLAT Total scores.....	102
Figure 7. The $d'$ scores for number agreement plotted against RAVEN scores.....	102
Figure 8. The $d'$ scores for number agreement plotted against proficiency scores. ....	103
Figure 9. The $d'$ scores for gender agreement plotted against MLAT Total scores. ....	103
Figure 10. The $d'$ scores for gender agreement plotted against RAVEN scores.....	103
Figure 11. The $d'$ scores for gender agreement plotted against proficiency scores.....	104

## 1 Introduction

Sensitivity to agreement violations has recently become one of the testing grounds for theories of adult second-language (L2) acquisition (e.g., Prévost and White, 2000; White et al., 2004; Hawkins, 1998). In particular, agreement involving features not present in learners' native language (L1) seems to be a locus of difficulty for learners even at high levels of proficiency. One such feature is grammatical gender, which often triggers agreement on a variety of syntactic elements in languages where it occurs. In the following example of Spanish gender agreement, the masculine gender present on the noun *vestido* 'dress' is also carried by the determiner *el* 'the' and the adjective *blanco* 'white' via morphosyntactic agreement.

(1)  $el_{\text{masc}}$   $vestido_{\text{masc}}$   $blanco_{\text{masc}}$   
the dress white  
'the white dress'

Adult L2 learners of languages with gender agreement exhibit difficulties, particularly when grammatical gender is not instantiated in their L1 (e.g., Hawkins, 1998; Franceschina, 2002). Similar difficulties have also been noted for number (e.g., McDonald, 2000; Jiang, 2004, 2007), *wh*-features (e.g., Hawkins and Chan, 1997), subject-verb agreement (e.g., McDonald, 2000), and more, often dependent upon the L1 of the learners. Accordingly, various theories of L2 acquisition can be differentiated based on their predictions regarding unique L2 features – predictions that result from differing approaches to limitations on adult L2 acquisition. Of interest here are two contrasting theories: the Interpretability Hypothesis (Tsimplici and Mastropavlou, 2007), which, as a further refinement of the Failed Functional Features Hypothesis (Hawkins and Chan, 1997), claims that learners cannot acquire uninterpretable features like gender if such features are not present in their L1, and the Full Transfer/Full Access

(Schwartz and Sprouse, 1994) model of L2 acquisition, under which even uninterpretable features not present in the L1 are ultimately acquirable.

A third theory is also relevant to the current study: the Fundamental Difference Hypothesis (FDH) (Bley-Vroman, 1989, 1990, 2009). Under this theory, adult learners must rely on domain-general, possibly shallow processing routines, along with their knowledge of the L1. An extension of the FDH is the prediction that individual differences will play a larger role in adult L2 acquisition than in child L1 or L2 acquisition. For example, a number of studies have found language analytic ability, as measured by verbal aptitude tests, to be a critical factor in the general success of adult L2 acquisition (e.g., Skehan, 1989; Harley and Hart, 1997; DeKeyser, 2000; Abrahamsson and Hyltenstam, 2008). DeKeyser (2000) suggests that aptitude may play a greater role in the acquisition of less salient structures, but little is known about the role of aptitude with regard to structures that are similar or dissimilar in the L1 and L2. Bley-Vroman's (1990) proposal regarding domain-general problem-solving capacities (versus domain-specific language abilities) warrants further investigation as well.

The current study investigated the acquisition of number and gender agreement in L2 Spanish by native English speakers. Learner sensitivity to agreement violations involving the unique L2 feature of grammatical gender on adjectival predicates was compared to sensitivity to number agreement in the same context. Under both the Interpretability Hypothesis and FTFA, number agreement should pose no problem for English-speaking learners, since the number feature is present in English. However, these theories make different predictions regarding English-speaking learners' acquisition of gender agreement. Under the Interpretability Hypothesis, learners will not demonstrate sensitivity to violations of gender agreement. On the other hand, if the Full Transfer/Full Access Hypothesis is on the right track, then learners should

show development toward native-like sensitivity to gender agreement violations. Additionally, both a verbal and a nonverbal aptitude test were also conducted and analyzed using correlations in order to test the predictions of the Fundamental Difference Hypothesis (Bley-Vroman, 1989, 1990, 2009) and explore the relationship between individual differences and the acquisition of specific structures that differ in their similarity to the L1.

The study is organized as follows. Section 2 presents the properties of gender and number features in Spanish that are being investigated, along with theoretical accounts of agreement involving these features. In Section 3, contrasting theories of L2 acquisition are outlined, especially with regard to limitations present in adult L2 acquisition and the role of uninterpretable morphosyntactic features. In light of the relevance of individual differences to L2 theory and to the study presented here, Section 4 discusses individual differences in verbal and nonverbal aptitude. Research questions and predictions are presented in Section 5, and the methods of investigation pursued in this study are discussed in Section 6. Section 7 presents the results of the study, followed by discussion of the results and future directions for research in Section 8, along with conclusions in Section 9.

## **2 Gender and Number Agreement**

### **2.1 Gender and Number Features in Spanish**

#### **2.1.1 Gender**

Although there is some debate about the locus of grammatical gender features in Spanish noun phrases, gender is argued to be an inherent property of nouns (Corbett, 1991). For animate nouns, especially those referring to humans, lexical gender often reflects biological gender. The lexical gender of inanimate nouns is arbitrary, however, and in fact the same object can

sometimes be referred to as either masculine or feminine (2), depending on the choice of vocabulary.

- (2) a. bolígrafo<sub>masc</sub>  
‘pen’  
b. pluma<sub>fem</sub>  
‘pen’ (historical derivation from quill and ink)

The morphological marking of gender is not straightforward either. Harris (1991) points out that the prototypical gender markers in Spanish also appear on words with no gender, suggesting that they are word markers rather than markers specific to gender. He nevertheless categorizes Spanish nouns based on their conformity to the prototypical markers of *-o* for masculine nouns (2a, above) and *-a* for feminine nouns (2b). While these markers are systematically applied in the language, there are also exceptions, i.e. some masculine nouns that end in *-a*, as well as some feminine nouns that end in *-o*, as seen in (3) below. Other non-canonical nouns can end in consonants (4a), *-e* (4b), or *-u* (4c).

- (3) a. mapa<sub>masc</sub>  
‘map’  
b. mano<sub>fem</sub>  
‘hand’  
(4) a. poción<sub>fem</sub>  
‘potion’  
b. infante<sub>masc</sub>  
‘infant’  
c. tribu<sub>masc</sub>  
‘tribe’

The current study employed only nouns with canonical gender (or word) markers, as in (2) above.



### 2.1.2 Number

In addition to gender, Spanish nouns uniformly bear number. Singular morphology is null or unmarked; plural morphology depends on the phonetic properties of the final segment in the root. If the root ends in an unstressed vowel (5a), the plural marker is [-s] (5b).

- (5) a. vestido<sub>sg</sub>  
      ‘dress’  
      b. vestidos<sub>pl</sub>  
          ‘dresses’

If the final segment is a stressed vowel<sup>1</sup> or consonant, as in (6a, 7a), the plural marker will be [-es] (6b, 7b). It should be noted that the lack of the diacritic in (7b) is a purely orthographic difference between the singular and plural forms.

- (6) a. tabú<sub>sg</sub>  
      ‘taboo’  
      b. tabúes<sub>pl</sub>  
          ‘taboos’  
(7) a. nación<sub>sg</sub>  
      ‘nation’  
      b. naciones<sub>pl</sub>  
          ‘nations’

A null (unmarked) plural is possible only in the case of roots of more than one syllable ending in [-s, -x]:

- (8) a. tesis<sub>sg</sub>  
      ‘thesis’  
      b. tesis<sub>pl</sub>  
          ‘theses’

The current study employed only nouns ending in unstressed vowels of the type in (5) above.

---

<sup>1</sup>One exception is the plural of *menú* “menu”, which is *menús*.

### 2.1.3 Agreement Marking in Spanish

In Spanish, nouns trigger number and gender agreement on other elements both within and across phrases. In the within-phrase examples in (9), the determiner and adjective must agree in both gender and number with the masculine, singular noun *vestido* ‘dress’ in (9a) and the feminine, plural noun *faldas* ‘skirts’ in (9b).

- (9) a. el            vestido      blanco  
      the<sub>masc.sg</sub> dress<sub>masc.sg</sub> white<sub>masc.sg</sub>  
      ‘the white dress’
- b. las            faldas      blancas  
      the<sub>fem.pl</sub> skirts<sub>fem.pl</sub> white<sub>fem.pl</sub>  
      ‘the white skirts’

Agreement marking also surfaces across phrases in predicates, as can be seen in (10), where the adjective must agree with the noun in gender and number. Note that the auxiliary verb, as is also true of all main verbs in Spanish, must agree with the subject in number, but not gender.

- (10) a. el            vestido      es      blanco  
      the<sub>masc.sg</sub> dress<sub>masc.sg</sub> is<sub>sg</sub> white<sub>masc.sg</sub>  
      ‘the dress is white’
- b. las            faldas      son      blancas  
      the<sub>fem.pl</sub> skirts<sub>fem.pl</sub> are<sub>pl</sub> white<sub>fem.pl</sub>  
      ‘the dresses are white’

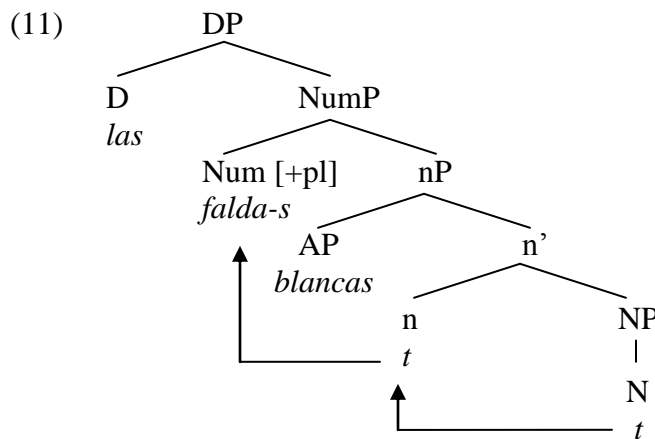
Gender and number on adjectives is marked in the same way as was described above for nouns. However, a difference emerges in how verbs are marked for number. Verbs display a number-marking system that is conflated with person-marking. Since person agreement is not under investigation here, only 3<sup>rd</sup>-person verb forms were used in the current study. Furthermore, as the only verb forms used in experimental stimuli here were *es* ‘is’ and *son* ‘are’, the examples in

(10) above will serve to demonstrate the differences between number-marking on the verbs and adjectives used in the current study.

## 2.2 Theoretical Accounts of Number and Gender Agreement

### 2.2.1 Number Agreement

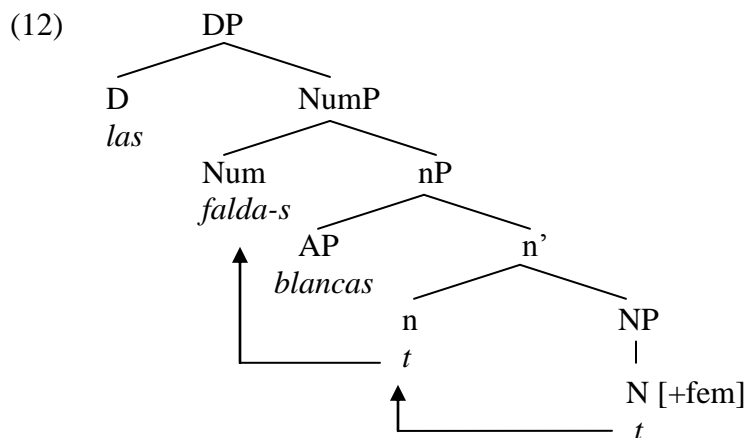
For Romance languages such as Spanish, Cinque (1994) has proposed that the noun moves within the Determiner Phrase (DP) from a position below adjectival modifiers to a position just below the DP. Given Ritter's (1991) proposal that number is instantiated in grammars as a functional category projecting its own phrase within the DP, it is generally accepted that the Number Phrase (NumP) is the location to which the noun moves and from which its uninterpretable number feature is checked by an interpretable number feature on Num (11), under the Chomskyan (1995) theory of feature-checking as the mechanism for agreement relations. Num also checks the uninterpretable number features on any Adjective Phrase (AP) below it<sup>2</sup>, as well as on the determiner (D) through covert movement of Num features to D.



<sup>2</sup> There is considerable debate regarding the position of APs within the NP structure. They are presented here in the position of spec-nP, following Carstens (2000).

### 2.2.2 Gender Agreement

There is a general consensus that gender on nouns is lexically determined<sup>3</sup> and interpretable<sup>4</sup> (Carroll, 1989; Corbett, 1991; Carstens, 2000). Other syntactic elements may bear uninterpretable features that must be checked, as in the examples of Spanish gender agreement on determiners and adjectives in (9, 10) above. For languages like Italian and Spanish, Carstens (2000) describes gender agreement as in (12) below. Here the gender feature raises overtly with the noun to NumP, where it can check an uninterpretable gender on any adjective in spec-nP position. The gender feature itself subsequently raises covertly to DP, where number and gender features on D are checked.

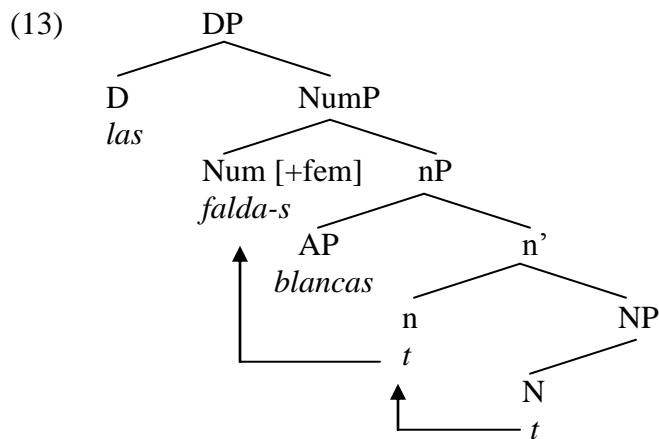


One other account of gender agreement should be considered. Ritter (1993) compares gender agreement in Modern Hebrew and Romance languages, concluding that differences in the properties of nouns in the two languages (and in particular word formation strategies) demonstrate parametric variation in the locus of interpretable gender features. Rather than attributing the gender feature in Romance to the noun, she argues that it must be a feature of the

<sup>3</sup> But see Picallo (1991), who claims that gender is not projected on N but heads its own projection above the noun phrase (NP), but below NumP. In his account, the noun must raise first to GenP and then to NumP to receive gender and number, respectively.

<sup>4</sup> Carstens (2000) presents gender as an interpretable feature on nouns, although she adds that [-inherent] gender could very well be an uninterpretable feature that does not erase (see footnote 12, p. 328).

NumP. For Spanish, the noun obligatorily raises to NumP to check number and gender features, as shown by the position of [+fem] in (13):



Regardless of the category on which the interpretable gender feature occurs, these theories have in common the lexical attribution of gender to some element within the DP phrase that then checks uninterpretable features on other elements, producing the relevant agreement paradigms. It is important to note that the task for learners involves not only the acquisition of the gender feature and its syntactic consequences, but also the correct assignment of gender to nouns as they are acquired. Having laid out the morphosyntactic features under consideration here, we now turn to theories of how these features are acquired by L2 learners.

### 3 Theories of Second Language Acquisition

According to both popular belief and theories in second-language (L2) acquisition, it is more difficult for adults to become near-native in their use of an L2 than for children to do so. One theory that attempts to account for this difference is adapted from Penfield and Roberts' (1959) and Lenneberg's (1967) proposals of a Critical Period (CP) for first language acquisition. In general, the Critical Period Hypothesis claims that language acquisition is impaired after a

certain age due to maturation of cognitive abilities and/or changes in the neuroanatomy of the brain. Although still under debate, child language acquisition is believed to be guided by an innate learning mechanism or Language Acquisition Device that facilitates the selection of features in a particular language, along with their parameter settings (often assumed to be binary options), from a universal set of features and parameters, or Universal Grammar (UG) (Chomsky, 1965; 1980; 1981). In second language research, a major question is whether or not that mechanism is available to post-Critical-Period learners. This question also provides the framework for distinctive theoretical predictions regarding linguistic features involved in morphosyntactic processes such as gender and number agreement discussed in the previous section.

### **3.1 Models of Access to Universal Grammar in Adult L2 Acquisition**

One of the primary differences between contrasting theories of L2 acquisition is their approach to Universal Grammar. Most theories based on an assumption of Universal Grammar can be categorized as either restricted-access or full-access theories. Theories of restricted access are of two varieties: some claim that UG is not available at all (no access) past the Critical Period, while others claim that adult L2 learners have partial access to UG, usually through the L1. Examples of partial-access theories include the No Parameter Setting Hypothesis (e.g., Clahsen and Muysken, 1986) and the Failed Functional Features Hypothesis (Hawkins and Chan, 1997), along with its more recent version, the Interpretability Hypothesis (Tsimplici and Dimitrakopoulou, 2007). On the other hand, the Fundamental Difference Hypothesis (Bley-Vroman, 1989, 1990) claims that adult learners have no access to Universal Grammar. These

theories have in common the claims that L1 and L2 acquisition are *qualitatively* different and that the adult learner's mental representation of the L2 will therefore be incomplete.

Accounts based on theories of partial or no access to UG contrast with theories that claim full access to UG for all learners, such as the Strong Continuity Hypothesis (Epstein, Flynn, and Martohardjono, 1996), the Full Transfer/Full Access Hypothesis (FTFA) (Schwartz and Sprouse, 1994; 1996), and the Missing Surface Inflection Hypothesis (Prévost and White, 2000). Under these theories, native-like grammatical representations can exist even in the absence of consistent accuracy in morpheme usage. The acquisition of native-like representations is theoretically possible for all learners, but transfer of L1 features is a source of difficulty in the initial stages of L2 acquisition. Additionally, some accounts, including the Missing Surface Inflection Hypothesis, have recently been discussed in terms of the processing constraints experienced by L2 learners. In other words, complete L2 representations may be present in the knowledge of the learner but may be underused due to heavy processing loads. As the current study does not report results of online processing, the focus will necessarily remain on whether or not learners show evidence of native-like sensitivity in offline tasks.

This section discusses the claims of the Full Transfer/Full Access Hypothesis, the Interpretability Hypothesis, and the Fundamental Difference Hypothesis as representative theories of full, partial, and no access approaches to UG, respectively.

### **3.2 Full Access: The Full Transfer/Full Access Theory**

Schwartz and Sprouse (1994, 1996) have proposed the Full Transfer/Full Access (FTFA) hypothesis, assuming a role for both UG and L1-specific representations. Under this theory, the initial state of L2 learning is considered to be the L1 grammar, but L1 parameter settings can be

abandoned in favor of UG-constrained settings as needed, in order to accommodate L2 input. Restructuring of the grammar occurs in cases where an L1-based analysis is not sufficient. It is important to note that restructuring may not always result in native-like L2 representations during development, since any combination of UG-constrained settings can be selected as long as the L2 input is accounted for. Furthermore, since the L2 input may be limited and L1 or other UG-constrained representations can continue to interfere with acquisition of native-like representations, “convergence on a grammar identical to that of a native speaker is not guaranteed” (White, 2003: 68). A number of studies have found that learners converge both qualitatively and quantitatively on native-like behavioral responses as their level of proficiency increases (e.g., White and Genesee, 1996; White, Valenzuela, Kozłowska-Macgregor, and Leung, 2004; Sagarra and Herschensohn, 2010), and in some cases, even lower-proficiency learners have exhibited native-like patterns in online measures (e.g., Renaud, 2010), providing support for theories claiming Full Access to UG.

As an example, White, Valenzuela, Kozłowska-MacGregor and Leung (2004) take advantage of the nominal agreement properties of Spanish to investigate the acquisition of L2 features. Recall that Spanish determiners and adjectives must agree in number and gender with their head noun. It is also possible in certain contexts in Spanish to refer to an object using only the determiner and adjective, a phenomenon known as N(oun)-drop:

- (14) Debo        comprar el        blanco?  
 Should<sub>1sg</sub> buy        the<sub>masc</sub> white<sub>masc</sub>?  
 “Should I buy the white one?”

Along with a control group of native Spanish speakers, White et al. (2004) tested adult learners of Spanish in Canada whose first language (L1) was either English or French and who were tested for proficiency level. In line with Full Transfer/Full Access, it was predicted that at



low proficiency, the L1 French group would outperform the L1 English group on gender agreement, since French but not English encodes gender in much the same way as Spanish. It was also predicted that the L1 English group would have less difficulty with number agreement than gender agreement since some determiners in English are marked for number. In line with Full Access, it was predicted that at higher proficiency levels, the L1 English group would perform as well on gender agreement as on number agreement and that differences between the L1 English and L1 French groups would disappear, perhaps converging on native speaker performance.

Tasks included production and comprehension tests. In one production task, the experimenter and participant were given the same cards depicting people, and the participant had to ask questions to figure out which person card the experimenter had chosen. Another production task involved describing what was going on in three different pictures. For the comprehension task, participants received a booklet with naturalistic conversations in which items were referred to through the use of only the determiner and the adjective, with appropriate number and gender endings. For each test sentence, the participants were asked to identify one of three pictures that was the referent in the sentence.

