

MORPHOLOGICAL VARIABILITY IN LONG-DISTANCE SUBJECT-VERB  
AGREEMENT: A STUDY OF NATIVE AND NONNATIVE PROCESSING

BY

Sally Ocampo

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 of Arts.

Committee:

---

Chairperson: Alison Gabriele, Ph.D.

---

Member: Robert Fiorentino, Ph.D.

---

Member: Utako Minai, Ph.D.

Date Defended: March 1, 2013

The Thesis Committee for Sally Ocampo  
certifies that this is the approved version of the following thesis:

MORPHOLOGICAL VARIABILITY IN LONG-DISTANCE SUBJECT-VERB  
AGREEMENT: A STUDY OF NATIVE AND NONNATIVE PROCESSING

Committee:

---

Chairperson: Alison Gabriele, Ph.D.

---

Member: Robert Fiorentino, Ph.D.

---

Member: Utako Minai, Ph.D.

Date Approved: February 13, 2013

## ABSTRACT

This study examined morphosyntactic variability in Spanish-speaking learners of English in order to determine the effects of two linguistic factors on the establishment of subject-verb agreement: structural distance and plural markedness. The study also investigated the effects of task demands on processing of agreement morphology. Participants completed an online reading task using a moving window self-paced reading paradigm. In order to explore whether increases in structural distance between agreeing elements leads to decreases in sensitivity to agreement violations (Clahsen & Felser, 2006; Gabriele et al., in press), agreement between subject and verb was established across a prepositional phrase (such as *in warm southern Mexico*) or a structurally more complex relative clause (*who hunted in Mexico*). Subject number (singular or plural) was also manipulated in the task in order to examine whether the marked plural feature facilitated agreement establishment even as distance increased (Wagers et al., 2009). In order to determine whether task effects cause native speakers to show learner-like patterns of agreement variability (MacDonald, 2006; Hopp, 2010; Lopez Prego, 2012), a group of native English speakers was placed under a concurrent digit load as they completed the self-paced reading task. The results revealed that L2 learner agreement was affected by structural distance, as learners became less sensitive to violations in the relative clause intervener condition. Weak effects of plural markedness emerged in the learner results as well, indicated by greater sensitivity to errors in pairwise comparisons to the plural subject-relative clause intervener condition over the singular subject counterpart. Finally, weak similarities in variability between the L2 learner and the native speaker group under a concurrent processing load tentatively suggest that learner variability may be caused by general processing limitations, not deficits in L2 grammatical knowledge.

## ACKNOWLEDGMENTS

So many people were crucial to the completion of this work. I would like to thank the Department of Linguistics at the University of Kansas for providing students like me with both the theoretical knowledge of linguistics and the critical thinking skills necessary to conduct research. I am especially grateful for my advisor, Alison Gabriele, for her constant advice, support, and patience. It is clear that you devote so much of yourself to your students, and I could not have possibly completed this project without your help. Thank you also to my committee members, Robert Fiorentino and Utako Minai, and Joan Sereno for your helpful feedback and support.

I would also like to extend thanks to all of the members of the Research in Acquisition and Processing group. Your kind and thoughtful constructive criticism and advice have been immensely helpful through the entire research process. I consider myself fortunate to have belonged to a group filled with such talented researchers. I would like to thank Alonso Canales especially for his help with the development of my stimuli, and I owe thanks to Jose Alemán-Bañón and María Teresa Martínez García for their willingness to help with participant recruitment.

I am grateful for the Latin American and Spanish communities at the University of Kansas and in Lawrence for providing such helpful participants. Not only did many take time from their jobs, teaching positions, or own research to participate, but many were helpful in recruiting their own friends and family members to help as well.

I am eternally grateful to my family and friends for their love and support throughout the past three years. A special thanks to my parents, Glen and Gail Nelson. I cannot even begin to

list in such a small space all the ways in which you have supported me, so I thank you here for teaching me to value education, hard work, and diligence. Sam Nelson, Emily Nelson, Santa Berlioz, Eduardo Ocampo, Nathalie Ocampo, Mark Weissman, Adrienne Johnson, and Jon Coffee: thank you for your encouragement and advice. Finally, I dedicate this work to my wonderful husband, Jorge Guillermo Ocampo. Your love and patience were absolutely essential to the completion of this project. Thank you for inspiring me, for challenging me, and for encouraging me.