Across tasks, the intermediate and advanced groups of both L1s performed at above 90% accuracy on gender. Furthermore, neither the English- nor the French-speaking advanced learners differed from the Spanish native speakers in comprehension or production. White et al. argued that L2 gender features are acquirable due to full access to UG regardless of the fact that these were adult learners. A problem for Full Transfer emerges here, however, in that there were no significant differences between the English- and French-speaking groups at low proficiency, both of which performed less well on gender than on number. The authors propose that perhaps

these were not absolute beginners, a level of proficiency that is difficult to test but would provide more solid evidence for transfer. White (2003) points to research indicating the presence of L1 properties in the developing grammar (e.g., Haznedar, 1997), as well as the behavior of different L1 groups with regard to properties of the same L2 (White, 1985b, 1986; Yuan, 1998; Slabakova, 2000), as evidence in favor of Full Transfer.

### 3.3 Partial Access: The Interpretability Hypothesis

One of the most current formulations of the partial-access approach is the Interpretability Hypothesis of Tsimpli and Mastropavlou (2007). This theory is grounded in the Failed Functional Features Hypothesis (FFFH) (Hawkins and Chan, 1997), which claims that learners have indirect access to UG only through their L1 and so are unable to incorporate features of the L2 that are not instantiated in the L1. As an example, Hawkins and Chan (1997) provide evidence that Chinese-speaking learners of English are unable to acquire the ( $\pm wh$ ) feature on complementizers<sup>5</sup>. Gender features have been argued to provide the same difficulty for learners whose L1s do not instantiate grammatical gender. Hawkins (1998) investigated gender agreement in English-speaking learners of French. Participants were asked to describe a short animated film. Overall, the learners were relatively accurate but displayed errors that suggested the overgeneralization of one form or the other, leading to the conclusion that gender was not accurately represented in the learners' L2 grammar. Franceschina (2001) also analyzed production data, but in a case study of Martin, a single high-proficiency L2 Spanish learner. Gender errors were minimal (8.3% on determiners, 7% on adjectives), but significantly higher than number errors (0.5% on determiners, 2% on adjectives). Given that Martin makes no

---

<sup>5</sup> But see White and Juffs (1998), who found evidence that Chinese-speaking learners can acquire this feature.

mistakes in gender assignment to nouns, his errors seem indicative of persistent problems with uninterpretable gender features on determiners and adjectives.

Franceschina (2002) provides data from both production and comprehension tasks in L2 Spanish that are claimed to support FFFH. Subjects for each of the two experiments were native Spanish speakers and two groups of advanced learners, one group whose various native languages include a gender feature and another group whose native English language does not. In the production experiment, participants had to provide the missing pronoun in sentences where either a masculine or feminine accusative-marked pronoun or a neutral dative-marked pronoun should have occurred. An example sentence is provided in (15), where the blank should be filled with the masculine pronoun *los* ‘them’, referring to the masculine word in italics:

- (15) Los dos *enchufes*<sub>masc</sub> que compré estaban fallados. ¿Será posible cambiar\_\_\_ por unos nuevos?  
‘The two plugs I bought were faulty. Could I change (them) for new ones?’

(Adapted from Franceschina, 2002:79)

A chi-square test demonstrated a significant association ( $p < .05$ ) between L1 and the type of mistakes made, with gender mistakes being more common than mistakes in number or structural Case only for the genderless L1 group.

Participants in the comprehension experiment were presented with sentences containing gender-bearing pronouns or adjectives and had to choose the referents from among three choices where only one of the choices was of the appropriate gender, as in the following practice item:

- (16) Los *trajo* Martín y dijo que son para usted  
them<sub>masc</sub> brought Martín and said that are for you  
‘Martín brought them and said that they were for you’
- a. flores            b. joyas            c. chocolates  
flowers<sub>fem</sub>            jewels<sub>fem</sub>            chocolates<sub>masc</sub>

In this experiment, there was a significant difference between the native Spanish speakers and

the genderless L1 group ( $p < .01$ ), but not the gendered L1 group ( $p = .311$ ). There was a marginally significant difference between the two L1 groups ( $p = .063$ ). Despite the fact that the learners in the genderless L1 groups in these experiments were fairly accurate<sup>6</sup>, Franceschina takes these results to indicate that they have not been able to acquire the gender feature of the L2.

Taking FFFH as their starting point, Tsimpli and Mastropavlou (2007) propose the Interpretability Hypothesis, under which it is only uninterpretable features that are constrained by Critical Period effects. Accordingly, those features which are unique to the L2 and uninterpretable are expected to be the ones that pose a problem for learners even at high levels of proficiency. Following this line of inquiry, Tsimpli and Dimitrakopoulou (2007) investigated Greek speakers' acceptance of illicit resumptive pronouns after *wh*-extraction in English, as in (17-18). Resumptive pronouns are allowed following *wh*-extraction in Greek, except in the case of *ti* 'what' in object position, which is argued to be a reflex of its lack of (uninterpretable) *phi*-features.

- (17) **Grammatical and ungrammatical object extraction**  
a. Which student/Who do you think that Jane likes *t/\*him*?  
b. Which book/What do you remember that Peter read *t/\*it* carefully?
- (18) **Grammatical and ungrammatical subject extraction ( $\pm$ that)**  
a. Which politician/Who have you suggested *t/\*he/\*that-he* should not resign?  
b. Which party/What does John think *t/\*it/\*that-it* was very boring?

(Adapted from Tsimpli and Dimitrakopoulou, 2007: 227)

In line with the Interpretability Hypothesis, results revealed an L1-like acceptance of illicit resumptive pronouns, except in cases where interpretable features like animacy and d-linking provided independent constraints on the resumptive strategy. The researchers propose that

---

<sup>6</sup> In the production experiment, the genderless L1 group produced only 11 gender errors and demonstrated an overall mean accuracy rate of 85% compared to 89% for native speakers. In the comprehension experiment, the mean score for the genderless L1 group was 12.20/18 points, while the native speakers were only at 14.69/18 points.

interpretable features are available even in the early stages of L2 acquisition and “can improve L2 performance so that it approximates target output” (Tsimpli and Dimitrakopoulou, 2007: 236), but that the uninterpretable features necessary for the rejection of the resumptive strategy had not been acquired by the learners in their study.

Importantly, the Interpretability Hypothesis makes a distinction between features which are interpretable in the underlying Logical Form (LF) and those which may be interpretable when spelled out in Phonetic Form (PF) but are not interpretable at LF. Only the former are considered to be acquirable. This distinction is an important one in the investigation of gender agreement, since gender marking on adjectives and determiners in Spanish is realized phonetically (as described previously) and thus could be argued to be PF-interpretable. However, uninterpretable features are argued to be those whose role is restricted to syntactic derivations with no realization at LF, regardless of whether or not they are realized phonetically (Tsimpli and Mastropavlou, 2007; Tsimpli and Dimitrakopoulou, 2007).

### **3.4 No Access: The Fundamental Difference Hypothesis**

While the Interpretability Hypothesis and the Full Transfer/Full Access Hypothesis both make clear predictions regarding specific roles for L1 features, the Fundamental Difference Hypothesis (FDH) (Bley-Vroman, 1989, 1990), which can be classified as a no-access approach, makes more general claims about the success of adult L2 learners. Bley-Vroman (1990) points to such factors as fossilization, the need for negative evidence during development, greater variation in adult second-language (L2) learners’ success (ultimate attainment), varied learning paths and strategies, and affective aspects of learning as evidence for a fundamental difference between adult L2 acquisition and first-language (L1) acquisition in children. The original

formulation of the FDH claimed that post-Critical-Period learners, having no access to UG, succeed in L2 acquisition only to the extent that they are able to rely on general problem-solving skills and their knowledge of the L1. Under this theory, learners with near-native abilities in the L2 are “pathologically” successful learners who exhibit high non-linguistic cognitive abilities that allow them to overcome CP limitations on acquisition (Bley-Vroman, 1990: 7).

Although these cognitive abilities were originally posited as non-specific to language, subsequent studies have taken this to mean that language aptitude plays a significant role in L2 acquisition. DeKeyser (2000) was among the first to test the hypothesis that the lack of sensitivity to L2 grammaticality attested in post-Critical-Period learners (Johnson and Newport, 1989) might be modulated by language aptitude. Adapting Johnson and Newport’s study, which required L2 learners to make grammaticality judgments (GJ) on sentences involving English morphosyntactic features, he tested Hungarian-speaking learners of English who had acquired the L2 at an early age versus those who had acquired it as adults. In order to test the prediction of FDH that individual cognitive abilities might modulate performance, DeKeyser also included a test of verbal analytical ability in the native language of the participants, specifically a Hungarian version of the Words in Sentences portion of the Modern Language Aptitude Test (MLAT). DeKeyser hypothesized that the results would show an overall significant negative correlation between age of acquisition and performance, and that any overlap between the early and late learners’ performance would provide evidence for the role of aptitude. Additionally, he predicted that not all structures would be subject to Critical Period effects. In general, these predictions were borne out. A negative correlation between age of acquisition and GJ performance was present ( $r = -.63, p < .001$ ), but in the late learners this correlation was not significant, in line with the Critical Period Hypothesis (but see Bialystok (2002) for arguments

that DeKeyser's results indicate a general decline over the lifespan rather than CP effects). Very few late acquirers performed near the level of the early learners, but of those who did so, all but one also had above-average scores on the aptitude test. Additionally, aptitude scores were significantly correlated with overall GJ performance in late learners ( $r=.33, p<.05$ ) but aptitude was not a factor for the early learners ( $r=.07, p>.05$ ). Finally, the late learners were only successful on structures that DeKeyser argued to be quite salient in L2 input. DeKeyser concluded that aptitude, i.e., language analytical ability, plays a role in adult L2 acquisition but not in early L1 or L2 acquisition.

In a more recent version of the Fundamental Difference Hypothesis, Bley-Vroman (2009) proposes that the fundamental difference between child L1 and adult L2 acquisition is a heavier reliance on patches and shallow processing after the Critical Period, a theory which is heavily influenced by processing accounts in general, and for L2, the Shallow Structures Hypothesis of Clahsen and Felser (2006). The reformulated FDH predicts that for L2 properties not present in the L1, learners will only be successful in the case of those properties that can be acquired through the use of such shallow processing, resulting in nonconvergent (non-native-like) processing and/or representations. If this is the case, it may be profitable to further consider the specific shallow processing strategies available to learners when investigating the acquisition of the syntactic consequences of a particular feature.

### **3.5 Summary**

While each of the no-, partial- and full-access approaches presented here assumes a role for the L1 grammar, each makes a different prediction with regard to features that are unique to the L2. Under the Interpretability Hypothesis, features that are both uninterpretable and unique

to the L2 are unacquirable. For proponents of the Fundamental Difference Hypothesis, there is also a role for the L1, but processing strategies and/or individual differences are argued to be highly relevant. Due to the lack of access to UG, unique L2 features may only be acquirable given the right combination of salient input and learner characteristics, particularly with regard to domain-general processing. Finally, under a full-access approach, such L2 features are theoretically acquirable despite the challenges they may present in initial stages of learning, and performance will improve as proficiency increases. It should be noted that while the current study focuses on behavioral studies involving agreement, a future extension of the study will address agreement in L2 processing.

#### **4 Individual Differences in Second Language Acquisition**

In a discussion of arguments for the Fundamental Difference Hypothesis (Bley-Vroman, 1989, 1990, 2009) in the previous section, individual differences in cognitive abilities were argued to play a substantial role in the late acquisition of L2 features. Indeed, measures of verbal aptitude typically yield predictive validity coefficients in the .40 to .60 range when correlated with L2 proficiency tests and measures of classroom learner success (Carroll, 1981). With regard to the predictions of the FDH, however, two gaps in the literature can be noted. First, the relationship between individual differences like aptitude and acquisition of *specific* structures within the L2 has not been adequately investigated. Second, the original formulation of the FDH has been largely supported by evidence from studies claiming that verbal aptitude can explain apparent exceptions to the Critical Period Hypothesis. Bley-Vroman's (1990) original hypothesis, however, gave preference to domain-general problem-solving capacities that may be better tested outside the realm of language. These two issues are considered in the current study,



which investigates whether individual learner characteristics related to either verbal or nonverbal aptitude explain a percentage of the variability observed in measures of sensitivity to specific L2 features that differ in their similarity to the L1. Accordingly, this section presents background on both verbal and nonverbal aptitude. Here, verbal aptitude and language aptitude are used synonymously. Nonverbal aptitude is used to refer to nonverbal cognitive abilities, which have most often been tested in conjunction with verbal measures in studies of the relationship of intelligence to L2 outcomes. Therefore, nonverbal aptitude will necessarily be linked to the concept of general intelligence.

#### **4.1 Verbal Aptitude**

Verbal aptitude is claimed to be the single most effective predictor of L2 success (Skehan, 1989). It is normally-distributed within the population and is believed to be a fixed characteristic of the learner<sup>7</sup> and distinct from but subsumed under a general intelligence factor (Carroll, 1981; Skehan, 1986b, Wesche et al., 1982). Carroll (1965) defined aptitude in terms of phonemic coding ability, associative memory (ability to match L1 to L2 words), grammatical sensitivity, and inductive language learning ability. Carroll and Sapon's (1959) Modern Languages Aptitude Test was designed to test for these four skill sets and is still in use today. Skehan (1986a) adds that the ability to handle de-contextualized language is also crucial, but it is not clear whether this skill is directly related to the stable characteristic of aptitude, since it appears to improve with instruction even in adults.

---

<sup>7</sup> But see Sternberg (2002) for arguments that language aptitude can fluctuate based on the experiences of the learner and Thompson (2009), who finds higher language aptitude in multilinguals than in bilinguals. She leaves open the question of whether these results stem from greater desire on the part of high-aptitude learners to acquire more languages, or whether the experience of learning more than one language influences aptitude.

Recent studies have demonstrated the role of verbal aptitude in successful adult L2 acquisition. In a study of native speakers of Swedish and native-like learners of Swedish whose onset of acquisition began either in childhood or adulthood, Abrahamsson and Hyltenstam (2008) claim support for a Critical Period in general and for DeKeyser's (2000) conclusions regarding the role of aptitude in adult L2 acquisition, although they also found a role for aptitude in child L2 acquisition. With regard to grammaticality judgments, 7 out of 11 late (adult) learners and 13 out of 31 early (child) learners fell outside of the native speaker range on a grammaticality judgment task involving very difficult Swedish structures, despite the fact that native speakers of Swedish judged all of these learners to be virtually indistinguishable from native speakers. The measure of aptitude used in this study was an adapted version of the Swansea Language Aptitude Test (Meara, Milton, and Lorenzo-Dus, 2003). Early learners patterned with native speakers with regard to a normal distribution of aptitude scores, and there was a significant correlation between aptitude and GJ scores for early learners ( $r=.70$ ,  $p<.001$ ). On the other hand, there was no significant correlation between GJ scores and aptitude in the late learners, but this was probably due to the fact that the aptitude scores for all of the late learners were higher than the average of the early learners, leading the researchers to conclude that while aptitude is helpful in child L2 acquisition, it may be absolutely necessary (although not entirely sufficient) for native-like attainment in late L2 acquisition.

Harley and Hart (1997) also make substantial contributions to the investigation of aptitude and L2 acquisition. Participants in their study were L2 French learners in partial immersion programs in Canadian schools. One group of participants began receiving 50% of their instruction in French in Grade 1, and the other in Grade 7. Testing included a variety of L2 tasks targeting vocabulary, comprehension, and production, along with three aptitude tests

designed to assess both memory and analytical skills related to language aptitude. Harley and Hart found significant correlations between proficiency and memory skills for the early immersion group, but even greater significant correlations between proficiency and analytical skills in the late immersion group, suggesting that language analytical ability in particular may play a stronger role in adult L2 acquisition. Unfortunately, it is not clear whether the reported differences between memory and analytic ability can be distinguished on the basis of age of acquisition rather than learning environment. Additionally, Harley and Hart report that early learners do not have a higher aptitude resulting from early exposure to a second language. In other words, aptitude seems to be fixed by the time second language acquisition begins, even for relatively early learners, supporting the stability of the construct over time. Clearly, language aptitude is related to the success (or lack of success) of adult L2 acquisition, but whether or not that relationship constitutes evidence for the FDH is not so clear, since the role of nonverbal factors must also be investigated.

Another question that has not been sufficiently tested is whether the correlation of aptitude and learner success in general is observable for specific structures, and whether L1/L2 similarities have any bearing on that correlation. If individual differences like aptitude can overcome the effects of age of acquisition, as is claimed under the Fundamental Difference Hypothesis, then the effects of aptitude should be observable even, and maybe especially, for features that are unique to the L2. Work in this area is sparse. Recall that DeKeyser (2000) predicted that not all structures would reveal age-related effects. He tested this prediction by analyzing correlations between age of arrival in the U.S. and performance on particular test items. Test structures for which there was a significant correlation ( $p < .01$ ) included auxiliaries, determiners, *wh*-questions, plurals, subcategorization, and adverb placement, which were not

compared to the L1 Hungarian. The structures that demonstrated no significant correlation with age of arrival ( $p > .05$ ) were argued to be highly salient either because the faulty use of these structures in production results in salient errors (word order), in particular on lexical verbs in sentence-initial position (subject-verb inversion, do-support in yes-no questions), or with regard to the likelihood that learners would be corrected by native speakers (gender on pronouns). It should be noted that DeKeyser did not test any direct correlation between performance on test items and aptitude test scores. However, overall performance was correlated with aptitude for late learners. Interestingly, with regard to the question of L1/L2 differences, the Hungarian L1 of these participants differs from the L2 English with respect to the structures for which learners were generally successful. Based on DeKeyser's overall results, then, it could be argued that both aptitude and salience (which is not explicitly defined by DeKeyser) can overcome L1/L2 differences even for post-CP learners. There is reason to believe, then, that aptitude effects could be found in adult learners for at least less salient structures. The current study employs number and gender agreement paradigms that do not appear to be salient in the input. Regardless, any attempt to test for aptitude effects on particular grammatical structures, as is undertaken here, should be considered exploratory in nature. In general, most studies of aptitude effects usually employ broad tests of L2 proficiency, without regard to the effects of aptitude for specific structures, either similar to the L1 or unique to the L2.

## **4.2 Nonverbal Aptitude**

The relationships of verbal and nonverbal aptitude to L2 acquisition have rarely been tested as separable constructs. In many cases where verbal aptitude and intelligence have both been tested for correlations with L2 outcomes, researchers often rely on tests of intelligence that

themselves include verbal measures without a factor analysis of the resulting correlations. For example, Robinson (2002) tested the relationship of intelligence, aptitude, and working memory (WM) to the L2 performance of native speakers of Japanese (n=160) under three conditions: implicit and explicit instruction in an artificial grammar, and incidental learning of aspects of Samoan grammar. Among other predictions, he hypothesized that individual differences in intelligence and aptitude would be most influential in explicit learning conditions. Aptitude was measured by Sasaki's (1996) Language Aptitude Battery for the Japanese (LABJ), and intelligence by the Short Form of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) in Japanese, which includes verbal, arithmetic, and block design subtests. While there were differences between aptitude and intelligence when correlated with learning under Implicit and Incidental conditions, both aptitude and intelligence were moderately correlated with a measure of Explicit learning<sup>8</sup>. The use of the WAIS-R test that included verbal tasks leaves room for the possibility that verbal intelligence was the underlying factor in the correlation between L2 performance in the Explicit group and general intelligence. Other investigations of general intelligence and verbal aptitude in L2 studies show the same trend of including verbal measures of intelligence. For example, Sasaki's (1996) construct of a general cognitive factor contributing to L2 proficiency lumps together the three factors of language aptitude, verbal intelligence and reasoning abilities, which were found to be correlated but not identical to an L2 proficiency construct in the best-fit model uncovered by Sasaki's structural equation modeling.

Research into other cognitive factors points to the fact that it might be reasonable to separate nonverbal and verbal factors in the investigation of L2 performance. Consider working

---

<sup>8</sup> Aptitude scores demonstrated no relationship to either Implicit or Incidental learning in immediate post-tests, but a 6-month post-test using guided production revealed a significant correlation with LABJ scores for participants in the Incidental group ( $r=.56, p<.01$ ). The WAIS-R measure of intelligence demonstrated a negative correlation with results for the Implicit group and no correlation in the Incidental group.

memory (WM), which is argued to play an important role in cognition generally and also in language processing (e.g., Miyake and Friedman, 1998). It is not clear that the WM resources recruited for language processing and those involved in more domain-general processes are the same because studies investigating the relationship of WM to successful L2 acquisition have largely used verbal measures of WM, again, not necessarily being concerned with the question of domain-specificity (e.g., McDonald, 2006; Sagarra, 2007; Sagarra and Herschensohn, 2010; Juffs, 2004; also see a description of WM measures in L2 studies presented in Ortega, 2009). For example, Sagarra (2007) used a reading span test to investigate the role of WM in the L2 processing of gender agreement. She found that beginning learners of Spanish (n=209) whose L1 English does not incorporate a gender feature were generally insensitive to gender agreement violations on adjectives, but taken together, the reading times and accuracy rates of the high-span individuals demonstrated a developing sensitivity to violations. Additionally, there is evidence that even L1-based measures of working memory may not completely capture individual differences affecting L2 processing. Service et al. (2002) reported that learners in their study had lower WM capacity in their L2 than in their L1, but that L2 but not L1 WM capacity correlated with L2 comprehension.

In order to test the relationship between verbal aptitude and intelligence, Wesche et al. (1982) conducted a factor analysis of the scores of 793 participants on the MLAT and the Primary Mental Abilities Test (PMA; Thurstone and Thurstone, 1965), a test of cognitive abilities that includes the following subtests: Verbal Meaning, Number Facility, Reasoning, and Spatial Relations. Unsurprisingly, they found evidence for a hierarchical model in which specific language-related skills are subsumed under a more general ability. They conclude that while verbal aptitude and general intelligence are not completely distinct in that they rely on a

general ability for abstract reasoning using symbols, verbal aptitude measures tap abilities that are specific to language learning and distinct from the cognitive abilities assessed by the PMA. They do not, however, test the relationship of verbal aptitude and intelligence to L2 outcomes. The only study where subtests of the MLAT and the PMA were examined in a factor analysis along with L2 outcomes was a study of 96 teen learners by Gardner and Lambert (1965), who examined 14 measures of French achievement, the MLAT, and an earlier version of the PMA including an additional subtest of Word Fluency (Thurstone and Thurstone, 1941). Like Wesche et al. (1982), they had also found high degrees of correlation between the MLAT and the PMA. However, in their study the MLAT subtests and the PMA subtests collapsed into separate factors. Indeed, four of the PMA measures loaded onto a factor that Gardner and Lambert identified as intelligence, and the researchers concluded that the four measures themselves “share little in common with measures of French achievement or language aptitude” (1965: 198-199).

Only one study of the relationship between proficiency and intelligence can be found that involves a nonverbal measure of intelligence. Flahive (1980) tested 20 learners of English from seven different language backgrounds to determine whether intelligence or proficiency correlated more precisely with scores on three reading comprehension tests: a multiple choice test, a paraphrase recognition test, and a cloze test. The measure of intelligence used in the study was the Raven’s Standard Progressive Matrices, which was chosen for its nonverbal format due to the varied language backgrounds of the participants. Proficiency was measured by scores on the Test of English as a Foreign Language (TOEFL). The RAVEN and the TOEFL measures were intercorrelated ( $r=.61, p<.01$ ). With regard to the reading comprehension tasks, the Raven was found to correlate most strongly with the multiple choice test ( $r=.84, p<.01$ ), but also with the other two tests (paraphrase:  $r=.68, p<.01$ ; cloze:  $r=.61, p<.01$ ). The opposite pattern was true of

the TOEFL even though it was also correlated with all three measures (multiple choice:  $r=.59$ ,  $p<.01$ ; paraphrase:  $r=.84$ ,  $p<.01$ ; cloze:  $r=.75$ ,  $p<.01$ ). Flahive concludes that reading comprehension tests, especially the multiple choice test used in his study, may test more than just reading comprehension. The main point here, though, is that a nonverbal measure of intelligence was found to be correlated with general measures of L2 performance. One factor that may limit the interpretation of these results, however, is that the learners in Flahive's study were not normally-distributed in terms of intelligence since their mean score on the Raven placed them at the 75<sup>th</sup> percentile. It is unknown whether the RAVEN would correlate with reading comprehension measures in a more normalized group of learners.

As pointed out by Skehan (1998), case studies of exceptional post-puberty learners may also provide evidence that intelligence is not a key factor in successful language learning. Opler (1989) and Novoa et al. (1988) present a case study of a learner who had rapidly acquired five foreign languages in informal or mixed formal/informal environments and was judged to be near-native by native speakers. Schneiderman and Desmarais (1988a, b) reported on two learners who were at native-like proficiency in three foreign languages, as judged by native speakers and grammaticality judgment tests. All of these learners demonstrated slightly above-average but not exceptional intelligence quotients (IQ). Additionally, Smith and Tsimpli (1995) studied a man who was gifted in his knowledge of ten languages, but was mentally retarded due to brain damage and did not score well on nonverbal measures in standard IQ tests.

In sum, there is no consensus on whether or not domain-general capacities are key to successful L2 acquisition. The only study to clearly demonstrate a relationship between L2 performance and a nonverbal measure of intelligence is Flahive (1980), whose participants were clearly of above average intelligence to begin with. This does not appear to be the case with



learners in a variety of case studies cited in Skehan (1998). The sophisticated analysis of Wesche et al. (1982) found that a general intelligence factor subsumes more language-specific aptitude measures, but those aptitude measures may correlate with L2 achievement without a corresponding relationship between L2 achievement and measures of intelligence (Gardner and Lambert, 1965). In other research on cognitive abilities, whether broadly defined (e.g., intelligence) or specific (e.g., working memory), the construct of domain-general capacities has not always been so clearly distinguished from domain-specific or verbal capacities. The question remains, then, as to whether there is independent evidence that general cognitive factors are correlated with L2 performance *outside the realm of verbal measures*. A secondary question to be explored in the current study is whether L1/L2 differences modulate the role of domain-general and domain-specific cognitive factors in L2 performance on particular structures.

## **5 The Current Study**

### **5.1 Research Questions**

Based on the literature reviewed above, at least two general claims can be made regarding the question of what the learner brings to second language acquisition. First, all of the theories investigated here (FTFA, Interpretability, FDH) have in common the supposition that the initial state of L2 acquisition at least involves knowledge of the L1, but they differ in their approaches to unique and uninterpretable L2 features. Secondly, it seems to be the case that at least some individual differences have a role to play in adult L2 acquisition regardless of theoretical approach, but the nature of that role and the specific individual differences impacting acquisition have not been determined. The aim of the current study is to investigate the impact of both of these factors – L1/L2 differences with regard to uninterpretable features, and

individual differences in verbal and nonverbal aptitude - on the acquisition of specific L2 structures involving morphosyntactic agreement, as measured by grammaticality judgments following word-by-word presentation of grammatical and ungrammatical sentences involving number and gender agreement. Research questions can thus be divided into two areas of interest: learner sensitivity to agreement violations, and the role of individual differences. The first two research questions have to do with sensitivity to agreement violations, with a particular focus in RQ2 on differences between violations involving features that are either present or absent in the L1:

RQ1: Do learners show the same pattern of sensitivity to violations as native speakers?

RQ2: Are learners equally sensitive to agreement violations for number features that are present in the L1 and for gender features that are unique to the L2?

The second area of interest brings in questions regarding the impact of individual differences and any interaction they may have with L1/L2 differences. These questions are exploratory in nature:

RQ3: Does verbal aptitude modulate sensitivity to agreement violations and/or to number and gender agreement differentially?

RQ4: Does nonverbal aptitude modulate sensitivity to agreement violations and/or to number and gender agreement differentially?

## **5.2 Predictions**

### **5.2.1 Learner Sensitivity to Agreement Violations**

The L2 under investigation here is Spanish [+num, +gen], which requires agreement between nouns and their modifiers in terms of both number and gender, as described above. The

learners recruited for this study were native speakers of English [+num, -gen], which includes a number feature that triggers agreement between nouns and verbs and within the DP between nouns and demonstratives. However, English is widely recognized as not instantiating grammatical gender agreement. With regard to the first two research questions, then, each of the L2 acquisition theories under investigation predicts that learners will be sensitive to number agreement violations in L2 Spanish since the number feature is present in their L1; therefore, learners should pattern with native speakers in this regard, despite the fact that number agreement does not occur on adjectival predicates in English. However, the three theories differ in their predictions regarding sensitivity to gender agreement violations. The Interpretability Hypothesis predicts that learners whose L1 English does not instantiate uninterpretable gender features will not be able to acquire the LF-uninterpretable gender feature on Spanish adjectives even though this feature is spelled out at PF, at least for canonical nouns. Thus, learners are expected to diverge from native speakers in that learners will not demonstrate sensitivity to gender agreement violations. On the other hand, the Fundamental Difference Hypothesis and Full Transfer/Full Access allow that at least a few learners could demonstrate sensitivity to gender agreement. If the FTFA model is on the right track, gender should ultimately be acquirable, and sensitivity to gender agreement should be largely a function of proficiency. On the other hand, the FDH predicts that only isolated cases of native-like responses to gender agreement would be observed, if any, and that their success would be attributable to exceptional ability, which will be tested as described below.

## **5.2.2 Aptitude**

Predictions are given below for research questions 3 and 4, which investigate the relationship of verbal and nonverbal aptitude to L2 sensitivity to agreement violations. While neither FTFA nor the Interpretability Hypothesis make specific predictions regarding correlations between measures of sensitivity to agreement violations and either verbal or nonverbal aptitude, such correlations might at least be expected under the Fundamental Difference Hypothesis.

### **5.2.2.1 Verbal Aptitude**

The results of the studies reviewed above (DeKeyser, 2000; Abrahamsson and Hyltenstam, 2008; Harley and Hart, 1997) show clear effects for verbal aptitude when it comes to L2 performance, at least for adult learners. DeKeyser's (2000) analysis of age-related effects for particular structures suggests that late learners' performance on less salient structures may in fact correlate with age of arrival. Furthermore, the overall responses of late learners in these studies demonstrated a significant correlation with aptitude scores, particularly with regard to language analytic ability (Harley and Hart, 1997; DeKeyser, 2000). Given these relationships, it is predicted here that in late learners, verbal aptitude (at least as it measures language analytic ability) might be correlated with measures of sensitivity to agreement violations, which are not generally salient. Since individual differences like aptitude may overcome age effects that inhibit acquisition of properties unique to the L2, the relationship of aptitude to sensitivity to violations will also be explored separately for number and gender, since these two features differ in their similarity to the L1.

### **5.2.2.2 Nonverbal Aptitude**

While several cognitive constructs arguably related to domain-general processing have been tapped in second-language research, the majority of these constructs are operationalized in the literature using at least partially domain-specific measures. The primary question raised here has been whether or not a general measure of nonverbal intelligence or aptitude outside of the verbal domain can provide evidence for the role of domain-general cognitive factors in L2 processing. The original formulation of the Fundamental Difference Hypothesis (Bley-Vroman, 1989, 1990) suggests that this should be the case, leading to the prediction that nonverbal aptitude scores will be correlated with measures of sensitivity to agreement violations. However, while full-access theories like FTFA do not make predictions in this regard, it could be argued that domain-specific as opposed to domain-general capacities would be more likely to be correlated with L2 performance if L2 acquisition is guided by UG; in this case, verbal aptitude may correlate more strongly than nonverbal aptitude with measures of sensitivity to agreement features. Just as with verbal aptitude, the relationship of nonverbal aptitude to sensitivity to violations will be explored separately for number and gender.

## **6 Methods**

### **6.1 Participants**

Participants were recruited for an ERP study where they were asked to make grammaticality judgments during simultaneous EEG recording (ERP results to be reported in a future study). A total of 34 people participated in the study: 22 English-speaking learners of Spanish (10 males) and a control group of 12 native Spanish speakers (3 male). Two L2 participants were excluded from analysis: one participant in the learner group indicated that she

had mistakenly rejected half of the stimuli, thinking that the structure used in the wrap-up material at the end of the sentence was not a grammatical structure in Spanish. Since that material was not the focus of investigation here, data from this participant were not included in the analysis. Data from another participant were excluded due to substantial previous exposure to a Romance language other than Spanish which also instantiates gender and number agreement. Due to the simultaneous collection of EEG data, all of the participants who were recruited for the study were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), had normal or corrected-to-normal vision, and reported no neurological impairment. All were financially compensated for their participation.

Participants in the learner group were students at the University of Kansas who were recruited from fourth-semester Spanish classes by means of a posting on the class website. A language background survey confirmed that all learners spoke English as their native language. The survey also collected information on the first age of exposure to Spanish, the amount of time each learner had spent in a Spanish-speaking country, and any previous exposure they may have had to a Romance language other than Spanish. A summary of findings is given in Table 1. It should be noted that the learner reporting an age of exposure of 5 years old was only taught some vocabulary by her English-speaking parents. This was the only learner reporting an age of exposure below 9 years of age, so it was concluded that all learners could be considered late learners.

**Table 1. Profile of L2 group characteristics**

	Age, in years (n=20)	Years of Spanish study <sup>a</sup> (n=20)	Age of exposure to Spanish, in years (n=17 <sup>b</sup> )
Mean	19.35	4.80	13.59
Standard Deviation	.93	1.96	3.32
Range	18-21	1.5-7.5	5-18

<sup>a</sup> This response includes classes taken in junior high and high school.

<sup>b</sup> Some participants did not give a numeric response.

Given the fact that the learners involved in this study were enrolled in Spanish classes at the time of testing, every effort was made to conclude the testing quickly so that no learner had an advantage over any other learner in this study. However, it was not expected that participants who were tested on a particular construct 2-3 weeks later than other participants would have received instruction on that specific construct since the third- and fourth-semester classes at the University of Kansas are taught using an online curriculum designed by the department (the *Acceso* Project, <http://www2.ku.edu/~spanish/acceso/>), which uses natural-language materials, with a focus on communication for functional tasks. Teaching assistants are also trained to avoid grammar explanations. Additionally, all but two participants were tested for Spanish proficiency within three weeks of one another. However, since the overall testing of the learners lasted for seven weeks, statistical tests were performed in order to ensure that there was no effect for the point in the semester when individual learners were tested. While there was a significant correlation between test date and proficiency scores, the correlation was negative,  $r=-.628$ ,  $p=.003$ . No significant correlation of test date was found for any grammaticality judgment measures.

A control group of twelve native Spanish speakers were also tested. These participants were recruited in one of three ways: by word of mouth, by visiting the Spanish Round Table discussions at the University of Kansas, and by emailing native Spanish speakers who had

previously registered themselves in a database of possible participants for research studies. A background questionnaire confirmed that all of these participants were native speakers of Spanish. With the exception of one participant who was born in the United States, all the native speakers were from Peru, Bolivia, or Costa Rica and were exposed to English after 9 years of age. One of these participants was visiting the United States; all others were living in the United States at the time of testing. For all native speakers, the average age at testing, length of residence in the United States, and age of exposure to English are provided in Table 2.

**Table 2. Characteristics of native speakers, in years**

	Age at testing (n=12)	Length of Residence in the U.S. (n=12)	Age of exposure to English (n=11 <sup>a</sup> )
Mean	28	3.9	13.7
Standard Deviation	7.92	4.37	5.87
Range	18-41	0-14.5	0-21

<sup>a</sup> One participant did not give a numeric response.

## 6.2 Materials

### 6.2.1 Stimuli

A grammaticality judgment task was administered to all participants during EEG recording as described under Procedures below. The stimuli were 120 triplet sets of sentences adapted from a subset of the stimuli in a previous event-related-potential (ERP) study of morphosyntactic agreement in native speakers of Spanish (Alemán Bañon, 2010: Exp. 2). The grammatical sentence in each triplet contained an adjectival predicate agreeing in number and gender with the subject of the sentence (19a), which was always singular. The ungrammatical versions of these sentences varied the morphological suffixes on the adjectives to provide a mismatch in either gender (19b) or number (19c) between the subject (underlined here) and the predicate (**in bold**).



(19) a. **Grammatical Number and Gender (NUM/GEN)**

La isla es **rocosa** y la península también.  
The<sub>sg.fem</sub> island<sub>sg.fem</sub> is<sub>3sg</sub> rocky<sub>sg.fem</sub> and the peninsula too.  
'The island is rocky and the peninsula too.'

b. **Ungrammatical Number (\*NUM)**

La isla es **\*rocosas** y la península también.  
The<sub>sg.fem</sub> island<sub>sg.fem</sub> is<sub>3sg</sub> rocky<sub>pl.fem</sub> and the peninsula too.

c. **Ungrammatical Gender (\*GEN)**

La isla es **\*rocoso** y la península también.  
The<sub>sg.fem</sub> island<sub>sg.fem</sub> is<sub>3sg</sub> rocky<sub>sg.masc</sub> and the peninsula too.

All sentences were presented in the present tense. With regard to gender, approximately half of the sentence subjects were feminine and the rest were masculine. Only nouns displaying canonical gender markings were used in the subject NPs, and all were inanimate so as to avoid bias toward natural gender. A total of 60 adjectives were used, each appearing in two different triplets. It should also be noted that none of the sentences contained violations of both number and gender agreement. In an attempt to eliminate any effects of unfamiliar vocabulary, efforts were made to ensure that the nouns and adjectives used in the stimuli were understandable by learners, either because they appeared in textbooks or because they are cognates with English. Because the adjectives appeared in the critical region, they were included in a computerized vocabulary test completed by learners after the experimental task was done. The triplet sets of target sentences can be found in Appendix 1.

Target sentences from a second experiment that also involved a set of 120 triplets were used as fillers. Since two-thirds of the experimental stimuli for the current study were ungrammatical, it was necessary to balance the overall number of grammatical and ungrammatical stimuli as seen in (20). All of the ungrammatical sentences (as in 20b) contained a violation of number agreement on a verb that was created by providing a plural suffix in a singular context. The sets of sentences used as fillers can be seen in Appendix 2.

(20) a. **Filler Type A: Grammatical Subject-Verb Agreement**

La viajera agotada descansa en el hotel.  
The<sub>3sg</sub> traveler<sub>3sg</sub> exhausted<sub>3sg</sub> rests<sub>3sg</sub> in the hotel.  
'The exhausted traveler rests in the hotel.'

b. **Filler Type B: Ungrammatical Subject-Verb Agreement**

La viajera agotada \*descansan en el hotel.  
The<sub>3sg</sub> traveler<sub>3sg</sub> exhausted<sub>3sg</sub> rests<sub>3pl</sub> in the hotel.

c. **Filler Type C: Grammatical Use of Copula**

La viajera agotada está en el hotel.  
The<sub>3sg</sub> traveler<sub>3sg</sub> exhausted<sub>3sg</sub> is<sub>3sg</sub> in the hotel.  
'The exhausted traveler is in the hotel.'

The 240 sentence sets (120 targets, 120 fillers) were counter-balanced across three lists in a Latin Square design such that participants read 40 tokens of each condition and never saw more than one version of each target item. Target sentences were mixed with fillers and presented in random order. Presentation of grammatical and ungrammatical sentences was therefore randomized across each list. To ensure that there had been no advantage for participants receiving any one list over the others, a one-way ANOVA was conducted following the experiment to compare accuracy rates for each list. No significant differences were found ( $F=1.018, p=.373$ )<sup>9</sup>.

### 6.2.2 Gender Assignment Task

All participants also completed an offline task in order to ensure that they were able to correctly assign the gender of each of the 60 nouns that appeared in the subject NP in the target stimuli. In this task, each noun was presented briefly in random order, followed by a prompt to which participants had to respond by mouse click, identifying the appropriate definite article to use before each noun - “El” for masculine nouns, and “La” for feminine nouns.

---

<sup>9</sup> Test results for list differences by condition were also non-significant.

### **6.2.3 Vocabulary Recognition Task**

In addition to the two tests described above, which were administered to both the native speaker controls and the learners, the learners were also required to take the remainder of the tests to be described in this section. A vocabulary recognition task was conducted in order to ensure that learners were familiar with the critical words used in the grammaticality judgment task. The task included all of the adjectives from the target stimuli like the one in (19) above, as well as all of the lexical verbs from the fillers like the one in (20), for a total of 120 items. All items were presented in random order. Learners responded by mouse click to choose the correct alternative between two possible English meanings for each Spanish word.

### **6.2.4 Verbal Aptitude: The Modern Languages Aptitude Test**

The measure of verbal aptitude chosen for this study was the short form of the Modern Language Aptitude Test (MLAT) (Carroll and Sapon, 1959), which consists of Parts 3, 4, and 5 of the longer version of the test. The MLAT3 (Spelling Clues) is a multiple-choice test that asks participants to select appropriate synonyms of words that are spelled as they are pronounced rather than by conventional orthography (e.g., an appropriate synonym for *ritn* might be ‘printed’, since *ritn* can represent the pronunciation of *written*). This test is designed to assess phonetic coding ability and memory for vocabulary. There are 50 items in this section, but only 5 minutes are allowed. The MLAT4 (Words in Sentences) is a test of grammatical sensitivity or a broader language analytic ability (Skehan, 1998), including sensitivity to grammatical roles and the ability to make analogies at a grammatical level. Participants must select from among several underlined words in each test sentence the one choice that functions in the same way as the underlined word in an example sentence. For example, if the subject is underlined in the

example sentence, the participant's correct response would also indicate the subject of the test sentence out of the multiple choices of underlined words. Participants are allowed 20 minutes to complete 45 items. The MLAT5 (Paired Associates) requires the rapid learning of a list of 24 vocabulary words in an adapted language, along with their associated English meanings. Following a 2-minute period of vocabulary memorization and a subsequent 2-minute practice period, participants have 4 minutes to choose the correct meaning from multiple choices given for each of the 24 items. The MLAT5 indexes associative memory and requires the storage and retrieval of a large amount of material in a short period of time.

#### **6.2.5 Nonverbal Aptitude: Raven's Advanced Progressive Matrices**

In order to contrast the contributions of skills that are domain-specific to language with those that are more domain-general, a test of nonverbal aptitude was also conducted. The Raven's Advanced Progressive Matrices (RAVEN) (Raven, 1965) is a multiple-choice test of nonverbal intelligence and reasoning skills, including the ability to decompose complex problems, search for rules and manage those rules in working memory (Carpenter, Just, and Shell, 1990). In a multidimensional scaling analysis by Snow, Kyllonen, and Marshalek (1984), where a wide variety of domain-general and domain-specific tests were placed in concentric circles representing the closeness of correlations among the tests, the RAVEN occupied the most central position, demonstrating that it is an optimal test for measuring domain-general reasoning. Each of the 12 practice items in Set I of the RAVEN and the 36 test items in Set II consists of a visual pattern with a piece missing, followed by an array of choices of patterns to fill the missing area in the test pattern. The test set can be administered as timed (usually for 40 minutes) or untimed; here it was administered in a fairly short amount of time due to time constraints and the

possibility of participant fatigue during the testing session. The time allowed for the practice set was five minutes, and twenty minutes were allowed for the test set<sup>10</sup>.