## TABLE OF CONTENTS

1	Introduction.....	1
2	Linguistic Factors Impacting Morphological Variability.....	4
2.1	Long-Distance Agreement in L2 Learners.....	4
2.2	Subject Noun Number.....	10
3	Processing Burdens in L2 Learners and Native Speakers.....	14
4	Research Questions and Predictions.....	17
5	Methods.....	20
5.1	Participants.....	20
5.2	Materials.....	21
5.3	Procedure.....	24
6	Results.....	25
6.1	Descriptive Statistics.....	25
6.1.1	Comprehension Accuracy.....	25
6.1.2	Digit Accuracy.....	26
6.1.3	Reading Times.....	27
6.2	2x2 ANOVA Results.....	31

6.2.1	L2 English Results: Prepositional Phrase Intervener.....	32
6.2.2	L2 English Results: Relative Clause Intervener.....	33
6.2.3	Native Speaker Results: Prepositional Phrase Intervener.....	33
6.2.4	Native Speaker Results: Relative Clause Intervener.....	34
6.2.5	Stressed Native Speaker Results: Prepositional Phrase Intervener.....	34
6.2.6	Stressed Native Speaker Results: Relative Clause Intervener.....	34
6.3	Paired Comparisons.....	35
6.3.1	L2 Learner Results.....	37
6.3.2	Native Speaker Results.....	37
6.3.3	Stressed Native Speaker Results.....	38
6.4	Summary of Results.....	39
7	Discussion.....	41
7.1	Effects of Structural Distance.....	42
7.2	Plural Markedness Effects.....	44
7.3	Processing Limitations.....	47
8	Conclusion.....	50

References.....52

Appendix: Target Stimuli.....55



## 1. Introduction

One of the primary differences between first and second language acquisition is the fact that late L2 learner performance, even at advanced stages of development, is quite variable. Although L1 learners can generally expect (barring impairment) to acquire their language to native levels, the same expectation cannot be held by adult second language learners even after years of study (e.g. Johnson & Newport, 1989; Birdsong & Mollis, 2001). Although this variability may be found in various domains of language use, variability in the use of inflectional morphology, such as subject-verb agreement and gender agreement, has been well-documented in second language acquisition literature (i.e., MacDonald, 2000, 2006; White et al., 2004; Ullman, 2004; Jiang, 2004,2007,2011; Clahsen & Felser, 2006; Keating, 2009; Hopp, 2010; Coughlin & Tremblay, in press). Production studies (Lardiere, 1998; White, 2003) have observed this faulty agreement resolution even in advanced learners with years of English exposure. For example, the sentence in (1) is an utterance from an advanced L1 Turkish/L2 English speaker (White, 2003, p. 134) who shows inflectional variability within just one sentence:

(1) And she cleans...the house. And *wash* the dishes. And, uh, she makes the bed.

Comprehension studies (Johnson & Newport, 1989; Keating, 2009; Clahsen et al., 2010; Jiang, 2004) have also found evidence of inflectional variability in late L2 learners in the form of a lack of sensitivity to agreement errors, using grammaticality judgments or reading times.

Although the presence of such inflectional variability is well-established in previous studies, the nature of this variability is still heavily debated. As a result, the primary goal of current research on the subject is to determine the source of variability in an effort to determine whether qualitative differences exist between native speaker and L2 learner storage and

processing of inflectional morphology and agreement resolution. Based on evidence from production and comprehension studies, some researchers have taken critical-period-based approaches to L2 morphological variability, proposing that deficits in the representation of L2 grammar prevents learners from using inflectional morphology accurately. One theory in this camp, the Shallow Structure Hypothesis (Clahsen & Felser, 2006; Clahsen et al., 2010; Keating, 2009) specifies that agreement can only be established locally for L2 learners. According to this proposal, long-distance agreement (for example, agreeing elements that occur in separate phrases) places such a burden on the L2 parser that it must parse “shallowly” and may be unable to establish a relationship between the agreeing elements. However, as Hopp (2010) discusses, an overview of the relevant literature reveals several open issues that must be considered when forming a theory of morphological variability, two of which will be discussed here:

First, studies have suggested that morphological variability is neither random nor universal across all grammatical constructions; rather it is systematic in that certain structures seem to be more likely to cause variability than others. For example, several studies have found that learners show more variability in detection and use of past-tense and subject-verb agreement morphology than present progressive inflections (MacDonald, 2006; Johnson & Newport, 1989; Jiang, 2004; MacDonald, 2006). Furthermore, research has indicated that various linguistic factors may affect variability seen within one structure. For example, Keating (2009) found that increasing the distance between a noun and adjective resulted in insensitivity to gender agreement errors among L2 learners of Spanish. On the other hand, Wagers et al. (2009) found that the presence of a plural demonstrative may actually facilitate accurate agreement for native English speakers, even as distance between demonstrative and noun increases (for example, *those face-making monkeys*).

Secondly, recent work has suggested that task demands are a factor of morphological variability in the L2 or even the native language (MacDonald, 2006; Hopp, 2010; Lopez-Prego, 2012). In order to explore task effects, Hopp (2010) and Lopez-Prego (2012) issued a grammaticality judgment task as well as a task that placed native speakers and L2 learners under an additional processing burden (speeded grammaticality judgment) while investigating sensitivity to agreement errors. Both studies found decreased sensitivity to errors in the speeded grammaticality judgment tasks for L2 learners and native speakers, suggesting that a more taxing task may significantly affect sensitivity to agreement even when sensitivity is clearly present in less strenuous tasks. Additional studies have found that this variability may be lessened as proficiency in the L2 increases (Coughlin & Tremblay, in press; Gabriele et al., in press).