### **6.2.6 Spanish Proficiency Test**

A short, written Spanish proficiency test was administered to learners that included the vocabulary section of the MLA Cooperative Foreign Language test (Educational Testing Service, Princeton, N.J.) and a cloze section from the Diploma de Español como Lengua Extranjera (DELE) test (Spanish Embassy, Washington D.C.). This test was chosen because it had previously been used in several other studies of L2 learners and heritage speakers (White et al., 2004; McCarthy, 2008; Montrul and Slabakova, 2003), allowing the comparison of learner groups across studies. The test includes a total of 50 multiple choice items. A score below 30 is considered to indicate low proficiency, intermediate proficiency ranges from 30-49 points, and advanced proficiency is reflected by scores at 40 points or higher. In both sections of the test, participants had to choose the word that correctly fills the blank in a sentence or passage from among the choices given. The items in this test involve knowledge of specific vocabulary, use of functional words, and/or specific grammatical knowledge. The test was untimed, but all participants finished it in about 30 minutes.

---

<sup>10</sup> The decision to shorten the amount of time allowed for the RAVEN was made after piloting the test on 13 native speakers of English recruited for extra credit in a linguistics class. Recorded data included the numbers of test items in Set II completed at 20 minutes (Mean=23.82, SD=4.21) and at 30 minutes (Mean=26.18, SD=5.00). While a paired-samples t-test did show that the scores after 30 minutes were significantly higher than the scores after 20 minutes ( $F=.580, p<.001$ ), there was a high level of correlation between the two sets of scores ( $r=.94$ ).

## 6.3 Procedures

### 6.3.1 Learners

The learners involved in the study attended two experimental sessions that lasted a total of approximately five hours. The first session was administered in the Second Language Acquisition Laboratory at the University of Kansas. During this session, learners provided informed consent for the entire study and completed a language background survey. The MLAT, RAVEN, and proficiency tests were then administered as described below, with breaks between each test. Before leaving the test session, learners scheduled their second session with the researcher. Participants were paid for their participation at the end of each session.

Given the intense nature of this testing session, the order of the three tests was varied between participants in order to eliminate the effects of participant fatigue on any single set of test scores. Half of the learners started their testing session with the Spanish proficiency test, the other half ended with it. In order to avoid any confound introduced by the order of testing languages, the Spanish test was not given between the other two tests. The order of the MLAT and RAVEN was also randomized, resulting in four different test orders, as presented in Table 3.

**Table 3. Test order between groups of L2 participants**

	Group 1 (n=4)	Group 2 (n=6)	Group 3 (n=5)	Group 4 (n=5)
Test order	1 RAVEN	MLAT	Proficiency	Proficiency
	2 MLAT	RAVEN	RAVEN	MLAT
	3 Proficiency	Proficiency	MLAT	RAVEN

A one-way analysis of variance (ANOVA) was conducted after the testing to explore possible differences in test outcomes based on test order. There was a significant effect for Test Order for the MLAT3,  $F(3,16)=6.321, p=.005$ . Post-hoc tests using Bonferroni adjustments for multiple comparisons showed that testing group 2 was found to have significantly higher scores

on the MLAT3 than Testing Groups 1 ( $p=.006$ ) and 4 ( $p=.047$ ), indicating a possible advantage to taking the MLAT3 first. However, the number of participants in each group is relatively low, so the possibility that significant differences could be produced by any one participant's score in any group was explored. For the MLAT3 results, one high score in Group 2 and one particularly low score in Group 4 may have produced the test order effect. For the MLAT4, a marginal effect of test order was found,  $F(3,16)=3.053$ ,  $p=.059$ . As demonstrated in post-hoc tests, this effect was driven by a marginally lower mean score for the MLAT4 in Testing Group 3 as compared to Group 1 ( $p=.058$ ), indicating a possible disadvantage for taking the MLAT last. However, one especially high score is present in Group 1. Accordingly, no statistical adjustments were made for test order in subsequent testing.

The second experimental session was conducted in the Neurolinguistics and Language Processing Laboratory at the University of Kansas. During this session, each learner was tested using the experimental stimuli in a grammaticality judgment task during simultaneous EEG recording, described below. The Edinburgh Handedness Inventory was completed first, in order to verify right-handedness (Oldfield, 1971). Participants were then fitted with an electrode cap and additional electrodes above, below, and to the outside of each eye and behind each ear. Participants were seated comfortably in front of a CRT monitor in a dimly-lit, sound-attenuated experiment room. The Paradigm experimental control system designed by Perception Research Systems, Inc. (Tagliaferri, 2005) was employed for randomized stimulus presentation and EEG interfacing. In order to familiarize the participant with the task, each recording session began with a practice session of nine trials, which included items targeting areas of grammar not investigated in the experiment. No words appeared in both the practice and experimental items. Participants received feedback after each of the first three practice items. For all practice and

experimental items, each trial was preceded by a blank screen for 500ms, allowing the participant time to blink, followed by a fixation cross for 500ms and then a 300-ms pause, after which the stimulus sentence was presented word-by-word in an RSVP (Rapid Serial Visual Presentation) paradigm. Each word appeared in black text on a dark gray screen for 450ms, with a pause between each word lasting 300ms. A pause of 1000ms followed the final word of the sentence before a response prompt appeared on the screen, asking participants to indicate whether the sentence was “Bien” (good) or “Mal” (bad). Participants were instructed to respond by button push for “Bien” if they felt the Spanish sentence was grammatical and “Mal” if they felt it was ungrammatical. The accuracy of the participant’s response was recorded by the software. After every 40 trials, participants were prompted to take a break. The task averaged around 55 minutes for learners to complete, including breaks.

Following the experimental task, the electrode cap was removed. Participants were given a break, followed by a short computerized test that included both the gender assignment and the vocabulary recognition tasks, in that order. Instructions and six practice items immediately preceded each task. No feedback was provided for the practice items. Again, the test items were presented in random order and responses were recorded by the Paradigm experimental control software (Tagliaferri, 2005). For the gender assignment task, each trial was preceded by a 300-ms pause and 500-ms fixation cross. In order to mimic the conditions of the grammaticality judgment task, each Spanish noun was presented for 450ms in black text on a dark gray screen. Presentation of the noun was followed by a prompt that presented the masculine determiner “El” on the left of the screen and the feminine determiner “La” on the right. Participants responded by mouse click in the area of the determiner of choice. For each item in the vocabulary recognition task, a Spanish word was presented in lower-case letters, with two possible English



translations for that word in capital letters below it. Participants were asked to use the mouse to select the appropriate translation for each item. A 1000-ms pause occurred between each trial. Participants finished these computerized tasks in approximately 10 minutes.

The total time for this session was around 2½ hours, including 10 minutes for paperwork, 45-60 minutes to set the electrode cap, approximately one hour for the EEG task, a break, and then another 10 minutes for the gender assignment and vocabulary recognition tasks.

### **6.3.2 Native Speakers**

The native speakers in this study were not required to take the aptitude and proficiency tests, so they only needed to attend one session. At the beginning of this session, they provided informed consent and then completed a background questionnaire and the Edinburgh Handedness Inventory (Oldfield, 1971). After being fitted with the electrode cap, they completed the grammaticality judgment test, which was administered in the same way as described above for the learner group. Including breaks, native speakers averaged 45 minutes to complete the grammaticality judgment task. After a short break, they were asked to complete the gender assignment task in an effort to ensure the validity of all the items included in that task. The total amount of time for this session was approximately 2½ hours. Each participant was compensated financially for their time.

## **7 Results**

In keeping with the dual nature of this investigation, this section is presented in two parts. In the first part, the descriptive results of the experimental tasks are presented, followed by planned analyses of effects for Group, Grammaticality, and Violation Type using a series of

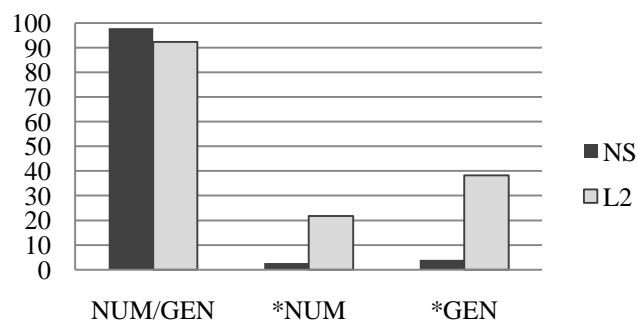
repeated-measures analysis of variance (ANOVA). The second part of this section presents analyses of the relationships between the experimental results and measures of verbal and nonverbal aptitude. The proficiency scores of the participants are also taken into consideration in these correlational analyses.

## 7.1 Learner Sensitivity to Agreement Violations

### 7.1.1 Descriptive Statistics

#### 7.1.1.1 Grammaticality Judgments

In the grammaticality judgment task, participants were presented with sentences containing either grammatical or ungrammatical instances of Noun-Adjective agreement, mixed with fillers. Acceptance rates reflecting the percentage of items that were accepted as grammatical were calculated for each participant for each experimental condition [grammatical (NUM/GEN), ungrammatical with respect to number (\*NUM), and ungrammatical with respect to gender (\*GEN)]. Thus, it was expected that if the participant performed well, the acceptance rate for grammatical conditions would approach 100% and for ungrammatical conditions, 0%. The mean acceptance rates for native speakers and learners are presented in Figure 1<sup>11</sup>.



**Figure 1. Mean acceptance rates for native speakers (NS) and learners (L2) across experimental conditions**

<sup>11</sup> See Appendices 3 and 4 for a table of means for learners and native speakers, respectively.

For the grammatical condition, both groups demonstrated high acceptance rates, with the native speakers at 98% (SD=2.09; Range=92.5-100) and the learners at 92% (SD=10.26; Range=60-100). Lower acceptance rates for the ungrammatical conditions were observed. The group of native speakers incorrectly accepted only 3% (SD=2.91; Range=0-10) for violations of number agreement and 4% (SD=3.28; Range=0-10) for violations of gender agreement. The learner group averaged 22% (SD=26.40; Range=0-92.5) and 38% (SD=28.72; Range=5-97.5) acceptance rates for number and gender violations, respectively. For the ungrammatical number condition, twelve learners demonstrated acceptance rates within the range of the native speakers. There were five learners with acceptance rates within the range of native speakers for the ungrammatical gender condition. Statistical tests for Group, Grammaticality, and Violation Types will be presented after a brief look at the other two experimental tasks: vocabulary recognition and gender assignment.

#### **7.1.1.2 Vocabulary Recognition**

Following the grammaticality judgment task, learners were tested on their knowledge of the adjectives and verbs used in the predicates in the experimental sentences and fillers. During this vocabulary recognition task, participants had to choose between two English meanings for the word presented in Spanish. The mean accuracy rate for all items, including both adjectives and verbs, was 93% (SD=4.69), with a range of 80-98%. The mean was slightly higher when only the experimental items were included (adjectives only): mean = 94% (SD=4.42), range = 83-100%. Thus, it was concluded that learners were sufficiently familiar with the vocabulary used in the experimental task.

### **7.1.1.3 Gender Assignment**

The gender assignment task was administered to both native speakers and learners following the experimental task. By making a choice between masculine and feminine determiners, participants identified the gender of each noun used as a subject in the target stimuli. The performance of the native speakers served to validate the test since they gave the expected responses at a mean rate of 99% (SD=1.11), with a range of 96-100%. Additionally, for this task, learners had to be able to assign the appropriate gender under conditions mimicking those of the experimental task. Learners responded with a mean accuracy rate of 98% (SD=3.28) and a range of 88-100%. Thus, it was assumed that overall errors observed in the gender condition in the experimental task were not due to inability on the part of the learners to correctly assign gender to the particular nouns used in the stimuli.

## **7.1.2 ANOVA Results**

### **7.1.2.1 Analysis of Group and Grammaticality Effects**

To test learner sensitivity to agreement violations, a repeated-measures analysis of variance (ANOVA) with Grammaticality as the within-subjects factor was conducted for each ungrammatical condition separately against the grammatical condition. Differences between the native speakers and learners were tested by including Group as the between-subjects factor. The results of these analyses are presented in Table 4.

**Table 4. Results of repeated-measures ANOVAs on acceptance rates for each violation type**

		Number	Gender
Group	F	3.817	12.685
	( <i>p</i> )	(.060)	(.001)
Grammaticality	F	305.658	232.910
	( <i>p</i> )	(.000)	(.000)
Group x Grammaticality	F	6.796	16.987
	( <i>p</i> )	(.014)	(.000)

When the grammatical (NUM/GEN) and ungrammatical number (\*NUM) conditions were compared, there was a significant effect for Grammaticality,  $F(1,30)=305.658, p<.001$ . The effect for Group was marginal,  $F(1,30)=3.817, p=.060$ , and a significant interaction was found between Group and Grammaticality,  $F(1,30)=6.796, p=.014$ . Because of the marginal effect for Group and the significant Group x Grammaticality interaction, post-hoc paired-samples t-tests were conducted, which demonstrated that the differences between acceptance rates for grammatical and ungrammatical sentences were significant for both native speakers ( $t_{(11)}=95.670, p<.001$ ) and learners ( $t_{(19)}=9.698, p<.001$ ). Both groups demonstrated sensitivity to violations of number agreement on adjectival predicates.

Similar results were found when the grammatical (NUM/GEN) and ungrammatical gender (\*GEN) conditions were compared. There was again a significant effect for Grammaticality,  $F(1,30)=232.910, p<.001$ , but in this case there was also a significant effect for Group,  $F(1,30)=12.685, p=.001$ . The Group x Grammaticality interaction was significant for gender agreement,  $F(1,30)=16.987, p<.001$ . However, paired-samples t-tests again showed that the differences between acceptance rates for grammatical and ungrammatical sentences were significant for both native speakers ( $t_{(11)}=86.504, p<.001$ ) and learners ( $t_{(19)}=7.265, p<.001$ ).

In summary, both groups generally demonstrated the same pattern of high acceptance rates for grammatical sentences and significantly lower acceptance rates for ungrammatical

sentences across both types of agreement. Although learner acceptance rates for ungrammatical sentences were higher than those of the native speakers, this difference was only significant between groups for gender agreement, while it was marginal for number agreement.

### **7.1.2.2 Analysis of Effects for Violation Type**

In the analysis presented above, it was not possible to include a test for Violation Type, since the same stimuli were used as grammatical counterparts for both the number and the gender conditions. Therefore, in order to directly compare learner sensitivity to violations across the two different types of agreement, addressing research question 2,  $d'$  scores for each participant were calculated to reflect the standardized differences in the average acceptance rates between each ungrammatical condition and its grammatical counterpart. A  $d'$  score near zero represents performance at chance, while perfect performance in this analysis results in a  $d'$  score of approximately 4.0.<sup>12</sup> The choice of  $d'$  analysis was also motivated by the desire to preserve both the differences between participants and the differences between each participant's acceptance rates for grammatical and ungrammatical sentences for each type of agreement. The average  $d'$  scores for the learner group were 2.042 (SD=1.11) for number agreement and 1.436 (SD=.855) for gender agreement. Individual  $d'$  scores for native speakers and learners are presented in Appendices 5 and 6, respectively.

---

<sup>12</sup>  $D'$  scores were calculated in Microsoft Excel using the following formula for forced-choice experimental paradigms, where H = the Hit rate for choosing Grammatical when the stimulus is grammatical, and FA = the False Alarm rate for choosing Grammatical when the stimulus is ungrammatical:  $(\text{NORMSINV}(H) - \text{NORMSINV}(FA)) / \text{SQRT}(2)$ . Following standard procedures, Hit rates of 1 (corresponding to 100% acceptance rates) were corrected to  $1 - 1/480$ , where 480 represented twice the number of items in the test. False Alarm rates of zero were corrected to  $1/480$ .

A repeated-measures ANOVA was conducted on  $d'$  scores for the learner group only, with Violation Type as the within-subjects factor. A significant effect of Violation Type was present,  $F(1,19)=19.389, p<.001$ . These results suggest that the learners were significantly worse at detecting violations of gender agreement than number agreement.

## **7.2 Individual Differences**

In addition to the grammaticality judgment task, L2 participants also completed tests for verbal aptitude (MLAT), nonverbal aptitude (RAVEN), and Spanish proficiency. The descriptive statistics for these independent variables are presented here, along with tests for correlations between test scores, and finally, for correlations between the test scores and the dependent measures of sensitivity to agreement violations (combined  $d'$  scores for both types of violations, plus  $d'$  scores for number and gender separately). Proficiency test scores are included in this analysis, given the fact that previous studies had reported a moderate correlation between verbal aptitude scores and L2 proficiency measures (Carroll, 1981).

### **7.2.1 Independent Variables**

#### **7.2.1.1 Verbal Aptitude**

The participants' MLAT scores ranged from 51 to 83 points out of a possible total of 119 points. The mean score was 62.45 ( $SD=8.56$ ). As could be expected for a standardized test, the scores were clustered around the 50<sup>th</sup> percentile, with an average of 48<sup>th</sup> percentile. The mean score for the MLAT 3 (Spelling Clues) was 19.00 ( $SD=5.60$ ) out of 50 points, with a range of 8 to 30 points. The mean score for the MLAT4 (Words in Sentences) was 22.40 ( $SD=6.24$ ) out of

45 points, with a range of 14 to 35 points. The mean score for the MLAT5 (Paired Associates) was 21.05 (SD=3.72) out of 24 points, with a range of 10 to 24 points.

### 7.2.1.2 Nonverbal Aptitude

The mean score for all participants on the Raven’s Advanced Progressive Matrices was 21.10 (SD=3.89) out of a possible 36 points. Scores on this measure ranged from 10 to 28 points. No standardized results can be reported due to the reduced time allowed for the test.

### 7.2.1.3 Proficiency

The combination MLA/DELE test for Spanish proficiency involved a total of 50 multiple choice items. The scores for the L2 participants in this study ranged from 13 to 27 points, so all scores were in the range designated as low proficiency (0-29 points). The mean score was 19.95 (SD=3.98).

### 7.2.1.4 Summary

The descriptive measures of aptitude and proficiency are summarized in Table 5:

**Table 5. Observed measures of aptitude and proficiency, along with the maximum possible scores for each measure**

Test	Mean	SD	Range	Max
MLAT3	19.00	5.60	8-30	out of 50
MLAT4	22.40	6.24	14-35	out of 45
MLAT5	21.05	3.72	10-24	out of 24
MLAT Total	62.45	8.56	51-83	out of 119
RAVEN	21.10	3.89	10-28	out of 36
Proficiency	19.95	3.98	13-27	out of 50



### 7.2.2 Intercorrelations of Independent Variables

In order to examine the interrelationship of verbal aptitude, nonverbal intelligence, and L2 proficiency, correlations between all independent variables were investigated. The results (see Table 6 below) indicate that there were no significant correlations between any measure on the MLAT and either nonverbal aptitude as measured by the RAVEN (MLAT3:  $r=-.135, p=.570$ ; MLAT4:  $r=.159, p=.504$ ; MLAT5:  $r=.131, p=.583$ ; MLAT Total:  $r=-.084, p=.725$ ) or proficiency scores on the MLA/DELE test (MLAT3:  $r=-.068, p=.774$ ; MLAT4:  $r=.329, p=.157$ ; MLAT5:  $r=.377, p=.101$ ; MLAT Total:  $r=.359, p=.120$ ). Likewise, nonverbal aptitude and proficiency demonstrated no significant correlation,  $r=.225, p=.341$ . Indeed, the only significant results were for correlations between the total MLAT score and each of the scores for its subsections (MLAT3:  $r=.492, p=.028$ ; MLAT4:  $r=.631, p=.003$ ; MLAT5:  $r=.500, p=.025$ ), as should be the case.

**Table 6. Intercorrelations of aptitude and proficiency scores**

	MLAT3	MLAT4	MLAT5	MLAT Total	RAVEN
MLAT3					
MLAT4	-.217 (.359)				
MLAT5	-.010 (.966)	.099 (.679)			
MLAT Total	<b>.492*</b> <b>(.028)</b>	<b>.631*</b> <b>(.003)</b>	<b>.500*</b> <b>(.025)</b>		
RAVEN	-.135 (.570)	.159 (.504)	.131 (.583)	-.084 (.725)	
Proficiency	-.068 (.774)	.329 (.157)	.377 (.101)	.359 (.120)	.225 (.341)

\* Correlation is significant at the  $p<.05$  level.

Given these results, no partial correlations were investigated in the analyses of correlations between these variables and the dependent variables that are presented below.

### **7.2.3 Correlations Between Independent and Dependent Variables**

This subsection presents tests for correlations that were conducted using each of the aptitude and proficiency measures. Correlational analyses of the relationships between learner sensitivity to agreement violations and both verbal and nonverbal aptitude are presented first, in order to address research questions 3 and 4. The measure of sensitivity to violations employed in these analyses is the mean for each participant of the combined gender and number  $d'$  scores. Since research question 5 asks whether sensitivity to violations for features that differ in their similarity to the L1 are differentially modulated by verbal or nonverbal aptitude, an analysis of correlations between the separate  $d'$  scores for number/gender and measures of aptitude are presented as well. Finally, correlations between  $d'$  scores and proficiency are presented. This subsection concludes by summarizing all of these results in one table<sup>13</sup>.

#### **7.2.3.1 Verbal Aptitude and Sensitivity to Agreement Violations**

Scores for all measures of verbal aptitude (MLAT3, MLAT4, MLAT5, and MLAT Total) were tested for correlation with the mean of  $d'$  scores for both types of agreement (Combined), as well as  $d'$  scores for each type of agreement (Number, Gender). The resulting Pearson correlation coefficients are presented in Table 7 and subsequently described.

---

<sup>13</sup> Additionally, scatterplots of all three types of  $d'$  scores (Combined, Number, and Gender) against the MLAT Total, RAVEN, and PROF can be found in Appendix 4.

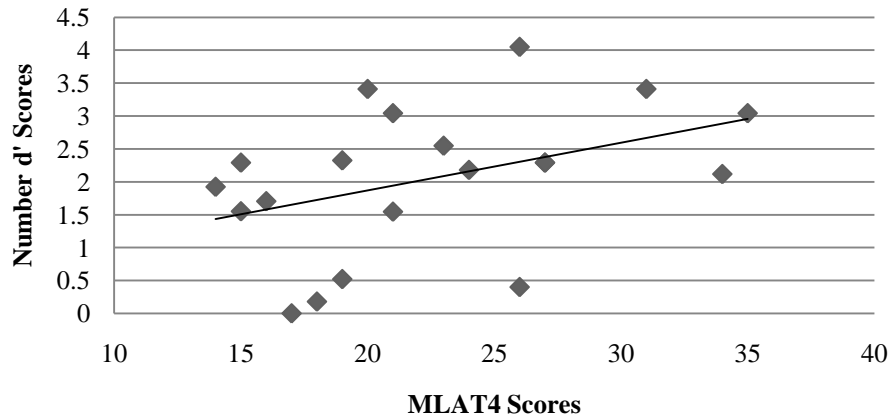
**Table 7. Correlations between d' scores and measures of verbal aptitude**

Test		d' Scores			
			Combined	Number	Gender
MLAT3	<i>r</i>	-.131	-.190	-.074	
	( <i>p</i> )	(.581)	(.423)	(.757)	
MLAT4	<i>r</i>	.309	<b>.407†</b>	.307	
	( <i>p</i> )	(.184)	<b>(.075)</b>	(.188)	
MLAT5	<i>r</i>	.301	.363	.251	
	( <i>p</i> )	(.198)	(.115)	(.287)	
MLAT Total	<i>r</i>	.270	.331	.285	
	( <i>p</i> )	(.249)	(.155)	(.224)	

† Correlation is marginally significant ( $.05 < p < .10$ ).

No significant correlation was found between the combined d' scores and any measure of verbal aptitude (MLAT3:  $r = -.131$ ,  $p = .581$ ; MLAT4:  $r = .309$ ,  $p = .184$ ; MLAT5:  $r = .301$ ,  $p = .198$ ; MLAT Total:  $r = .270$ ,  $p = .249$ ).

For number agreement, the MLAT3, MLAT5, and MLAT Total scores were not correlated with d' scores (MLAT3:  $r = -.190$ ,  $p = .423$ ; MLAT5:  $r = .363$ ,  $p = .115$ ; MLAT Total:  $r = .331$ ,  $p = .155$ ). The MLAT4 was marginally correlated with d' scores for number agreement,  $r = .407$ ,  $p = .075$ . While the effect was marginal, it was present with roughly the same strength of correlation as has been reported previously for the MLAT and various proficiency measures in classroom learners (Carroll, 1981). A scatterplot of d' scores for number agreement against the MLAT4 scores can be visually inspected in Figure 2.



**Figure 2. The d' scores for number agreement plotted against MLAT4 scores**

As can also be seen in Table 7, no correlations were found between any measure of verbal aptitude and d' scores for gender agreement (MLAT3:  $r=-.074$ ,  $p=.757$ ; MLAT4:  $r=.307$ ,  $p=.188$ ; MLAT5:  $r=.251$ ,  $p=.287$ ; MLAT Total:  $r=.285$ ,  $p=.224$ ).

### 7.2.3.2 Nonverbal Aptitude and Sensitivity to Agreement Violations

Nonverbal aptitude, as measured by the RAVEN, was not found to be significantly correlated with any measure of L2 sensitivity to agreement violations (Combined:  $r=-.213$ ,  $p=.368$ ; Number:  $r=-.121$ ,  $p=.611$ ; Gender:  $r=-.123$ ,  $p=.605$ ).

### 7.2.3.3 Proficiency and Sensitivity to Agreement Violations

In contrast to the general lack of correlations between aptitude measures and d' scores, proficiency was found to be significantly correlated with combined d' scores ( $r=.497$ ,  $p=.026$ ) as well as d' scores for gender agreement ( $r=.484$ ,  $p=.030$ ), and marginally correlated with d' scores for number agreement ( $r=.397$ ,  $p=.083$ ). Thus,  $R^2$  calculations show that proficiency explains 25% of the variance in d' scores overall and 23% of variance in d' scores for gender agreement.

### 7.2.3.4 Summary

In summary, the only significant correlations uncovered in this analysis were between proficiency and the combined  $d'$  scores, as well as the  $d'$  scores for gender agreement. Marginal correlations were also found between proficiency and  $d'$  scores for number agreement. Except for a marginal correlation between  $d'$  scores for number agreement and the MLAT4, which tests language analytic ability, no other correlations were found for either verbal or nonverbal aptitude. For ease of reference, all of the results of the analyses presented in this subsection are presented in Table 8 below.

**Table 8. Correlations between  $d'$  scores and measures of aptitude and proficiency**

Test		$r$	$d'$ Scores		
			Combined	Number	Gender
MLAT3	$r$	-.131	-.190	-.074	
	( $p$ )	(.581)	(.423)	(.757)	
MLAT4	$r$	.309	<b>.407†</b>	.307	
	( $p$ )	(.184)	<b>(.075)</b>	(.188)	
MLAT5	$r$	.301	.363	.251	
	( $p$ )	(.198)	(.115)	(.287)	
MLAT Total	$r$	.270	.331	.285	
	( $p$ )	(.249)	(.155)	(.224)	
RAVEN	$r$	-.213	-.121	-.123	
	( $p$ )	(.368)	(.611)	(.605)	
Proficiency	$r$	<b>.497*</b>	<b>.397†</b>	<b>.484*</b>	
	( $p$ )	<b>(.026)</b>	<b>(.083)</b>	<b>(.030)</b>	

\* Correlation is significant at the  $p < .05$  level.

† Correlation is marginally significant ( $.05 < p < .10$ ).

## 8 Discussion

The current study was designed to investigate learner sensitivity to violations of number and gender agreement in L2 Spanish. Additionally, the role of individual differences in modulating responses to number and gender agreement was investigated. The following research questions were proposed:

RQ1: Do learners show the same pattern of sensitivity as native speakers when it comes to violations involving uninterpretable features?

RQ2: Are learners equally sensitive to agreement violations for number features that are present in the L1 and for gender features that are unique to the L2?

RQ3: Does verbal aptitude modulate sensitivity to agreement violations and/or to number and gender agreement differentially?

RQ4: Does nonverbal aptitude modulate sensitivity to agreement violations and/or to number and gender agreement differentially?

This section will discuss each of these questions in terms of the results reported here, along with possible limitations in interpreting the results given the constraints of the research design and stimuli.

With regard to the first research question, it was hypothesized that learners would pattern with native speakers with regard to number agreement even though number agreement does not occur on adjectival predicates in English. Since grammatical gender is a feature unique to the L2 Spanish, it was expected to be more difficult for learners, at least at the level of proficiency of the learners in this study, which was relatively low as confirmed by the MLA/DELE proficiency test. While Interpretability Hypothesis predicts that no learners should be able to acquire the gender feature, the FDH and FTFA theories allow that some learners at this stage may demonstrate sensitivity to gender agreement violations. Across the entire participant base, these predictions were tested separately for the two types of agreement using repeated-measures ANOVA, which revealed significant group differences between learners and native speakers only for gender agreement, while a marginal group effect was present in the number agreement conditions. These differences were tied to the poorer performance of learners on ungrammatical

conditions, but follow-up tests confirmed the same general pattern in both groups, with significantly higher acceptance rates for grammatical than ungrammatical conditions for both number and gender agreement. Thus the results confirm that learners differ from native speakers only quantitatively with regard to sensitivity to both number and gender violations.

Additionally, there were five learners who demonstrated native-like performance on the ungrammatical gender conditions. These findings are more in line with Full Transfer/Full Access than with either the Interpretability Hypothesis or the Fundamental Difference Hypothesis.

In an effort to address the second research question by comparing number and gender agreement more directly, a repeated-measures ANOVA was used to test for significant differences between sensitivity to number and gender violations as measured by  $d'$  scores. Learners were expected to perform better on number agreement than gender agreement, and this prediction was borne out. Results indicated that learners were significantly worse at detecting gender violations. A gender assignment task demonstrated that learners were able to correctly assign gender to the nouns used in the study under the same conditions present in the grammaticality judgment task, indicating that errors were not due to faulty gender assignment.

Overall, the results with regard to learner sensitivity to agreement violations suggest that unique L2 features like gender may present more of a problem than number features for learners at low levels of proficiency, but that learners are able to acquire both number and gender agreement. Contrary to the Interpretability Hypothesis, some learners seem to be able to perform at native-like levels even in the case of gender agreement, which is not present in their L1. Whether or not those learners are of exceptional ability, as predicted by the Fundamental Difference Hypothesis, will be addressed below.

Research question 3 posited the question of whether sensitivity to specific L2 structures would be modulated by verbal aptitude. Based on the results of previous studies (DeKeyser, 2000; Abrahamsson and Hyltenstam, 2008; Harley and Hart, 1997), it was hypothesized that sensitivity to agreement violations (here,  $d'$  scores) would be correlated with MLAT scores - at least the MLAT4, which tests language analytic ability. The analysis revealed no significant correlations between  $d'$  scores and any verbal aptitude measures, including the MLAT Total score and scores for each subsection of the MLAT. There was a marginal correlation between  $d'$  scores for number agreement and the MLAT4, leaving room for the possibility that this correlation might be significant given larger numbers of participants. Of course, it is possible that a lack of variation in either  $d'$  scores or MLAT scores could result in the lack of correlations seen here, but a visual inspection of scatterplots for these variables (see Appendix 7) shows no evidence for such a lack of variation.

One surprising result is that no measure of verbal aptitude was found to correlate with proficiency test scores, even though previous studies involving adult learners had reported a moderate correlation between verbal aptitude scores and broad measures of L2 ability (e.g., Carroll, 1981; Harley and Hart, 1997). The lack of correlation between aptitude and general proficiency raises the question of whether the lack of correlations here between aptitude and specific L2 measures can be generalized to other learners. Since the sample tested here ( $n=20$ ), while small, is large enough to make correlations possible and shows no evidence of abnormality in terms of either the standardized MLAT scores or the proficiency scores, it does not seem likely that the lack of correlation between aptitude and proficiency is due to sampling error. One possibility, however, is that the aptitude tests and proficiency tests are different enough that they do not test the same constructs, where the opposite may have been the case in other studies. The



MLA/DELE test, as has been previously described, is largely a multiple choice test of lexical knowledge or knowledge of specific uses of functional words to which learners with greater exposure to the L2 would be more familiar. It could be argued that such a proficiency test may not tap grammatical inferencing/verbal analytic ability, or verbal working memory (as tested by the MLAT) to the same extent as the grammar-based tasks in other studies (e.g., DeKeyser, 2000; Abrahamsson and Hyltenstam, 2008).

Another possibility is that the instructional environment of the learners may play a role in the overall lack of correlations for aptitude with either broad proficiency or specific L2 measures. The MLAT, or verbal aptitude measures more generally, has previously been argued to bear more predictive validity in formal learning environments due to the explicit nature of instruction (Krashen, 1981). Although learners in the current study were acquiring Spanish in a formal environment, they were being taught under a curriculum that applies as few explicit grammatical explanations as possible. While the lack of correlations may thus be attributed to the communicative learning undertaken by these participants, there is reason to believe that this argument may not be valid. Robinson (2002), described above, found aptitude to be correlated with L2 outcomes both immediately in an explicit condition and at a delay in an incidental learning condition that was focused on meaning. Furthermore, DeGraaf (1997) conducted a series of tests on beginning learners of an artificial language and found that aptitude scores correlated with mean scores on three of four tasks, including a grammaticality judgment task under time pressure. Finally, Skehan (1982) showed correlations between aptitude and L2 outcomes in an environment based on communicative competence rather than formal instruction.

While the lack of correlation between aptitude and proficiency in this group of learners is surprising, it may, on the other hand, make it possible to investigate their roles separately. While

proficiency was a factor in L2 sensitivity to violations involving features that are both similar to the L1 and unique to the L2, verbal aptitude - particularly language analytic ability as measured by the MLAT4 - may have influenced learner performance only on measures of sensitivity to number features. As the correlation between number  $d'$  scores and the MLAT4 was only marginal, any further discussion with regard to verbal aptitude and L1/L2 differences would be speculative. A future extension of this study will investigate the role of aptitude with regard to L1/L2 similarities and differences by testing more participants and including for analysis the number agreement violations in the subject-verb agreement conditions from this experiment's fillers. English and Spanish are similar in this regard, thus correlations between MLAT4 scores and  $d'$  scores for subject-verb agreement are expected if verbal analytical ability does indeed play a role in the processing of L1-like features. If the relationship between aptitude and sensitivity to number violations is found to be significant, given more participants, this might indicate a differential impact of verbal aptitude and proficiency on the acquisition of different structures.

In response to research question 4, the question was raised as to whether or not nonverbal measures would reveal a role for domain-general cognitive factors. The FDH predicts such a role, while it could be argued that a role for domain-specific factors would be more appropriately predicted under full-access theories. Indeed, the results of the current study find no evidence that nonverbal aptitude plays a role, even in the acquisition of unique L2 features, as seen in the lack of correlations between the RAVEN and  $d'$  scores. While the matter cannot be settled on the basis of one test, it may be useful to consider here the nature of that test<sup>14</sup>. The RAVEN was chosen for the study due to its broad use in cognitive studies as a general measure of nonverbal

---

<sup>14</sup> Indeed, Robinson (2002) argues that tests of individual differences will be correlated with L2 performance only where they tap the same constructs as the L2 tasks involved.

intelligence and mental reasoning. Carpenter, Just, and Shell (1990) cite research into correlations between the RAVEN and other intelligence measures (Court and Raven, 1982; Jensen, 1987) suggesting that the processing that underlies the RAVEN may not be specific just to that test, but rather, general in nature and likely more central to the testing of domain-general analytic ability than a number of other tests, confirming the analysis of Snow, Kyllonen, and Marshalek (1984) that was cited previously. Carpenter et al. also report on a detailed analysis of factors relevant to performance on the RAVEN. Their findings are framed in terms of two simulated models of RAVEN performance, one for median performance and one for best performance, based on the verbal protocols, eye-fixation patterns, and errors of 79 students recorded in a series of experiments involving the RAVEN test. Results indicate that the ability to decompose problems into smaller parts is central to analytic ability, and that variation in RAVEN test scores arises from individual differences in working memory and abstract reasoning. It is thus surprising that if domain-general capacities are at work, as argued by Bley-Vroman (1990, 2009), the RAVEN would not in some way reflect morphological decomposition, maintenance of features in verbal working memory, and abstraction of patterns resulting from agreement.

One final point that needs to be addressed in terms of the results of the study is that five learners performed within the range of native speakers in rejecting violations of gender agreement. Since this condition involved a feature not present in the L1, the Fundamental Difference Hypothesis might predict that these learners should demonstrate exceptional capacities for language learning. Furthermore, a correlational analysis may not be sufficient to uncover the role of aptitude, since the findings of Abrahamsson and Hyltenstam (2008) indicate that high aptitude may be necessary but not sufficient for successful L2 acquisition, in which

case high performers might demonstrate high aptitude, but not all high-aptitude learners might perform well. Therefore, a brief analysis of the five high performers in this study is presented here, following DeKeyser (2000). DeKeyser found that of the six adult learners who obtained high scores on a grammaticality judgment task, five had verbal aptitude scores above an arbitrary cutoff (6 points out of a possible 20) that represented at least .46 SD from the overall average, and DeKeyser argued that the verbal aptitude score of the sixth may not have been indicative of his above-average analytical ability, as suggested by his career path. Of the five learners in the current study who performed within the range of native speakers on gender agreement, only two demonstrated verbal aptitude scores above .50 SD of the group average, and one of these is only at the 55<sup>th</sup> percentile in the standardized MLAT scores. Of the other three, only one has a RAVEN score higher than the group average, right around .50 SD above average. However, four of the five high performers in terms of gender agreement demonstrated Spanish proficiency scores at or above .50 SD above the group average, suggesting that proficiency played a greater role than either verbal or nonverbal aptitude. This generalization also held true when the proficiency scores of the high performers were compared to those of the rest of the group using one-tailed independent-samples *t*-tests, assuming unequal variances ( $t_{(8)}=2.154$ ,  $p=.032$ ). No significant differences were found when the same comparison was made based on RAVEN or MLAT test or subtest scores (RAVEN:  $t_{(5)}=-.228$ ,  $p=.414$ ; MLAT Total:  $t_{(6)}=-.171$ ,  $p=.435$ ; MLAT3:  $t_{(6)}=-.961$ ,  $p=.187$ , MLAT4:  $t_{(10)}=.885$ ,  $p=.198$ , MLAT5:  $t_{(8)}=-.184$ ,  $p=.429$ ). In sum, both an analysis of high performers on gender agreement and a correlational analysis across the entire group of learners provide evidence corroborating the findings of other studies on gender agreement that show increasing accuracy with proficiency (e.g., White and Genesee, 1996; White et al., 2004; Sagarra and Herschensohn, 2010). Even in the relatively low-proficiency

group tested here, proficiency appears to have played the greatest role with regard to grammaticality judgments. These findings are most in line with the Full Transfer/Full Access theory.

There are a number of limitations to the current study. First, the comparison of number and gender agreement here is based on stimuli that may not be equitable with regard to the two features. One possible issue is that the subject DPs in the experimental stimuli were approximately 50% masculine and 50% feminine, but all were singular with regard to number. The use of both masculine and feminine subjects may have made processing gender more demanding. Additionally, the verb that intervenes between the DP subject and the adjective in these stimuli, as in the sample given above and repeated here in (21), bears a number feature but not gender, providing an additional cue for number agreement on the following adjective.

- (21) La isla es **\*rocoso** y la península también.  
The<sub>sg.fem</sub> island<sub>sg.fem</sub> is<sub>3sg</sub> rocky<sub>sg.masc</sub> and the peninsula too.

Finally, number features and mismatches are also present in the fillers, which involved subject-verb agreement. Thus the number of sentences exhibiting number agreement or mismatches would have been double the number of sentences involving gender, possibly biasing learners to focus on number features.

On the other hand, it could be argued that learner performance with regard to gender agreement may be inflated due to a variety of factors. Only nouns and adjectives demonstrating canonical gender marking were employed here, making it possible for learners to have simply noticed matching patterns of *-o* endings on masculine nouns and adjectives and *-a* endings on feminine nouns and adjectives. This matching strategy might be just the type of heuristic or shallow processing that learners might rely on in L2 acquisition if the more recent version of the FDH is accurate (Bley-Vroman, 2009). The matching hypothesis can possibly be countered by

three arguments, one of which is that gender errors might actually have been more difficult for learners to detect, as discussed in the paragraph above. The correlation discussed above between proficiency and sensitivity to gender agreement violations might not be expected either, if the learners were simply using a matching strategy. Finally, a third argument merits some discussion. A heuristic relying on the detection and analysis of visual patterns and the maintenance of visual material in working memory might have been expected to produce correlations between performance and RAVEN scores, given the visual nature of the RAVEN<sup>15</sup>. In the current study, no correlations were present between the RAVEN and sensitivity to gender agreement violations. Caution is recommended in interpreting results based on negative evidence, but it is at least possible that a visual matching strategy was not in use by most learners.

Since online measures may provide a better picture of what is going on during processing, future analysis of the EEG data collected during the grammaticality judgment task may provide additional clues to address this issue. Studies investigating event-related potentials (ERPs) have identified characteristic responses in native speakers to violations of agreement features like gender or number on adjectives, or subject-verb agreement, including a late positivity (P600) in posterior regions of the brain, peaking around 600ms after presentation of the word containing the violation (e.g., Osterhout and Mobley, 1995; Barber and Carreiras, 2005; Alemán Bañón, 2010). Indeed, the stimuli from which the current experiment derives its stimuli were presented in the second of three experiments by Alemán Bañón (2010), who found that both number and gender violations on the adjectives in question produced a significantly more

---

<sup>15</sup> There are a considerable number of neurocognitive studies that, among other results, demonstrate the visual-perceptual nature of the RAVEN. Thoma, et al. (2000) found RAVEN scores to be correlated with processing speed in a test of sensory-motor control in response to visual stimuli. Prabhakaran et al. (1997) noted that a number of areas involving visual perception are activated both for RAVEN problems requiring analytical reasoning and for those requiring only figural (visual perceptual) acuity.

positive-going waveform between 400 and 900 ms, peaking at 600ms. The P600 has also been attested in learners in a variety of studies (e.g., Rossi et al., 2006; Sabourin and Stowe, 2008; Tokowicz and MacWhinney, 2005), although the response sometimes differs in amplitude or peaks later than in native speakers. Therefore, by examining P600 responses for learners and native speakers in the current study, it may be possible to examine whether or not learners are processing gender (and number) violations in a native-like manner versus relying on a matching strategy. Future directions for this research, then, include adding more participants, analyzing the role of verbal and nonverbal aptitude in sensitivity to subject-verb agreement violations, and analyzing online measures of sensitivity to agreement violations.

## **9 Conclusion**

The current study investigated sensitivity to violations of number, a feature present in the L1 of the learners, and gender, a feature that is unique to their L2. The results of a grammaticality judgment task indicated that learners are able to acquire both number and gender agreement, but that gender agreement is more problematic at low levels of proficiency. Contrary to the Interpretability Hypothesis, five learners were able to perform at native-like levels even in the case of gender agreement on adjectives, where the gender feature is uninterpretable. Those learners did not demonstrate exceptional ability on nonverbal or even verbal aptitude tests, which would have been predicted by the Fundamental Difference Hypothesis. Indeed, for the entire learner group, no correlation was found for either verbal or nonverbal aptitude against either overall L2 proficiency or measures of sensitivity to agreement violations. Indeed, proficiency was the only variable found to correlate with measures of sensitivity to violations. The importance of proficiency in the current study is in line with previous results of gender studies

claiming support for the Full Transfer/Full Access theory. In sum, evidence for the adult L2 acquisition of uninterpretable features, especially those unique to the L2, in the absence of mitigating factors such as high verbal analytical ability or nonverbal intelligence, may be taken as evidence *against* restricted access to Universal Grammar past any supposed Critical Period for language-learning and *for* Full Access, which predicts that low sensitivity to unique L2 features will increase with proficiency.