While these open issues are still in need of further investigation, they do indicate that a general theory of morphological variability such as the Shallow Structure Hypothesis (Clahsen & Felser, 2006) may very well be too broad to answer for the effects of linguistic factors and general cognitive processing capabilities on successful use of inflectional morphology. The current study, therefore, aims to investigate the effects of several different factors on advanced L2 English learners' sensitivity to subject-verb agreement errors, in which incorrect verb number inflections create agreement mismatches. Sensitivity to errors of subject-verb agreement errors will be examined since verb number morphology seems to be particularly vulnerable in L2 English (MacDonald 2000, 2006; Jiang, 2004). First, we examine the role of structural distance between the agreeing elements. To this end, two different types of intervening material—a prepositional phrase (e.g., *from scenic western France*) or relative clause (e.g., *who moved from France*)—will be inserted between subject nouns and verbs in order to investigate the effects of increasing distance and complexity on sensitivity to agreement violations. Second, we examine

the role of subject noun number in order to learn whether plural number facilitates agreement establishment for advanced L2 learners, as Wagers et al. (2009) finds is the case for native speakers. Finally, we examine whether native speakers perform similarly to L2 learners when placed under a processing burden in order to further contribute to the discussion of morphological variability as a problem either in grammatical deficiency or processing limitations. In the following sections, I review the relevant literature in each of these three domains.

## **2. Linguistic Factors Impacting Morphological Variability**

### **2.1 Long-Distance Agreement in L2 learners**

Recent research (Keating, 2009, 2010; Coughlin & Tremblay in press; Gabriele et al., in press) has suggested that increasing the distance between two agreeing elements further exacerbates morphosyntactic variability in L2 learner comprehension. Some researchers (Clahsen et al., 2010; Keating, 2009) have proposed that the lack of agreement error sensitivity in the presence of long-distance agreement is predicted by the Shallow Structure Hypothesis, which proposes that non-local agreement dependencies (i.e., cases in which agreeing elements are located in different syntactic phrases) result in non-native like processing (Clahsen & Felser, 2006). Keating (2009) investigates the predictions of the Shallow Structure Hypothesis in an eye-tracking study of gender agreement in beginner, intermediate, and advanced L2 Spanish. The study manipulates distance between a subject noun and its modifying adjective by placing the adjective in one of three syntactic positions: within the same noun phrase as the subject noun, in an adjacent verb phrase, or in a subordinate clause next to the verb phrase (see 2a-2c).

2a) **Within-phrase agreement:**

Un libro aburrido es mas difícil de leer que un  
AMSG bookMSG boringMSG issG more difficultMSG to read than aMSG

libro interesante.  
bookMSG interestingsG

‘A boring book is more difficult to read than an interesting book.’

2b) **Agreement across one phrase:**

Un libro es bastante pequeño cuando tiene solo treinta  
AMSG bookMSG is quite smallMSG when hasG only thirty

páginas.  
pagesFPL

‘A book is quite small when it has only thirty pages.’

2c) **Agreement across two phrases:**

Un libro recibe bastante atención cuando es nuevo  
AMSG bookMSG receiveMSG much attention when issG newMSG

y popularMSG.  
and popular

‘A book receives a lot of attention when it is new and popular.’

His regression analysis of total reading times (first fixation and first pass) indicated that native speakers of Spanish were sensitive to agreement mismatches in all conditions. Advanced L2 Spanish learners were sensitive to errors only in the shortest condition (within-phrase), while intermediate and beginner L2 learners displayed no sensitivity to ungrammaticality in any of the distance manipulations. Keating takes the results as evidence for the Shallow Structure Hypothesis, since no group of learners was successful in resolving agreement dependencies within non-local domains. He also interprets his results as tentative support for the proposal that learners may never achieve native-like levels of processing, since even the advanced learners lacked sensitivity to long-distance agreement errors. However, Keating admits that a key limitation to the application of these results is that, in the stimuli, linear distance (number of intervening words) and structural distance (number of intervening syntactic nodes) were

confounded. For example, (2a) above shows that the agreeing adjective, *aburrido*, is adjacent to the subject noun, *libro*; furthermore, the two words occur within the same syntactic phrase, the determiner phrase (DP). However, in sentence (2b) the subject noun (*libro*) and agreeing adjective (*pequeño*) are separated linearly by two words (*es bastante*) and structurally by a phrase, since *libro* occurs in the DP while *pequeño* occurs in the adjacent verb phrase (VP). The noun and adjective in sentence (2c) are even farther apart linearly and structurally, with 5 words and 2 phrases (the VP and a subordinate clause) separating them. Because both linear and structural distance increase in each of the distance manipulations, it is impossible to determine from Keating (2009)'s results whether sensitivity to errors decreases due simply to an increasing number of words between the noun and adjective or whether the insensitivity lies in the increase of syntactic nodes separating agreeing elements. While the Shallow Structure Hypothesis does predict that learners will process non-local agreement in a nonnative-like way, it also proposes that L2 learners may not be sensitive to hierarchical structure (Clahsen & Felser, 2006). The confound of distance in Keating's (2009) study becomes problematic when attempting to determine whether or not this second prediction of the Shallow Structure Hypothesis holds true for advanced L2 learners, since it is unclear from the results whether learners lack sensitivity to errors due to an increase in linear or structural distance between the two agreeing elements.

In a self-paced reading study examining the roles of individual differences in antecedent-clitic agreement in L2 French, Coughlin & Tremblay (in press) eliminate this confound by successfully manipulating linear distance while holding structural distance constant (see examples 3a-3b).

3a) **Short Distance:**

Ce fruit Marie le mangera pour sa collation avant  
ThatMSG fruitMSG Marie Cl.MSG eatsG-FT for her snackFSG before

l'entretien

theMSG-interviewMSG

'Marie will eat that fruit for her snack before the interview.'

3b) **Long Distance:**

Ce fruit avant l'entretien Marie le mangera pour  
ThatMSG fruitMSG before theMSG-interviewMSG Marie Cl.MSG eatsG-FT for

sa collation

her snackFSG

'Marie will eat that fruit for her snack before the interview.'

Native speakers of French, intermediate L1 English/L2 French learners, and advanced L1 English/L2 French learners completed an offline acceptability judgment task, a self-paced reading task, a proficiency test, and working memory tasks in both French and English. The researchers found that advanced and intermediate learners displayed native-like error detection in the offline acceptability judgment task, suggesting that learners do in fact have access to inflectional morphology in the L2. In the self-paced reading task, intermediate learners showed no sensitivity to agreement errors; advanced learners, however, were sensitive to agreement errors regardless of the distance separating the antecedent and clitic. The researchers conclude that these results do not support the Shallow Structure Hypothesis, since successful non-local distance agreement resolution was possible for the advanced French learners. Furthermore, mixed model analyses revealed that high proficiency learners contributed more than intermediate learners to effects of grammaticality found, and correlative analyses revealed a weak positive correlation between working memory and sensitivity to agreement errors. These results suggest that individual differences, such as proficiency and working memory capacity, may be an important source of variability in L2 agreement.

Because Coughlin & Tremblay (in press) eliminated the linear-structural distance confound by isolating linear distance and found no distance effects, it is beneficial to consider other studies that isolate the effects of structural distance. Gabriele et al., (in press) examine the predictions of the Shallow Structure Hypothesis using noun-adjective agreement across two structural distances. The effects of L1-L2 differences and proficiency on sensitivity to errors were also investigated. The researchers conducted an evoked response potential (ERP) study of gender and number agreement in L2 Spanish, manipulating structural distance by placing the agreeing noun and adjective either within the same noun phrase (4a) or across one verb phrase boundary (4b). Linear distance was held constant at one intervening word. Four participant groups (L1 Spanish, low L1 English/L2 Spanish, intermediate L1 English/L2 Spanish, advanced L1 English/L2 Spanish) completed an RSVP grammaticality judgment task for the study.

(4a) El banco es un edificio muy seguro y  
 TheMSG bankMSG isSG aMSG buildingMSG very safeMSG and  
 el juzgado también.  
 theMSG courthouseMSG too  
 ‘The bank is a very safe building and the courthouse too.’

4b) El cuento es anónimo y el manuscrito  
 TheMSG storyMSG isSG anonymousMSG and theMSG manuscriptMSG  
 también.  
 too  
 ‘The story is anonymous and the manuscript too.’

The researchers propose that two possible results would satisfy the predictions of the Shallow Structure Hypothesis and therefore provide evidence in support of it: 1) L2 learners will not show sensitivity to agreement errors in either structural distance condition, regardless of proficiency. If learners are not sensitive to hierarchical structure, as the Shallow Structure



Hypothesis predicts, then the within-phrase and across-phrase conditions would be equally difficult for learners to process simply because linear distance is controlled in both conditions. 2) L2 learners may show sensitivity to agreement mismatches in the within-phrase condition, but not in the across-phrase condition due to the difficulty of establishing agreement across a phrase in a non-local dependency. Based on the proposals of the Shallow Structure Hypothesis, it is non-local agreement (i.e., across-phrase agreement) that would reduce learner sensitivity to agreement errors.

The study's ERP results revealed a consistent P600 in response for native Spanish speakers. The P600 is a positive waveform typically occurring within the 500-900ms window. Although it is argued to emerge in various contexts, the P600 is understood to occur in this window as a result of syntactic violations (Friederici et al., 1996) including (at least where native speakers are concerned) violations of agreement (Osterhout et al., 2004). While the native Spanish speaker group in Gabriele et al.'s (in press) study revealed a P600 in response to agreement errors overall, greater waveform positivities were elicited for within-phrase agreement errors relative to across-phrase errors. They interpret this result as evidence for reduced sensitivity overall in the establishment of agreement in the across-phrase condition. Both advanced and intermediate L2 Spanish learners displayed the same patterns, with consistent P600s in both groups for agreement errors and greater positivities in the within-phrase condition. The only group lacking the P600 pattern was the low proficiency L2 Spanish learner group. These participants displayed emerging positivities, with no difference in waveform amplitude observed based on the distance manipulation. The authors conclude that the results provide evidence for the proposal that accurate resolution of agreement dependencies is possible for L2 learners, even in non-local domains (contra the Shallow Structure Hypothesis). The results also

indicate, similar to Coughlin & Tremblay (in press), that sensitivity to agreement violations increases with proficiency. However, while Coughlin & Tremblay (in press) find that sensitivity to ungrammaticality is not a function of linear distance for learners of French, Gabriele et al. (in press) find that structural distance affects sensitivity to agreement errors for both learners of Spanish and native speakers.