## References

- Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30, 481-509.
- Alemán Bañón, J. (2010). The processing of number and gender agreement in Spanish: An ERP investigation. MA Thesis, University of Kansas, 2010.
- Barber, H. & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: an ERP comparison. *Journal of Cognitive Neuroscience*, 17, 137-153.
- Bialystok, E. (2002). On the reliability of robustness: A reply to DeKeyser. *Studies in Second Language Acquisition*, 24, 481-488.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. Gass & J. Schacter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 41-68). Cambridge: Cambridge University Press.
- Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis*, 20, 3-49.
- Bley-Vroman, R. (2009). The evolving context of the Fundamental Difference Hypothesis. *Studies in Second Language Acquisition*, 31, 175-198.
- Carroll, J. B. (1965). The prediction of success in foreign language training. In R. Glaser (Ed.), *Training, research, and education*. New York: Wiley.
- Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in foreign language aptitude* (pp. 83-118). Rowley, MA: Newbury House.
- Carroll, S. (1989). Second-language acquisition and the computational paradigm. *Language*



*Learning*, 39, 535-594.

- Carroll, J. B., & Sapon, S. (1959). *Modern Language Aptitude Test: Form A*. New York: Psychological Corporation.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.
- Carstens, V. (2000). Concord in minimalist theory. *Linguistic Inquiry*, 31, 319–355.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1980). *Rules and Representations*. Oxford: Blackwell.
- Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein & D. Lightfoot (Eds.), *Explanation in linguistics: the logical problem of language acquisition* (pp. 32-75). London: Longman.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 1994. On the evidence for partial N-movement in the Romance DP. In G. Cinque, J. Koster, J. –Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths towards Universal Grammar: Studies in Honor of Richard S. Kayne* (pp. 85-110). Washington, DC: Georgetown University Press.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27, 3–42.
- Clahsen, H., & Muysken, P. (1986). The availability of universal grammar to adult and child learners: A study of the acquisition of German word order. *Second Language Research*, 2, 93-119.
- Corbett, G. G. (1991). *Gender*. New York: Cambridge University Press.
- Court, J. H., & Raven, J. (1982). *Manual for Raven's Progressive Matrices and Vocabulary Scales, Research Supplement No. 2* (Pt. 3, Section 7). London: H. K. Lewis.
- DeGraaff, R. (1997). The eXperanto experiment. *Studies in Second Language Acquisition*, 19, 249-276.
- DeKeyser, R.M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 499-533.
- Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and*

- Brain Sciences*, 19, 677-714.
- Flahive, D.E. (1980). Separating the *g* factor from reading comprehension. In J. W. Oller, Jr., & K. Perkins (Eds.), *Research in Language Testing* (pp. 34-46). Rowley, MA: Newbury House.
- Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17, 213-247.
- Franceschina, F. (2002). Case and (phi)-feature agreement in advanced L2 Spanish grammars. *EUROSLA Yearbook*, 2, 71-86.
- Gardner, R. C., & Lambert, W. E. (1965). Language aptitude, intelligence, and second language achievement. *Journal of Educational Psychology*, 56, 191-199.
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19, 379-400.
- Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22, 27-62.
- Hawkins, R. (1998). Explaining the difficulty of French gender attribution for speakers of English. Paper presented at the 8th annual EUROSLA conference, Paris.
- Hawkins, R., & Chan, C. Y.-H. (1997). The partial availability of Universal Grammar in second language acquisition: The “failed functional features hypothesis.” *Second Language Research*, 13, 187-226.
- Haznedar, B. (1997). L2 acquisition by a Turkish-speaking child: Evidence for L1 influence. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the 21<sup>st</sup> Annual Boston University Conference on Language Development* (pp. 245-256). Somerville, MA: Cascadilla Press.
- Jensen, A. R. (1987). The *g* beyond factor analysis. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing* (pp. 87-142). Hillsdale, NJ: Erlbaum.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25, 603-634.
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57, 1-33.
- Johnson J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.

- Juffs, A., 2004. Representation, processing, and working memory in a second language. *Transactions of the Philological Society*, 102, 199–225.
- Krashen, S. (1981). *Principles and Practice in Second Language Acquisition*. Oxford, UK: Pergamon Press.
- Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley.
- McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, 24, 459-486.
- McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, 21, 395–423.
- McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381-401.
- Meara, P., Milton, J., & Lorenzo-Dus, N. (2003). *Swansea language aptitude tests (LAT) v2.0*. Swansea, UK: Lognostics.
- Miyake, A. & Friedman, N. F. (1998). Individual differences in second language proficiency: Working memory as language “aptitude”. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339-364). Mahwah, NJ: Erlbaum.
- Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite/imperfect contrast in Spanish. *Studies in Second Language Acquisition*, 25, 351-398.
- Novoa, L., Fein, D., & Obler, L. (1988). Talent in foreign languages: A case study. In L. Obler, & D. Fein (Eds.), *The Exceptional Brain: Neuropsychology of Talent and Special Abilities*. New York: Guilford Press.
- Obler, L. (1989). Exceptional second language learners. In S. Gass, C. Madden, D. Preston, & L. Selinker (Eds.), *Variation in Second Language Acquisition: Psycholinguistic Issues*. Clevedon, Avon: Multilingual Matters.
- Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97-113.
- Ortega, L. (2009). *Understanding Second Language Acquisition*. London: Hodder Education.
- Osterhout, L., & Mobley, L.A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739-773.

- Penfield, W. & Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton, NJ: Princeton University Press.
- Picallo, M. C. (1991). Nominals and nominalizations in Catalan. *Probus*, 3, 279-316.
- Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices test. *Cognitive Psychology*, 33, 43-63.
- Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 103-133.
- Raven, J. C. (1965). *Advanced Progressive Matrices, Sets I and II*. London: H. K. Lewis. (Distributed in the United States by The Psychological Corporation, San Antonio, TX).
- Renaud, C. (2010). On the nature of agreement in English-French acquisition: A processing investigation in the verbal and nominal domains. PhD. Dissertation, Indiana University.
- Ritter, E. (1991). Two functional categories in noun phrases: Evidence from Modern Hebrew. In S. D. Rothstein (Ed.), *Perspectives on Phrase Structure: Heads and Licensing, Syntax and Semantics Vol. 25* (pp. 37-62). New York: Academic Press.
- Ritter, E. (1993). Where is gender? *Linguistic Inquiry*, 24, 795-803.
- Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning*. (pp. 211-266). Amsterdam: John Benjamins.
- Rossi, S., Gugler, M.F., Friederici, A.D., & Hahne, A. (2006). The impact of proficiency on syntactic second language processing of German and Italian: evidence from event-related potentials. *Journal of Cognitive Neuroscience*, 18, 2030-2048.
- Sabourin, L. , & Stowe, L. (2008). Second language processing: when are first and second languages processed similarly? *Second Language Research*, 24, 397-430.
- Sagarra, N. (2007). Online processing of gender agreement in low proficient English-Spanish late bilinguals. In M. J. Cabrera, J. Camacho, V. Déprez, N. Flores, L. Sánchez (Eds.), *Current Issues in Linguistic Theory Series* (pp. 240-253). John Benjamins, Amsterdam.
- Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, 120, 2022-2039.
- Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses*. New York: Peter Lang.

- Schwartz, B. D., & Sprouse, R. (1994). Word order and nominative case in nonnative language acquisition. In T. Hoekstra & B. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp. 317–368). Amsterdam: John Benjamins.
- Schwartz, B. D., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, 12, 40–72.
- Schneiderman, E., & Desmarais, C. (1988a). The talented language learner: Some preliminary findings. *Second Language Research*, 4, 91-109.
- Schneiderman, E., & Desmarais, C. (1988b). A neuropsychological substrate for talent in second language acquisition. In L. Obler & D. Fein (Eds.), *The Exceptional Brain: Neuropsychology of Talent and Special Abilities*. New York: Guilford Press.
- Service, E., Simola, M., Metsanheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology*, 14, 383–408.
- Skehan, P. (1982). Memory and motivation in language aptitude testing. PhD Dissertation, University of London.
- Skehan, P. (1986a). The role of foreign language aptitude in a model of school learning. *Language Testing*, 3, 188–221.
- Skehan, P. (1986b). Where does language aptitude come from? In P. Meara (Ed.), *Spoken Language*. London: Centre for Information on Language Teaching.
- Skehan, P. (1989). *Individual differences in second language learning*. London: Arnold.
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. New York: Oxford University Press.
- Slabakova, R. (2000). L1 transfer revisited: The L2 acquisition of telicity marking in English by Spanish and Bulgarian native speakers. *Linguistics*, 38, 739-770.
- Smith, N., & Tsimpli, X. (1995). *The Mind of a Savant*. Oxford, UK: Blackwell.
- Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence*, Vol. 2 (pp. 47-103). Hillsdale, NJ: Erlbaum.
- Sternberg, R. J. (2002). The theory of successful intelligence and its implications for language-aptitude testing. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning* (pp. 13-44). Amsterdam: John Benjamins.
- Tagliaferri, B. (2005). *Paradigm*. Perception Research Systems, Inc.

The *Acceso* Project. Available: <http://www2.ku.edu/~spanish/acceso/>.

Thoma, R. J., Lewine, J. D., Davis, J. T., & Paulson, K. (2000). Duration of parietal MEG source dipole correlated with Ravens advanced progressive matrices score. *Journal of Cognitive Neuroscience, (Suppl. S)*, 86–87.

Thompson, A. S. (2009). The multilingual/bilingual dichotomy: An exploration of individual differences. PhD Dissertation, Michigan State University.

Thurstone, L. L., & Thurstone, T. G. (1941). *The Chicago Tests of Primary Mental Abilities, Ages 11-17, Manual*. Washington, D.C.: American Council on Education.

Thurstone, L. L., & Thurstone, T. G. (1965). *Primary Mental Abilities Test*. Chicago: Science Research Associates.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: an event related potential investigation. *Studies in Second Language Acquisition*, 27, 173-204.

Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: Evidence from *wh*-interrogatives in second language acquisition. *Second Language Research*, 23, 215-242.

Tsimpli, I. M., & Mastropavlou, M. (2007). Feature interpretability in L2 acquisition and SLI: Greek clitics and determiners. In J. Liceras, H. Zobl, & H. Goodluck, (Eds.), *The Role of Formal Features in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Wesche, M. B., Edwards, H., & Wells, W. (1982). Foreign language aptitude and intelligence. *Applied Psycholinguistics*, 3, 127-140.

White, L. (1985). The pro-drop parameter in adult second language acquisition. *Language Learning*, 35, 47-62.

White, L. (1986). Implications of parametric variation for adult second language acquisition: An investigation of the 'pro-drop' parameter. In V. Cook (Ed.), *Experimental Approaches to Second Language Acquisition* (pp. 55-72). Oxford, UK: Pergamon Press.

White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge, UK: Cambridge University Press.

White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, 12 (3), 233-265.

White, L., & Juffs, A. (1998). Constraints on wh-movement in two different contexts of non-native language acquisition: competence and processing. In S. Flynn, G. Martohardjono, & W. O'Neil (Eds.), *The Generative Study of Second Language Acquisition* (pp. 111-129). Mahwah, NJ: Lawrence Erlbaum.

White, L., Valenzuela, E., Kozłowska-Macgregor, M. & Leung, Y.-K.I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 25, 105–33.

Yuan, B. (1998). Interpretation of binding and orientation of the Chinese reflexive *ziji* by English and Japanese speakers. *Second Language Research*, 14, 324-340.

## APPENDIX 1: Target stimuli

1. El Mediterráneo es profundo y el Adriático también.
2. El Mediterráneo es \*profundos y el Adriático también.
3. El Mediterráneo es \*profunda y el Adriático también.
  
4. El arroyo es profundo y el charco también.
5. El arroyo es \*profundos y el charco también.
6. El arroyo es \*profunda y el charco también.
  
7. El Mediterráneo es tranquilo y el Caspio también.
8. El Mediterráneo es \*tranquilos y el Caspio también.
9. El Mediterráneo es \*tranquila y el Caspio también.
  
10. La boda es tranquila y la gala también.
11. La boda es \*tranquilas y la gala también.
12. La boda es \*tranquilo y la gala también.
  
13. La naranja es redonda y la sandía también.
14. La naranja es \*redondas y la sandía también.
15. La naranja es \*redondo y la sandía también.
  
16. La Tierra es redonda y la Luna también.
17. La Tierra es \*redondas y la Luna también.
18. La Tierra es \*redondo y la Luna también.
  
19. El Atlántico es gigantesco y el Egeo también.
20. El Atlántico es \*gigantescos y el Egeo también.
21. El Atlántico es \*gigantesca y el Egeo también.
  
22. La Tierra es gigantesca y la atmósfera también.
23. La Tierra es \*gigantescas y la atmósfera también.
24. La Tierra es \*gigantesco y la atmósfera también.
  
25. El colegio es gratuito y el instituto también.
26. El colegio es \*gratuitos y el instituto también.
27. El colegio es \*gratuita y el instituto también.
  
28. La biblioteca es gratuita y la sauna también.
29. La biblioteca es \*gratuitas y la sauna también.
30. La biblioteca es \*gratuito y la sauna también.
  
31. El colegio es antiguo y el liceo también.
32. El colegio es \*antiguos y el liceo también.
33. El colegio es \*antigua y el liceo también.



34. La huella es antigua y la calavera también.  
35. La huella es \*antiguas y la calavera también.  
36. La huella es \*antiguo y la calavera también.
37. La isla es preciosa y la bahía también.  
38. La isla es \*preciosas y la bahía también.  
39. La isla es \*precioso y la bahía también.
40. El espejo es precioso y el mosaico también.  
41. El espejo es \*preciosos y el mosaico también.  
42. El espejo es \*preciosa y el mosaico también.
43. La isla es rocosa y la península también.  
44. La isla es \*rocosas y la península también.  
45. La isla es \*rocoso y la península también.
46. La montaña es rocosa y la sierra también.  
47. La montaña es \*rocosas y la sierra también.  
48. La montaña es \*rocoso y la sierra también.
49. La iglesia es emblemática y la abadía también.  
50. La iglesia es \*emblemáticas y la abadía también.  
51. La iglesia es \*emblemático y la abadía también.
52. El Coliseo es emblemático y el Foro también.  
53. El Coliseo es \*emblemáticos y el Foro también.  
54. El Coliseo es \*emblemática y el Foro también.
55. La pastelería es famosa y la bombonería también.  
56. La pastelería es \*famosas y la bombonería también.  
57. La pastelería es \*famoso y la bombonería también.
58. El Coliseo es famoso y el Capitolio también.  
59. El Coliseo es \*famosos y el Capitolio también.  
60. El Coliseo es \*famosa y el Capitolio también.
61. La fresa es ácida y la piña también.  
62. La fresa es \*ácidas y la piña también.  
63. La fresa es \*ácido y la piña también.
64. La naranja es ácida y la mandarina también.  
65. La naranja es \*ácidas y la mandarina también.  
66. La naranja es \*ácido y la mandarina también.
67. El Atlántico es frío y el Báltico también.

68. El Atlántico es \*fríos y el Báltico también.  
69. El Atlántico es \*fría y el Báltico también.
70. El otoño es frío y el invierno también.  
71. El otoño es \*fríos y el invierno también.  
72. El otoño es \*fría y el invierno también.
73. La falda es florida y la blusa también.  
74. La falda es \*floridas y la blusa también.  
75. La falda es \*florido y la blusa también.
76. El trópico es florido y el prado también.  
77. El trópico es \*floridos y el prado también.  
78. El trópico es \*florida y el prado también.
79. El vestido es azulado y el manto también.  
80. El vestido es \*azulados y el manto también.  
81. El vestido es \*azulada y el manto también.
82. La falda es azulada y la corbata también.  
83. La falda es \*azuladas y la corbata también.  
84. La falda es \*azulado y la corbata también.
85. La montaña es grandiosa y la colina también.  
86. La montaña es \*grandiosas y la colina también.  
87. La montaña es \*grandioso y la colina también.
88. El faro es grandioso y el obelisco también.  
89. El faro es \*grandiosos y el obelisco también.  
90. El faro es \*grandiosa y el obelisco también.
91. El faro es luminoso y el fuego también.  
92. El faro es \*luminosos y el fuego también.  
93. El faro es \*luminosa y el fuego también.
94. La cocina es luminosa y la terraza también.  
95. La cocina es \*luminosas y la terraza también.  
96. La cocina es \*luminoso y la terraza también.
97. La pastelería es minúscula y la panadería también.  
98. La pastelería es \*minúsculas y la panadería también.  
99. La pastelería es \*minúsculo y la panadería también.
100. La casa es minúscula y la cochera también.  
101. La casa es \*minúsculas y la cochera también.  
102. La casa es \*minúsculo y la cochera también.

103. La maleta es espaciosa y la caja también.  
104. La maleta es \*espaciosas y la caja también.  
105. La maleta es \*espacioso y la caja también.
106. La casa es espaciosa y la bodega también.  
107. La casa es \*espaciosas y la bodega también.  
108. La casa es \*espacioso y la bodega también.
109. El arroyo es estrecho y el riachuelo también.  
110. El arroyo es \*estrechos y el riachuelo también.  
111. El arroyo es \*estrecha y el riachuelo también.
112. El lago es estrecho y el barranco también.  
113. El lago es \*estrechos y el barranco también.  
114. El lago es \*estrecha y el barranco también.
115. La cocina es amplia y la entrada también.  
116. La cocina es \*amplias y la entrada también.  
117. La cocina es \*amplio y la entrada también.
118. El baño es amplio y el pasillo también.  
119. El baño es \*amplios y el pasillo también.  
120. El baño es \*amplia y el pasillo también.
121. El lago es oscuro y el pozo también.  
122. El lago es \*oscuros y el pozo también.  
123. El lago es \*oscura y el pozo también.
124. El otoño es oscuro y el invierno también.  
125. El otoño es \*oscuros y el invierno también.  
126. El otoño es \*oscura y el invierno también.
127. La maleta es pesada y la cartera también.  
128. La maleta es \*pesadas y la cartera también.  
129. La maleta es \*pesado y la cartera también.
130. La mesa es pesada y la silla también.  
131. La mesa es \*pesadas y la silla también.  
132. La mesa es \*pesado y la silla también.
133. La mesa es inmensa y la cama también.  
134. La mesa es \*inmensas y la cama también.  
135. La mesa es \*inmenso y la cama también.
136. El submarino es inmenso y el pesquero también.

137. El submarino es \*inmensos y el pesquero también.  
 138. El submarino es \*inmensa y el pesquero también.
139. El baño es rosado y el dormitorio también.  
 140. El baño es \*rosados y el dormitorio también.  
 141. El baño es \*rosada y el dormitorio también.
142. El narciso es rosado y el gladiolo también.  
 143. El narciso es \*rosados y el gladiolo también.  
 144. El narciso es \*rosada y el gladiolo también.
145. El disco es clásico y el concierto también.  
 146. El disco es \*clásicos y el concierto también.  
 147. El disco es \*clásica y el concierto también.
148. El abrigo es clásico y el chaleco también.  
 149. El abrigo es \*clásicos y el chaleco también.  
 150. El abrigo es \*clásica y el chaleco también.
151. El abrigo es caluroso y el gorro también.  
 152. El abrigo es \*calurosos y el gorro también.  
 153. El abrigo es \*calurosa y el gorro también.
154. El trópico es caluroso y el desierto también.  
 155. El trópico es \*calurosos y el desierto también.  
 156. El trópico es \*calurosa y el desierto también.
157. La biblioteca es moderna y la escuela también.  
 158. La biblioteca es \*modernas y la escuela también.  
 159. La biblioteca es \*moderno y la escuela también.
160. La cámara es moderna y la agenda también.  
 161. La cámara es \*modernas y la agenda también.  
 162. La cámara es \*moderno y la agenda también.
163. La batalla es destructiva y la lucha también.  
 164. La batalla es \*destructivas y la lucha también.  
 165. La batalla es \*destructivo y la lucha también.
166. El terremoto es destructivo y el tornado también.  
 167. El terremoto es \*destructivos y el tornado también.  
 168. El terremoto es \*destructiva y el tornado también.
169. La batalla es catastrófica y la guerra también.  
 170. La batalla es \*catastróficas y la guerra también.  
 171. La batalla es \*catastrófico y la guerra también.

172. El terremoto es catastrófico y el maremoto también.  
173. El terremoto es \*catastróficos y el maremoto también.  
174. El terremoto es \*catastrófica y el maremoto también.
175. El plátano es beneficioso y el coco también.  
176. El plátano es \*beneficiosos y el coco también.  
177. El plátano es \*beneficiosa y el coco también.
178. La fresa es beneficiosa y la zanahoria también.  
179. La fresa es \*beneficiosas y la zanahoria también.  
180. La fresa es \*beneficioso y la zanahoria también.
181. El libro es didáctico y el periódico también.  
182. El libro es \*didácticos y el periódico también.  
183. El libro es \*didáctica y el periódico también.
184. El diccionario es didáctico y el glosario también.  
185. El diccionario es \*didácticos y el glosario también.  
186. El diccionario es \*didáctica y el glosario también.
187. El libro es anónimo y el artículo también.  
188. El libro es \*anónimos y el artículo también.  
189. El libro es \*anónima y el artículo también.
190. El cuento es anónimo y el manuscrito también.  
191. El cuento es \*anónimos y el manuscrito también.  
192. El cuento es \*anónima y el manuscrito también.
193. La guitarra es melodiosa y la flauta también.  
194. La guitarra es \*melodiosas y la flauta también.  
195. La guitarra es \*melodioso y la flauta también.
196. La samba es melodiosa y la rumba también.  
197. La samba es \*melodiosas y la rumba también.  
198. La samba es \*melodioso y la rumba también.
199. La cortina es bonita y la moqueta también.  
200. La cortina es \*bonitas y la moqueta también.  
201. La cortina es \*bonito y la moqueta también.
202. La guitarra es bonita y la armónica también.  
203. La guitarra es \*bonitas y la armónica también.  
204. La guitarra es \*bonito y la armónica también.
205. El laboratorio es conocido y el departamento también.