The above studies all manipulate distance between agreeing elements in order to provide evidence either for or against the Shallow Structure Hypothesis. However, considering the available research on long-distance agreement as a whole, some issues arise. First, relatively few studies in the SLA literature examine long-distance agreement as a factor of morphosyntactic variability, and even fewer examine the effects of distance on subject-verb agreement. Second, in order to determine the accuracy of the predictions of the Shallow Structure Hypothesis, it is necessary to further examine the effects of linear distance and structural distance independently. For example, Coughlin & Tremblay (in press) find no distance effects while examining linear distance, and Gabriele et al. (in press) find that both native speakers and learners of Spanish are sensitive to hierarchical structure. In order to determine if learners and native speakers process agreement similarly or differently, additional research is needed to further tease apart the effects of each type of distance. The current study builds on these previous findings by examining the role of structural distance while linear distance is controlled with a grammatical structure (subject-verb agreement) that has not yet been examined within this domain.

## **2.2 Subject Noun Number**

Previous research in both the L1 literature (Bock & Miller, 1991; Bock & Cutting, 1992, etc.) and L2 literature (McCarthy, 2008; Lopez-Prego, 2012) has suggested that number features

on nouns may facilitate or impede accurate agreement dependency resolution between two elements. We will begin by reviewing native speaker research which suggests that in English, the plural is a marked feature that can impact the processing of agreement in different ways. First, the agreement attraction literature (Bock & Miller, 1991; Bock & Cutting, 1992; Eberhard, 1997; Solomon & Pearlmutter, 2004) offers evidence that a plurally marked intervening noun may disrupt accurate agreement resolution between a subject and verb. In the sentences below, only (5a) is a grammatical sentence in English. However, seminal production studies have found that, when native speakers produce subject-verb agreement errors, they tend to occur in sentences such as (5b), not (5a). For example, Bock & Miller (1991) found that, when native English speakers were provided with a preamble such as “*The key(keys) to the cabinets...*”, 79% of the subject-verb agreement errors that occurred were found in sentences with a singular subject and a plural intervening noun (for example, sentence 5b):

5a) The key to the cabinets is on the table.

5b) \*The key to the cabinets are on the table.

These results suggest that the plurally marked local noun *cabinets* disrupts the number of the singular subject noun *key* and results in the production of agreement errors. Comprehension studies (for example, Wagers, Phillips, & Lau, 2009) have also reported that the same subject noun-intervening noun number combination results in decreased sensitivity to agreement errors for native speakers.

One interesting facet of the agreement attraction proposal is that, while *subject noun SG-local noun PL* combinations are troublesome for native speakers of English, other combinations do not seem to cause similar disruptions in subject-verb agreement dependency resolution.

The sentence in (5d) is also a possible example of an agreement attraction error, since the presence of a singularly-marked local noun can potentially cause disruption in the agreement of the plurally-marked subject noun and verb:

5c) The keys to the cabinet are on the table.

5d) \*The keys to the cabinet is on the table.

However, according to production studies of agreement attraction (Eberhard et al, 2005), these errors are rarely elicited from native speakers. Comprehension studies (Pearlmutter et al., 1999; Wagers, Phillips, & Lau, 2009) similarly find that native speakers are significantly more sensitive to cases of agreement mismatch in sentences with *subject noun PL-intervening noun SG* constructions. These findings as a whole have led some researchers to suggest that the plural noun in English is marked; therefore an unmarked singular intervening noun does not disrupt the establishment of agreement while a plurally marked intervening noun may (Wagers, Lau, & Phillips, 2009; Eberhard, 1997; Pearlmutter et al., 1999).

A second possible impact of the plural feature in English is examined in Wagers et al., (2009), which proposes that plural markedness stands up to manipulations of long-distance agreement such that a plural number will facilitate the establishment of agreement even when linear distance separates two elements. The researchers examine the effect using agreement between English demonstratives and nouns (see 6a-6b) predicting that if the plural number feature is indeed marked on the demonstrative, the processor will be better able to “predict” a plurally-marked noun and the feature will be held in focal attention even as distance increases between the two elements. In a multi-response speed-accuracy trade-off task, in which linear distance between the agreeing elements was manipulated by the addition of one or two

adjectives, Wagers et al. found that native speakers were significantly faster in responding to agreement violations in adjacent conditions than non-adjacent conditions when the demonstrative was singular (6a).

**6a) Singular Demonstrative:**

That monkey/\*monkeys...