206. El laboratorio es \*conocidos y el departamento también.  
 207. El laboratorio es \*conocida y el departamento también.
208. El disco es conocido y el grupo también.  
 209. El disco es \*conocidos y el grupo también.  
 210. El disco es \*conocida y el grupo también.
211. El laboratorio es privado y el archivo también.  
 212. El laboratorio es \*privados y el archivo también.  
 213. El laboratorio es \*privada y el archivo también.
214. La conferencia es privada y la fiesta también.  
 215. La conferencia es \*privadas y la fiesta también.  
 216. La conferencia es \*privado y la fiesta también.
217. La película es pedagógica y la novela también.  
 218. La película es \*pedagógicas y la novela también.  
 219. La película es \*pedagógico y la novela también.
220. El diccionario es pedagógico y el tesoro también.  
 221. El diccionario es \*pedagógicos y el tesoro también.  
 222. El diccionario es \*pedagógica y el tesoro también.
223. El empleo es patético y el sueldo también.  
 224. El empleo es \*patéticos y el sueldo también.  
 225. El empleo es \*patética y el sueldo también.
226. La película es patética y la crítica también.  
 227. La película es \*patéticas y la crítica también.  
 228. La película es \*patético y la crítica también.
229. El narciso es decorativo y el lirio también.  
 230. El narciso es \*decorativos y el lirio también.  
 231. El narciso es \*decorativa y el lirio también.
232. La cortina es decorativa y la alfombra también.  
 233. La cortina es \*decorativas y la alfombra también.  
 234. La cortina es \*decorativo y la alfombra también.
235. El espejo es plateado y el florero también.  
 236. El espejo es \*plateados y el florero también.  
 237. El espejo es \*plateada y el florero también.
238. La espada es plateada y la diadema también.  
 239. La espada es \*plateadas y la diadema también.  
 240. La espada es \*plateado y la diadema también.

241. La cafetera es metálica y la tetera también.  
242. La cafetera es \*metálicas y la tetera también.  
243. La cafetera es \*metálico y la tetera también.
244. La ventana es metálica y la puerta también.  
245. La ventana es \*metálicas y la puerta también.  
246. La ventana es \*metálico y la puerta también.
247. El gramófono es viejo y el cronómetro también.  
248. El gramófono es \*viejos y el cronómetro también.  
249. El gramófono es \*vieja y el cronómetro también.
250. La cafetera es vieja y la tostadora también.  
251. La cafetera es \*viejas y la tostadora también.  
252. La cafetera es \*viejo y la tostadora también.
253. La ventana es amarilla y la escalera también.  
254. La ventana es \*amarillas y la escalera también.  
255. La ventana es \*amarillo y la escalera también.
256. El plátano es amarillo y el mango también.  
257. El plátano es \*amarillos y el mango también.  
258. El plátano es \*amarilla y el mango también.
259. La cámara es automática y la calculadora también.  
260. La cámara es \*automáticas y la calculadora también.  
261. La cámara es \*automático y la calculadora también.
262. El gramófono es automático y el estéreo también.  
263. El gramófono es \*automáticos y el estéreo también.  
264. El gramófono es \*automática y el estéreo también.
265. La corona es dorada y la cadena también.  
266. La corona es \*doradas y la cadena también.  
267. La corona es \*dorado y la cadena también.
268. La espada es dorada y la lanza también.  
269. La espada es \*doradas y la lanza también.  
270. La espada es \*dorado y la lanza también.
271. La corona es auténtica y la sortija también.  
272. La corona es \*auténticas y la sortija también.  
273. La corona es \*auténtico y la sortija también.
274. El cuadro es auténtico y el grabado también.

275. El cuadro es \*auténticos y el grabado también.  
 276. El cuadro es \*auténtica y el grabado también.
277. El motociclismo es costoso y el automovilismo también.  
 278. El motociclismo es \*costosos y el automovilismo también.  
 279. El motociclismo es \*costosa y el automovilismo también.
280. El cuadro es costoso y el retrato también.  
 281. El cuadro es \*costosos y el retrato también.  
 282. El cuadro es \*costosa y el retrato también.
283. El empleo es prestigioso y el título también.  
 284. El empleo es \*prestigiosos y el título también.  
 285. El empleo es \*prestigiosa y el título también.
286. La academia es prestigiosa y la galería también.  
 287. La academia es \*prestigiosas y la galería también.  
 288. La academia es \*prestigioso y la galería también.
289. La academia es pública y la guardería también.  
 290. La academia es \*públicas y la guardería también.  
 291. La academia es \*público y la guardería también.
292. La plaza es pública y la avenida también.  
 293. La plaza es \*públicas y la avenida también.  
 294. La plaza es \*público y la avenida también.
295. El cuento es entretenido y el relato también.  
 296. El cuento es \*entretenidos y el relato también.  
 297. El cuento es \*entretenida y el relato también.
298. El motociclismo es entretenido y el judo también.  
 299. El motociclismo es \*entretenidos y el judo también.  
 300. El motociclismo es \*entretenida y el judo también.
301. El camino es corto y el atajo también.  
 302. El camino es \*cortos y el atajo también.  
 303. El camino es \*corta y el atajo también.
304. La conferencia es corta y la entrevista también.  
 305. La conferencia es \*cortas y la entrevista también.  
 306. La conferencia es \*corto y la entrevista también.
307. La samba es erótica y la lambada también.  
 308. La samba es \*eróticas y la lambada también.  
 309. La samba es \*erótico y la lambada también.



310. El tango es erótico y el flamenco también.  
 311. El tango es \*eróticos y el flamenco también.  
 312. El tango es \*erótica y el flamenco también.
313. El tango es rápido y el mambo también.  
 314. El tango es \*rápidos y el mambo también.  
 315. El tango es \*rápida y el mambo también.
316. El submarino es rápido y el hidroplano también.  
 317. El submarino es \*rápidos y el hidroplano también.  
 318. El submarino es \*rápida y el hidroplano también.
319. El contrato es justo y el pago también.  
 320. El contrato es \*justos y el pago también.  
 321. El contrato es \*justa y el pago también.
322. La sentencia es justa y la condena también.  
 323. La sentencia es \*justas y la condena también.  
 324. La sentencia es \*justo y la condena también.
325. El contrato es estricto y el reglamento también.  
 326. El contrato es \*estrictos y el reglamento también.  
 327. El contrato es \*estricta y el reglamento también.
328. La sentencia es estricta y la pena también.  
 329. La sentencia es \*estrictas y la pena también.  
 330. La sentencia es \*estricto y la pena también.
331. La boda es sencilla y la ceremonia también.  
 332. La boda es \*sencillas y la ceremonia también.  
 333. La boda es \*sencillo y la ceremonia también.
334. El vestido es sencillo y el velo también.  
 335. El vestido es \*sencillos y el velo también.  
 336. El vestido es \*sencilla y el velo también.
337. La huella es misteriosa y la reliquia también.  
 338. La huella es \*misteriosas y la reliquia también.  
 339. La huella es \*misterioso y la reliquia también.
340. La iglesia es misteriosa y la cripta también.  
 341. La iglesia es \*misteriosas y la cripta también.  
 342. La iglesia es \*misterioso y la cripta también.
343. El camino es feo y el pueblo también.

344. El camino es \*feos y el pueblo también.  
345. El camino es \*fea y el pueblo también.
346. La plaza es fea y la basílica también.  
347. La plaza es \*feas y la basílica también.  
348. La plaza es \*feo y la basílica también.
349. El gobierno es autoritario y el ejército también.  
350. El gobierno es \*autoritarios y el ejército también.  
351. El gobierno es \*autoritaria y el ejército también.
352. La aristocracia es autoritaria y la realeza también.  
353. La aristocracia es \*autoritarias y la realeza también.  
354. La aristocracia es \*autoritario y la realeza también.
355. El gobierno es poderoso y el parlamento también.  
356. El gobierno es \*poderosos y el parlamento también.  
357. El gobierno es \*poderosa y el parlamento también.
358. La aristocracia es poderosa y la burguesía también.  
359. La aristocracia es \*poderosas y la burguesía también.  
360. La aristocracia es \*poderoso y la burguesía también.

## APPENDIX 2: Fillers

1. El mecánico astuto trabaja en el garaje.
2. El mecánico astuto \*trabajan en el garaje.
3. El mecánico astuto está en el garaje.
  
4. La embajadora exiliada trabaja en Nueva York.
5. La embajadora exiliada \*trabajan en Nueva York.
6. La embajadora exiliada está en Nueva York.
  
7. El sociólogo ocupado escribe en la oficina.
8. El sociólogo ocupado \*escriben en la oficina.
9. El sociólogo ocupado está en la oficina.
  
10. La novelista prolífica escribe en la cabaña.
11. La novelista prolífica \*escriben en la cabaña.
12. La novelista prolífica está en la cabaña.
  
13. El toro bravo corre en el campo.
14. El toro bravo \*corren en el campo.
15. El toro bravo está en el campo.
  
16. La competidora sueca corre en la pista.
17. La competidora sueca \*corren en la pista.
18. La competidora sueca está en la pista.
  
19. El cirujano rico almuerza en el restaurante.
20. El cirujano rico \*almuerzan en el restaurante.
21. El cirujano rico está en el restaurante.
  
22. La empleada bancaria almuerza en la cafetería.
23. La empleada bancaria \*almuerzan en la cafetería.
24. La empleada bancaria está en la cafetería.
  
25. El niño creativo baila en la escuela.
26. El niño creativo \*bailan en la escuela.
27. El niño creativo está en la escuela.
  
28. La coreógrafa francesa baila en Los Ángeles.
29. La coreógrafa francesa \*bailan en Los Ángeles.
30. La coreógrafa francesa está en Los Ángeles.
  
31. La solista asignada canta en el escenario.
32. La solista asignada \*cantan en el escenario.

33. La solista asignada está en el escenario.
34. La monja bondadosa canta en el coro.  
35. La monja bondadosa \*cantan en el coro.  
36. La monja bondadosa está en el coro.
37. El perro descuidado duerme en el garaje.  
38. El perro descuidado \*duermen en el garaje.  
39. El perro descuidado está en el garaje.
40. La psicóloga cansada duerme en el sofá.  
41. La psicóloga cansada \*duermen en el sofa.  
42. La psicóloga cansada está en el sofá.
43. El cocodrilo americano nada en la laguna.  
44. El cocodrilo americano \*nadan en la laguna.  
45. El cocodrilo americano está en la laguna.
46. La ballena blanca nada en el mar.  
47. La ballena blanca \*nadan en el mar.  
48. La ballena blanca está en el mar.
49. El loro colorado vive en la selva.  
50. El loro colorado \*viven en la selva.  
51. El loro colorado está en la selva.
52. La doncella secuestrada vive en la torre.  
53. La doncella secuestrada \*viven en la torre.  
54. La doncella secuestrada está en la torre.
55. El ajo sabroso crece en el jardín.  
56. El ajo sabroso \*crecen en el jardín.  
57. El ajo sabroso está en el jardín.
58. La planta nutritiva crece en la selva.  
59. La planta nutritiva \*crecen en la selva.  
60. La planta nutritiva está en la selva.
61. El anillo lustroso brilla en el estante.  
62. El anillo lustroso \*brillan en el estante.  
63. El anillo lustroso está en el estante.
64. La joya egipcia brilla en el museo.  
65. La joya egipcia \*brillan en el museo.  
66. La joya egipcia está en el museo.

67. El leopardo africano caza en la sabana.  
68. El leopardo africano \*cazan en la sabana.  
69. El leopardo africano está en la sabana.
70. La pantera negra caza en el valle.  
71. La pantera negra \*cazan en el valle.  
72. La pantera negra está en el valle.
73. El académico venezolano estudia en la universidad.  
74. El académico venezolano \*estudian en la universidad.  
75. El académico venezolano está en la universidad.
76. La investigadora meticulosa estudia en el archivo.  
77. La investigadora meticulosa \*estudian en el archivo.  
78. La investigadora meticulosa está en el archivo.
79. El neurólogo pediátrico enseña en la universidad.  
80. El neurólogo pediátrico \*enseñan en la universidad.  
81. El neurólogo pediátrico está en la universidad.
82. La profesora simpática enseña en el instituto.  
83. La profesora simpática \*enseñan en el instituto.  
84. La profesora simpática está en el instituto.
85. El panadero perezoso lee en la cafetería.  
86. El panadero perezoso \*leen en la cafetería.  
87. El panadero perezoso está en la cafetería.
88. La bibliotecaria desocupada lee en la sala.  
89. La bibliotecaria desocupada \*leen en la sala.  
90. La bibliotecaria desocupada está en la sala.
91. El caballero heróico lucha en el castillo.  
92. El caballero heróico \*luchan en el castillo.  
93. El caballero heróico está en el castillo.
94. La armada británica lucha en el Pacífico.  
95. La armada británica \*luchan en el Pacífico.  
96. La armada británica está en el Pacífico.
97. El obispo católico medita en la capilla.  
98. El obispo católico \*meditan en la capilla.  
99. El obispo católico está en la capilla.

100. La abuela piadosa medita en la catedral.  
101. La abuela piadosa \*meditan en la catedral.  
102. La abuela piadosa está en la catedral.
103. El muchacho travieso grita en la calle.  
104. El muchacho travieso \*gritan en la calle.  
105. El muchacho travieso está en la calle.
106. La maestra enojada grita en el pasillo.  
107. La maestra enojada \*gritan en el pasillo.  
108. La maestra enojada está en el pasillo.
109. El soldado herido descansa en el hospital.  
110. El soldado herido \*descansan en el hospital.  
111. El soldado herido está en el hospital.
112. La turista agotada descansa en el hotel.  
113. La turista agotada \*descansan en el hotel.  
114. La turista agotada está en el hotel.
115. El genio artístico pinta en la galería.  
116. El genio artístico \*pintan en la galería.  
117. El genio artístico está en la galería.
118. La retratista boliviana pinta en el estudio.  
119. La retratista boliviana \*pintan en el estudio.  
120. La retratista boliviana está en el estudio.
121. El alumno aburrido dibuja en su cuaderno.  
122. El alumno aburrido \*dibujan en su cuaderno.  
123. El alumno aburrido está en su escritorio.
124. La artista talentosa dibuja en el parque.  
125. La artista talentosa \*dibujan en el parque.  
126. La artista talentosa está en el parque.
127. El equipo colombiano juega en el estadio.  
128. El equipo colombiano \*juegan en el estadio.  
129. El equipo colombiano está en el estadio.
130. La tenista japonesa juega en Nueva York.  
131. La tenista japonesa \*juegan en Nueva York.  
132. La tenista japonesa está en Nueva York.
133. El músico nervioso fuma en el camerino.

134. El músico nervioso \*fuman en el camerino.  
135. El músico nervioso está en el camerino.
136. La administradora estresada fuma en su oficina.  
137. La administradora estresada \*fuman en su oficina.  
138. La administradora estresada está en su oficina.
139. El novio abandonado llora en su habitación.  
140. El novio abandonado \*lloran en su habitación.  
141. El novio abandonado está en su habitación.
142. La viuda deprimida llora en el cementerio.  
143. La viuda deprimida \*lloran en el cementerio.  
144. La viuda deprimida está en el cementerio.
145. El aficionado expulsado protesta en la entrada.  
146. El aficionado expulsado \*protestan en la entrada.  
147. El aficionado expulsado está en la entrada.
148. La traidora sentenciada protesta en la corte.  
149. La traidora sentenciada \*protestan en la corte.  
150. La traidora sentenciada está en la corte.
151. El aventurero chileno esquía en los Andes.  
152. El aventurero chileno \*esquían en los Andes.  
153. El aventurero chileno está en los Andes.
154. La deportista suiza esquía en los Alpes.  
155. La deportista suiza \*esquían en los Alpes.  
156. La deportista suiza está en los Alpes.
157. El abogado australiano practica en Buenos Aires.  
158. El abogado australiano \*practican en Buenos Aires.  
159. El abogado australiano está en Buenos Aires.
160. La oradora inquieta practica en el auditorio.  
161. La oradora inquieta \*practican en el auditorio.  
162. La oradora inquieta está en el auditorio.
163. El auto nuevo ruge en la pista.  
164. El auto nuevo \*rugen en la pista.  
165. El auto nuevo está en la pista.
166. La turbina anticuada ruge en el hangar.  
167. La turbina anticuada \*rugen en el hangar.

168. La turbina anticuada está en el hangar.
169. El farmacéutico argentino reside en Buenos Aires.  
170. El farmacéutico argentino \*residen en Buenos Aires.  
171. El farmacéutico argentino está en Buenos Aires.
172. La directora excéntrica reside en Los Ángeles.  
173. La directora excéntrica \*residen en Los Ángeles.  
174. La directora excéntrica está en Los Ángeles.
175. El asesino contratado espera en el vestíbulo.  
176. El asesino contratado \*esperan en el vestíbulo.  
177. El asesino contratado está en el vestíbulo.
178. La prisionera condenada espera en su celda.  
179. La prisionera condenada \*esperan en su celda.  
180. La prisionera condenada está en su celda.
181. El barco vacío flota en la laguna.  
182. El barco vacío \*flotan en la laguna.  
183. El barco vacío está en la laguna.
184. La balsa perdida flota en el río.  
185. La balsa perdida \*flotan en el río.  
186. La balsa perdida está en el río.
187. El paso peligroso comienza en el cañón.  
188. El paso peligroso \*comienzan en el cañón.  
189. El paso peligroso está en el cañón.
190. La carretera montañosa comienza en los Pirineos.  
191. La carretera montañosa \*comienzan en los Pirineos.  
192. La carretera montañosa está en los Pirineos.
193. El sendero romántico termina en el jardín.  
194. El sendero romántico \*terminan en el jardín.  
195. El sendero romántico está en el jardín.
196. La ruta arriesgada termina en la sierra.  
197. La ruta arriesgada \*terminan en la sierra.  
198. La ruta arriesgada está en la sierra.
199. El teléfono ruidoso suena en el vestíbulo.  
200. El teléfono ruidoso \*suenan en el vestíbulo.  
201. El teléfono ruidoso está en el vestíbulo.



202. La campana hermosa suena en el campanario.  
 203. La campana hermosa \*suenan en el campanario.  
 204. La campana hermosa está en el campanario.
205. El espectro espantoso aparece en la torre.  
 206. El espectro espantoso \*aparecen en la torre.  
 207. El espectro espantoso está en la torre.
208. La tabla informativa aparece en el apéndice.  
 209. La tabla informativa \*aparecen en el apéndice.  
 210. La tabla informativa está en el apéndice.
211. El ogro adormilado ronca en la cueva.  
 212. El ogro adormilado \*roncan en la cueva.  
 213. El ogro adormilado está en la cueva.
214. La oficinista dormida ronca en su asiento.  
 215. La oficinista dormida \*roncan en su asiento.  
 216. La oficinista dormida está en su asiento.
217. La arqueóloga fanática acampa en el Sáhara.  
 218. La arqueóloga fanática \*acampan en el Sáhara.  
 219. La arqueóloga fanática está en el Sáhara.
220. El montañero peruano acampa en los Andes.  
 221. El montañero peruano \*acampan en los Andes.  
 222. El montañero peruano está en los Andes.
223. El rabino sincero ora en la sinagoga.  
 224. El rabino sincero \*oran en la sinagoga.  
 225. El rabino sincero está en la sinagoga.
226. La pecadora arrepentida ora en la catedral.  
 227. La pecadora arrepentida \*oran en la catedral.  
 228. La pecadora arrepentida está en la catedral.
229. El científico marino bucea en las Bahamas.  
 230. El científico marino \*bucean en las Bahamas.  
 231. El científico marino está en las Bahamas.
232. La bióloga mexicana bucea en el Caribe.  
 233. La bióloga mexicana \*bucean en el Caribe.  
 234. La bióloga mexicana está en el Caribe.

235. El oso hambriento pesca en el río.  
236. El oso hambriento \*pescan en el río.  
237. El oso hambriento está en el río.
238. La financiera retirada pesca en las Bahamas.  
239. La financiera retirada \*pescan en las Bahamas.  
240. La financiera retirada está en las Bahamas.
241. El extranjero testarudo regatea en el mercado.  
242. El extranjero testarudo \*regatean en el mercado.  
243. El extranjero testarudo está en el mercado.
244. La vendedora sabia regatea en el quiosco.  
245. La vendedora sabia \*regatean en el quiosco.  
246. La vendedora sabia está en el quiosco.
247. El cocinero renombrado cocina en el restaurante.  
248. El cocinero renombrado \*cocinan en el restaurante.  
249. El cocinero renombrado está en el restaurante.
250. La criada italiana cocina en el hotel.  
251. La criada italiana \*cocinan en el hotel.  
252. La criada italiana está en el hotel.
253. El ejecutivo obeso desayuna en su despacho.  
254. El ejecutivo obeso \*desayunan en su despacho.  
255. El ejecutivo obeso está en su despacho.
256. La conductora apresurada desayuna en su vehículo.  
257. La conductora apresurada \*desayunan en su vehículo.  
258. La conductora apresurada está en su vehículo.
259. El canguro capturado boxea en su jaula.  
260. El canguro capturado \*boxean en su jaula.  
261. El canguro capturado está en su jaula.
262. La campeona cubana boxea en el gimnasio.  
263. La campeona cubana \*boxean en el gimnasio.  
264. La campeona cubana está en el gimnasio.
265. El caballo enfermo convalece en el establo.  
266. El caballo enfermo \*convalecen en el establo.  
267. El caballo enfermo está en el establo.
268. La víctima asustada convalece en el hospital.