That mischievous monkey/\*monkeys...

That mischievous face-making monkey/\*monkeys...

**6b) Plural Demonstrative:**

Those monkeys/\*monkey...

Those mischievous monkeys/\*monkey...

Those mischievous face-making monkeys/\*monkey...

However, when the demonstrative was plural (6b), participants were equally fast at responding to all distance conditions. These results suggest that the plural number feature significantly reduces the effects of distance on processing agreement, and add corroborating evidence to the proposal that the plural is marked in English and can affect the way in which native speakers process agreement.

While these studies have contributed important findings to the L1 processing literature, no study has examined whether the facilitative effects of plural markedness will be observed when structural distance between two agreeing elements, instead of linear distance, is manipulated. Furthermore, the L2 literature includes several studies that examine the effects of marked features on learner agreement and find that learners do tend to behave differently depending on whether they encounter a marked or unmarked feature (McCarthy, 2008; Lopez-Prego, 2012). However, as far as we are aware, no L2 study has examined the effects of the plural subject on agreement establishment as structural distance between the agreeing elements is increased. The current study aims to address this potential interaction between number and structural distance by manipulating both structural distance (with prepositional phrase or relative

clause interveners) and subject noun number while examining sensitivity to agreement errors in all conditions.

### **3. Processing Burdens in L2 Learners and Native Speakers**

Recent studies (MacDonald, 2006; Hopp, 2010; Lopez-Prego, 2012) have suggested that morphological variability in learners may be tied more to processing limitations and not to grammatical deficits (as proposed by the Shallow Structure Hypothesis). In support of this proposal, these researchers have shown that native speakers perform similarly to learners when placed under a processing burden. In earlier work with native speakers only, researchers placed participants under a memory load (in the form of a concurrent digit load) while administering a reading comprehension task (Blackwell & Bates, 1995; Dick et al., 2001; Waters et al., 1995). These studies found that participants' accuracy in grammaticality judgment or comprehension decreased under the memory load, although not all syntactic structures were affected. Blackwell & Bates (1995), for example, found that participants were less successful in detecting agreement errors under the digit load, while other features such as word order were preserved until a higher memory load was presented. Some second language acquisition researchers (MacDonald, 2006; Hopp, 2010; Lopez-Prego, 2012) have taken these native speaker results as evidence that processing may rely, at least in part, on non-linguistic cognitive skills such as working memory capacity. Their studies place native speakers under similar processing burdens and compare their performance in grammaticality judgment tasks to that of learners.

MacDonald (2006) conducted two experiments in which she compared the performance of L2 English learners to that of stressed native speakers. In Experiment 1, a group of native speakers and a group of L2 learners were administered three individual difference measures: a

verbal working memory task, a gating task (testing phonological decoding) and a word detection task (testing processing speed). All subjects also participated in an auditorily presented grammaticality judgment task in which a variety of syntactic errors were possible (including subject-verb agreement errors). MacDonald predicted that the learners would show lower performance than native speakers on all individual tasks as well as the judgment task, since all tests were conducted in the participants' second language. Another prediction, based on the proposal that non-linguistic capacities are involved in linguistic processing, was that comparatively more complex linguistic structures should decrease the availability of general processing capabilities and therefore reduce sensitivity to grammatical errors (Blackwell & Bates, 1995; Dick et al., 2001; Waters et al., 1995; Waters et al., 2003). Results indicated that learners were significantly poorer than native speakers in all tasks, and that certain errors (i.e., word order) were more easily detected in the learner group than errors involving more complex morphosyntactic relationships (subject-verb agreement, plural markings, etc.).

The same predictions were extended to native speakers under an additional processing burden in Experiment 2. The participants completed the same cognitive measure tasks and were administered the grammaticality judgment task, this time with one of four processing burdens added: a low or high digit load (working memory), white noise (phonological decoding), or a speeded response requirement (processing speed). The results suggest that native speakers under a processing burden became more learner-like in their sensitivity to grammatical errors, as performance fell significantly in all processing burden conditions except for the low memory load. When analyzed by syntactic construction, results indicated that errors involving word order were not significantly affected by any of the three stressors, and were therefore detected easily by native speakers. However, quite similarly to the L2 learners in Experiment 1, stressed native

speakers displayed significantly lower accuracy in detecting errors involving articles or regular morphology. Stressed native speaker performance became most learner-like under the high memory load and the white noise condition. Due to the patterns of sensitivity found in both L2 learners and stressed native speakers, MacDonald concludes that the results most accurately correspond to the predictions of a proposal that includes general cognitive abilities as a factor in L2 morphological and syntactic variability.

In a study of near-native L2 German, Hopp (2010) compares learner and stressed native speaker performance on constructions such as subject-verb agreement and case marking and corroborates MacDonald's (2006) results. Hopp (2010) found that both near-native L2 learners and native speakers showed sensitivity to errors in an offline task and a self-paced reading task. The results also indicate that both learners and native speakers showed similar patterns of variability in error detection (with sensitivity to word order decreasing the least and sensitivity to case and gender decreasing the most) when placed under a processing burden in the form of a speeded grammaticality judgment task. Lopez-Prego (2012) reveals similar results with advanced L2 learners of Spanish that performed at native-like levels when detecting number and gender agreement errors in an untimed grammaticality judgment task. Although L2 learner sensitivity to errors decreased in a speeded grammaticality judgment task, Lopez-Prego (2012) also found that native speaker sensitivity decreased significantly in a similar task as speed increased.

Though taxing the processing capabilities of native speakers provides a different way to test general cognitive proposals, this method of testing within the field of second language acquisition is relatively nascent. As can be deduced from examining the current literature, extensive research must be conducted still in order to discover which structures are most



susceptible to processing burdens and which types of burdens are most effective in creating learner-like sensitivity to grammatical errors in native speakers. For example, although the studies discussed above have provided strong support for a processing limitation proposal, none have compared L2 learner and stressed native speaker performance while also examining the interaction of plural markedness effects and structural distance effects on subject-verb agreement. The current study aims to further contribute to the discussion of processing limitations by placing native speakers under a concurrent memory load while examining native speaker and learner sensitivity to long-distance subject-verb agreement errors.

#### 4. Research Questions and Predictions

The current study aims to further investigate potential sources of L2 variability in the processing of subject-verb agreement by L1 Spanish/ L2 English learners. The study addresses three main questions:

RQ1: Will L2 learners of English be less sensitive to subject-verb agreement errors when agreement is established across a more structurally complex intervening phrase?

We address this question by comparing learners' ability to detect agreement violations across two different types of phrases: prepositional phrases (7a) and relative clauses (7b).

7a) The tourist PP[in NP[ADJP[warm] NP[ADJP[southern] NP[Mexico]]]] often fishes in the ocean.

7b) The tourist CP[who TP[(*ed*) VP[VP[(*who*) hunted] PP[in NP[Mexico]]]]] often fishes in the ocean.

In (7a), the prepositional phrase *in warm southern Mexico* separates the subject *tourist* from the verb phrase *often fishes*. This intervening phrase requires agreement to be established over the distance of two main syntactic nodes, which include the prepositional phrase (*in warm*

*southern Mexico*) and the embedded noun phrase (*warm southern Mexico*). The relative clause *who hunted in Mexico* in sentence (7b), on the other hand, requires agreement to be established over the more complex structural distance of five syntactic nodes: the complementizer phrase (*who hunted in Mexico*), the tense phrase (TP, marked for past tense), the verb phrase (*hunt in Mexico*), the prepositional phrase (*in Mexico*), and the noun phrase (*Mexico*). The Shallow Structure Hypothesis (Clahsen & Felser 2006, 2010; Keating, 2009) proposes that learners do not process non-local agreement in a native-like way: specifically, learners are not able to establish agreement across phrases and show a lack of sensitivity to hierarchical structures. In the current study, one pattern of results would support these proposals. Since both distance conditions in the current study require agreement establishment across phrases, finding that learners are less sensitive to agreement errors than native speakers in both conditions (prepositional phrase intervener and relative clause intervener) could potentially provide support for the Shallow Structure Hypothesis. Finding that learners display a lack of sensitivity to agreement errors in both distance conditions would also suggest that learners are not sensitive to hierarchical structure, since the prepositional phrase intervener is relatively less complex compared to the relative clause intervener.

RQ2: Will the effects of plural markedness facilitate successful subject-verb agreement resolution for L2 learners as structural distance between the two agreeing elements is increased?

Wagers et al. (2009) found that the plural noun feature was able to survive long-distance agreement as linear distance increased, suggesting that the feature is marked in native English. However, the effects of plural markedness in conjunction with structural distance on inflectional variability have not yet been investigated in L2 learner groups. The present study examines this

question by comparing the effect of the plural subject within two structural distance conditions (8 and 9).

**8) Prepositional Phrase Intervener Conditions: Singular Subject and Plural Subject**

- a) The tourist in warm southern Mexico often fishes in the ocean.
- b) The tourists in warm southern Mexico often fish in the ocean.

**9) Relative Clause Intervener Conditions: Singular Subject and Plural Subject**

- a) The tourist who hunted in Mexico often fishes in the ocean.
- b) The tourists who hunted in Mexico often fish in the ocean.

In order to extend the predictions of plural markedness effects to L2 learners and structural distance, learners should show greater sensitivity to agreement errors in the plural subject conditions compared to the singular subject conditions, particularly as distance between the two agreeing elements increases (i.e., the relative clause intervener conditions).

RQ3: Will native speakers placed under a processing burden show variability similar to the L2 learners in their processing of agreement?