269. La víctima asustada \*convalecen en el hospital.  
 270. La víctima asustada está en el hospital.
271. El enfermero infectado estornuda en la clínica.  
 272. El enfermero infectado \*estornudan en la clínica.  
 273. El enfermero infectado está en la clínica.
274. La señora delgada estornuda en el aeropuerto.  
 275. La señora delgada \*estornudan en el aeropuerto.  
 276. La señora delgada está en el aeropuerto.
277. El político protegido desembarca en el aeropuerto.  
 278. El político protegido \*desembarcan en el aeropuerto.  
 279. El político protegido está en el aeropuerto.
280. La pasajera brasileña desembarca en San Francisco.  
 281. La pasajera brasileña \*desembarcan en San Francisco.  
 282. La pasajera brasileña está en San Francisco.
283. El monstruo acuático reaparece en la bahía.  
 284. El monstruo acuático \*reaparecen en la bahía.  
 285. El monstruo acuático está en la bahía.
286. La enemiga vengativa reaparece en la secuela.  
 287. La enemiga vengativa \*reaparecen en la secuela.  
 288. La enemiga vengativa está en la cárcel.
289. El tirano malvado reina en el archipiélago.  
 290. El tirano malvado \*reinan en el archipiélago.  
 291. El tirano malvado está en el archipiélago.
292. La soberana tiránica reina en el castillo.  
 293. La soberana tiránica \*reinan en el castillo.  
 294. La soberana tiránica está en el castillo.
295. El destacamento especializado ataca en el este.  
 296. El destacamento especializado \*atacan en el este.  
 297. El destacamento especializado está en el este.
298. La brigada rusa ataca en el norte.  
 299. La brigada rusa \*atacan en el norte.  
 300. La brigada rusa está en el norte.
301. El anciano respetado habla en el coloquio.  
 302. El anciano respetado \*hablan en el coloquio.

303. El anciano respetado está en el coloquio.
304. La arquitecta exitosa habla en el auditorio.
305. La arquitecta exitosa \*hablan en el auditorio.
306. La arquitecta exitosa está en el auditorio.
307. El cautivo torturado sufre en su celda.
308. El cautivo torturado \*sufren en su celda.
309. El cautivo turturado está en su celda.
310. La huérfana traumatizada sufre en el orfanato.
311. La huérfana traumatizada \*sufren en el orfanato.
312. La huérfana traumatizada está en el orfanato.
313. El juego entero cabe en la caja.
314. El juego entero \*cabén en la caja.
315. El juego entero está en la caja.
316. La bolsa llena cabe en el baúl.
317. La bolsa llena \*cabén en el baúl.
318. La bolsa llena está en el baúl.
319. El geranio híbrido florece en el patio.
320. El geranio híbrido \*florecen en el patio.
321. El geranio híbrido está en el patio.
322. La orquídea púrpura florece en la ribera.
323. La orquídea púrpura \*florecen en la ribera.
324. La orquídea púrpura está en la ribera.
325. El candidato próspero triunfa en la elección.
326. El candidato próspero \*triunfan en la elección.
327. El candidato próspero está en San Francisco.
328. La modista imaginativa triunfa en la pasarela.
329. La modista imaginativa \*triunfan en la pasarela.
330. La modista imaginativa está en la pasarela.
331. El funcionario designado participa en la investigación.
332. El funcionario designado \*participan en la investigación.
333. El funcionario designado está en la capital.
334. La secretaria discreta participa en la transacción.
335. La secretaria discreta \*participan en la transacción.
336. La secretaria discreta está en el banco.

337. El cartero maltratado persevera en el trabajo.  
338. El cartero maltratado \*perseveran en el trabajo.  
339. El cartero maltratado está en la camioneta.
340. La carpintera trabajadora persevera en el taller.  
341. La carpintera trabajadora \*perseveran en el taller.  
342. La carpintera trabajadora está en el taller.
343. El fotógrafo experto colabora en el proyecto.  
344. El fotógrafo experto \*colaboran en el proyecto.  
345. El fotógrafo experto está en la gala.
346. La reportera inquisitiva colabora en la investigación.  
347. La reportera inquisitiva \*colaboran en la investigación.  
348. La reportera inquisitiva está en la capital.
349. El paramédico musculoso compite en el torneo.  
350. El paramédico musculoso \*compiten en el torneo.  
351. El paramédico musculoso está en la ambulancia.
352. La poetisa española compite en el concurso.  
353. La poetisa española \*compiten en el concurso.  
354. La poetisa española está en la librería.
355. El árbitro frustrado intercede en la pelea.  
356. El árbitro frustrado \*interceden en la pelea.  
357. El árbitro frustrado está en el vestuario.
358. La consejera respetuosa intercede en el conflicto.  
359. La consejera respetuosa \*interceden en el conflicto.  
360. La consejera respetuosa está en la sala.

**APPENDIX 3: Mean Acceptance Rates by Condition for Native Speakers**

**Table 9. Mean acceptance rates by condition for native speakers**

		Condition		
		NUM/GEN	*NUM	*GEN
Participant	NS001	97.50	0.00	0.00
	NS002	100.00	0.00	5.00
	NS003	100.00	5.00	2.50
	NS004	97.50	2.50	7.50
	NS005	97.50	0.00	0.00
	NS006	97.50	2.50	0.00
	NS007	97.50	10.00	2.50
	NS008	100.00	2.50	2.50
	NS009	100.00	5.00	10.00
	NS010	92.50	2.50	5.00
	NS011	97.50	2.50	7.50
	NS012	97.50	0.00	5.00
<i>Average</i>		97.92	2.71	3.96
<i>(SD)</i>		(2.09)	(2.91)	(3.28)
<i>(Range)</i>		(92.5-100)	(0-10)	(0-10)

#### APPENDIX 4: Mean Acceptance Rates by Condition for Learners

**Table 10. Mean acceptance rates by condition for learners**

Participant	Condition	Condition		
		NUM/GEN	*NUM	*GEN
L2001		100.00	67.50	97.50
L2003		100.00	2.50	67.50
L2004		87.50	15.00	45.00
L2006		95.00	7.50	7.50
L2007		77.50	57.50	87.50
L2008		97.50	5.00	15.00
L2009		92.50	92.50	55.00
L2010		97.50	15.00	17.50
L2011		75.00	47.50	57.50
L2012		95.00	5.00	7.50
L2014		92.50	10.00	22.50
L2016		100.00	0.00	5.00
L2017		100.00	7.50	5.00
L2018		97.50	10.00	22.50
L2019		97.50	0.00	7.50
L2022		92.50	22.50	55.00
L2023		97.50	10.00	42.50
L2024		97.50	10.00	72.50
L2025		60.00	50.00	47.50
L2026		92.50	0.00	27.50
<i>Average</i>		92.25	21.75	38.25
<i>(SD)</i>		(10.26)	(26.40)	(28.72)
<i>(Range)</i>		(60-100)	(0-92.5)	(5-97.5)

**APPENDIX 5: Combined, Number, and Gender d' Scores for Native Speakers**

**Table 11. Native speaker d' scores for combined conditions, and for both number and gender agreement separately**

		Combined d' scores	Number d' scores	Gender d' scores
Participant	NS001	3.412	3.412	3.412
	NS002	3.412	4.052	3.189
	NS003	3.285	3.189	3.412
	NS004	2.549	2.772	2.404
	NS005	3.412	3.412	3.412
	NS006	2.971	2.772	3.412
	NS007	2.471	2.292	2.772
	NS008	3.412	3.412	3.412
	NS009	3.044	3.189	2.932
	NS010	2.277	2.404	2.181
	NS011	2.549	2.772	2.404
	NS012	2.772	3.412	2.549
	<i>Average</i>	2.964	3.091	2.958
	<i>(SD)</i>	(.427)	(.502)	(.477)
	<i>(Range)</i>	(2.277-3.412)	(2.292-4.052)	(2.181-3.412)

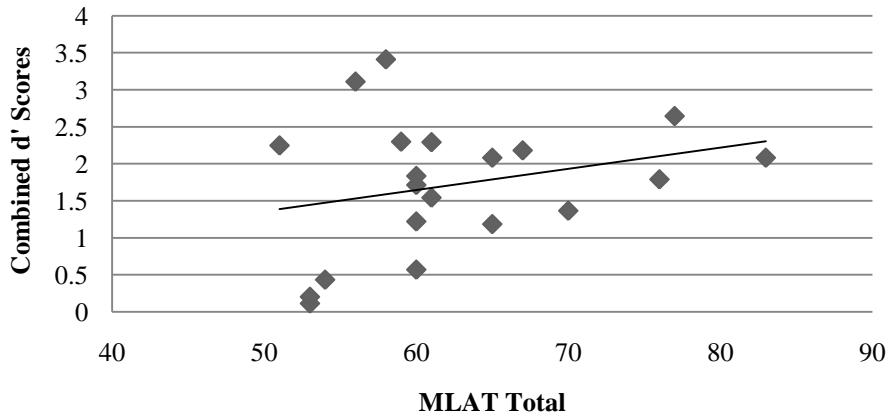


**APPENDIX 6: Combined, Number, and Gender d' Scores for Learners**

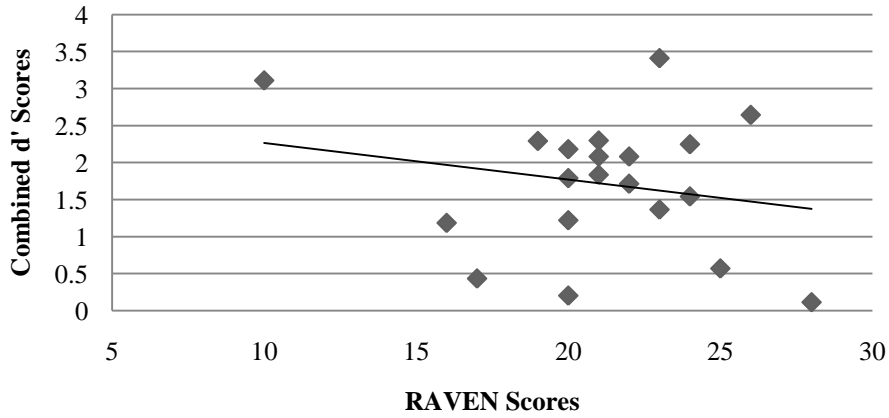
**Table 12. Learner group d' scores for combined conditions, and for both number and gender agreement separately**

		Combined d' scores	Number d' scores	Gender d' scores
Participant	L2001	1.365	1.705	0.640
	L2003	2.299	3.412	1.705
	L2004	1.184	1.546	0.902
	L2006	2.181	2.181	2.181
	L2007	0.111	0.400	-0.279
	L2008	2.292	2.549	2.119
	L2009	0.568	0.000	0.929
	L2010	2.082	2.119	2.047
	L2011	0.433	0.521	0.343
	L2012	2.248	2.326	2.181
	L2014	1.714	1.924	1.552
	L2016	3.412	4.052	3.189
	L2017	3.111	3.044	2.181
	L2018	2.082	2.292	1.552
	L2019	2.645	3.412	2.404
	L2022	1.220	1.552	0.929
	L2023	1.835	2.292	1.520
	L2024	1.542	2.292	0.963
	L2025	0.201	0.179	0.223
	L2026	1.790	3.044	1.441
	<i>Average</i>	1.716	2.042	1.436
	<i>(SD)</i>	(0.908)	(1.115)	(0.855)
	<i>(Range)</i>	(0.111-3.412)	(0.000-4.052)	(-0.279-3.189)

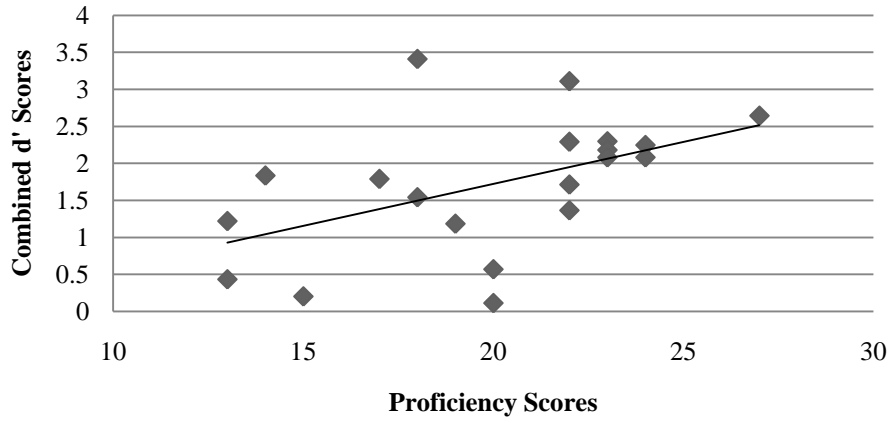
**APPENDIX 7: Combined, Number, and Gender d' Scores Plotted Against MLAT Total, RAVEN, and Proficiency Scores**



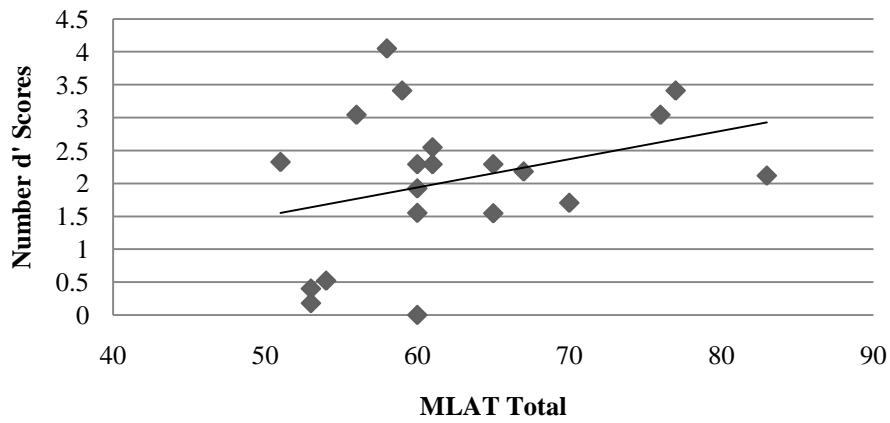
**Figure 3. Combined d' scores plotted against MLAT Total scores**



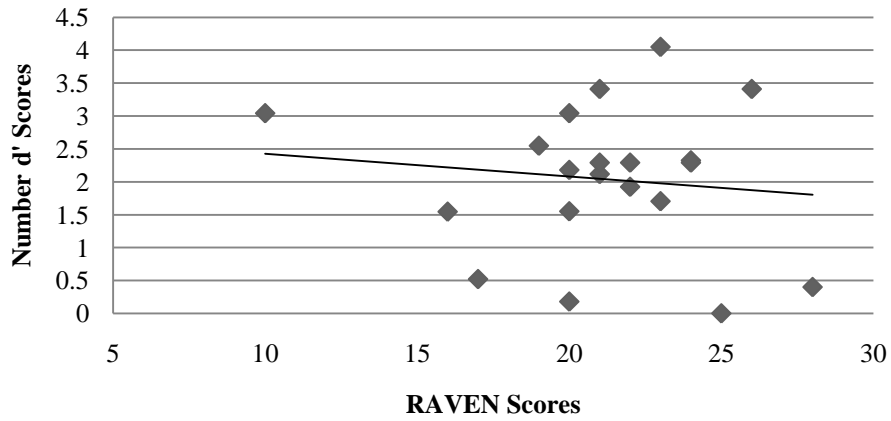
**Figure 4. Combined d' scores plotted against RAVEN scores**



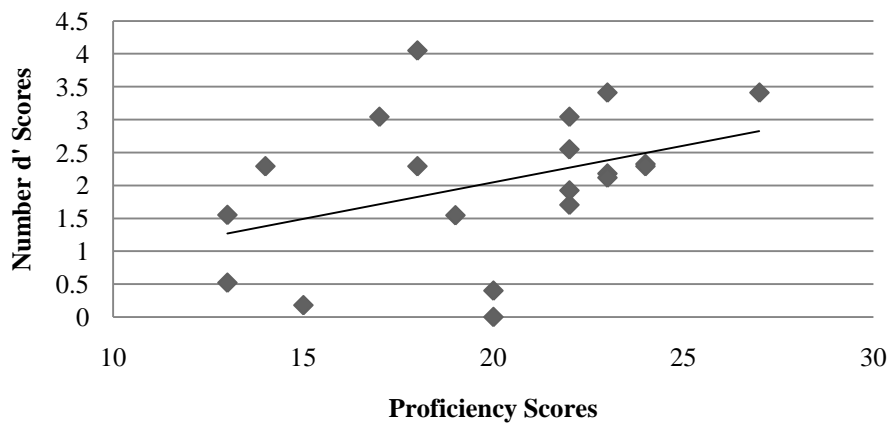
**Figure 5. Combined d' scores plotted against proficiency scores**



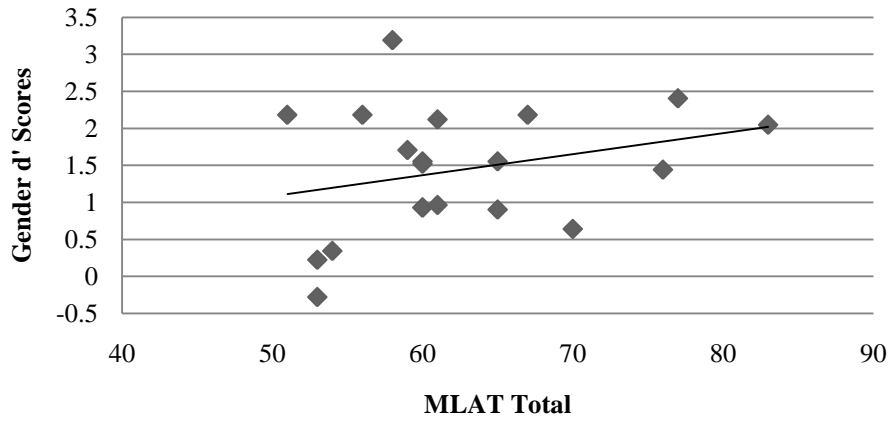
**Figure 6. The d' scores for number agreement plotted against MLAT Total scores**



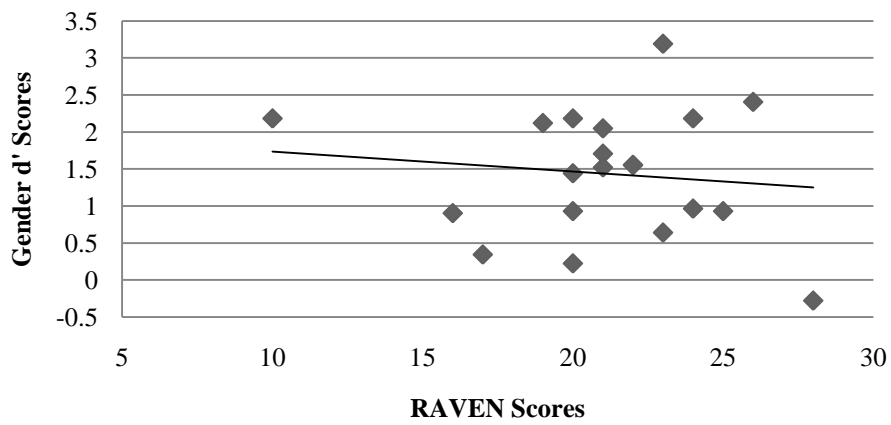
**Figure 7.** The d' scores for number agreement plotted against RAVEN scores



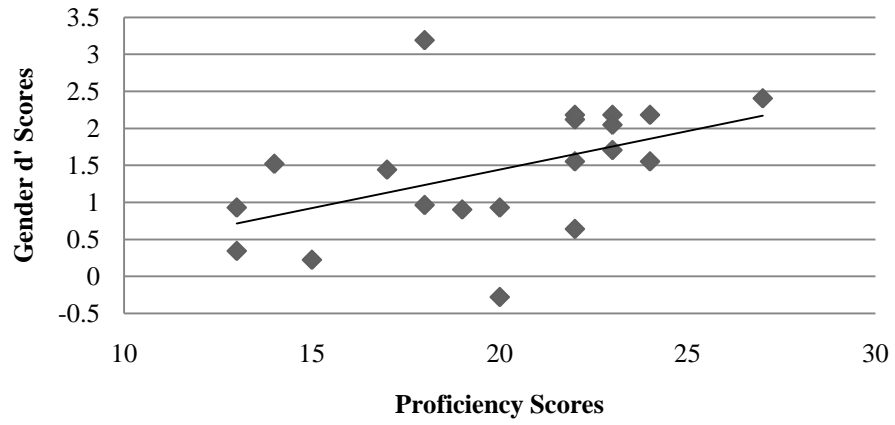
**Figure 8.** The d' scores for number agreement plotted against proficiency scores



**Figure 9.** The d' scores for gender agreement plotted against MLAT Total scores



**Figure 10.** The d' scores for gender agreement plotted against RAVEN scores



**Figure 11.** The  $d'$  scores for gender agreement plotted against proficiency scores