Several studies (MacDonald, 2006; Hopp, 2010; Lopez-Prego, 2012) have found that placing native speakers under a processing burden (such as speeded judgment tasks or memory loads) causes native speaker processing of agreement to become similar to L2 learner processing. However, these studies also reveal that not all grammatical constructions are equally susceptible to processing burdens. For example, MacDonald (2006) finds that word order errors are still easily detected by stressed native speakers, while regular past-tense morphology is less resilient. One grammatical construction that is difficult for even advanced L2 English learners to acquire is long-distance subject-verb agreement. At this point, no L2 study has attempted to compare L2 learner processing of long-distance subject-verb agreement errors with that of native speakers under a processing burden. To this end, the current study aims to place native speakers of English under a concurrent memory load while examining the same manipulations of structural

distance and subject number as are investigated in the L2 learner and control group. Finding that stressed native English speakers display similar patterns of sensitivity compared to the L2 learners will suggest (in line with MacDonald, 2006) that morphological variability in learners is due more to general processing limitations than to deficient representations in the L2 grammar.

Although the Shallow Structure Hypothesis does not predict a role for the L1 in morphosyntactic variability (Clahsen & Felser, 2006), other recent research (Hopp, 2010; Jiang, 2011; Gabriele et al., in press, etc.) has suggested that L1 transfer effects may influence learners' abilities to successfully establish agreement. Since the current study aims to investigate other factors (i.e., structural distance, plural markedness, and processing limitations) that may also influence the acquisition of agreement morphology, an L1/L2 pairing was selected such that both languages follow similar patterns of subject-verb agreement. Spanish was selected as the L1 language because, similar to English, it requires subjects and verbs to agree in number.

## **5. Methods**

### **5.1 Participants**

Three groups of participants took part in the current study. There were 28 native English speakers in the control group, and 28 native English speakers in the stressed native speaker group. All native English speakers were students at the University of Kansas, and received either extra credit in a linguistics or speech pathology course or were paid as compensation for their participation. 20 L1 Spanish/L2 English speakers formed the L2 learner group. The large majority of the Spanish-speaking participants were born and raised in Spanish-speaking countries throughout Latin America, and all participants' dialects of Spanish mark subject-verb agreement in writing. Although a few of the Spanish-speaking participants were initially exposed

to English at a young age, none received immersion instruction until after the age of 11 and all received instruction in Spanish in primary school. For all Spanish-speaking participants, Spanish was the language used at home.

**Table 1.** L1 Spanish/L2 English Participant Background Information  
**L2 English Participants (N=20)**

	<b>Mean</b>	<b>Min</b>	<b>Max</b>	<b>SD</b>
Age	28.95	18	40	7.09
AOI <sup>a</sup>	11.5	3	27	6.1
YOR <sup>b</sup>	3.57	.1667	12.83	3.52
Proficiency (50)	38.65	21	49	8.49

<sup>a</sup> Age of Acquisition, referring to the age when participants began studying English in any capacity

<sup>b</sup>Years of Residence in the United States

L2-English participants completed a standardized multiple-choice test (University of Cambridge) to assess their English proficiency. The test consisted of 50 incomplete sentences targeting either grammar or vocabulary. Participants were asked to choose the most appropriate word from four offered choices per item to correctly complete the sentence. Based on the performance of the majority of participants (*Mean*=38.5, see Table 1), the group was classified as consisting of advanced learners of English.

## 5.2 Materials

In order to examine the processing of subject-verb agreement, we used a non-cumulative, moving window, self-paced reading task. The task used a 2 (structural distance: prepositional phrase, relative clause) x 2 (subject noun number: singular, plural) x 2 (grammaticality:

grammatical, ungrammatical) design, resulting in eight experimental conditions as is shown in Table 2. Structural distance between the subject and verb was manipulated by inserting either a prepositional phrase or subject relative clause between the two elements, while linear distance was held constant at four words between subject and verb. We generated 24 sets of eight sentences. These sentences were then divided into four lists, each list containing one of the four sentences from the Prepositional Phrase conditions and one of the four sentences from the Relative Clause conditions. The sentences with Prepositional Phrase interveners and the sentences with Relative Clause interveners were further divided on each list into two experimental blocks.

**Table 2. Stimuli**

	<b>Prepositional Phrase Intervener</b>	<b>Relative Clause Intervener</b>
<b>SG-</b>	The tourist <i>in warm southern Mexico</i>	The tourist <i>who hunted in Mexico</i> often fishes
<b>GR</b>	often fishes in the ocean.	in the ocean.
<b>SG-</b>	The tourist <i>in warm southern Mexico</i>	The tourist <i>who hunted in Mexico</i> often *fish
<b>UG</b>	often *fish in the ocean.	in the ocean.
<b>PL-</b>	The tourists <i>in warm southern Mexico</i>	The tourists <i>who hunted in Mexico</i> often fish
<b>GR</b>	often fish in the ocean.	in the ocean.
<b>PL-</b>	The tourists <i>in warm southern Mexico</i>	The tourists <i>who hunted in Mexico</i> often
<b>UG</b>	often *fishes in the ocean.	*fishes in the ocean.

These sentences were counterbalanced across the four different lists, so that participants saw only one version of a sentence within each experimental block, and never saw two versions of a sentence on one list that were identical in both subject noun number and grammaticality. Within each list, the 48 target stimuli were interspersed among 96 fillers that were based on stimuli from

Bond et al. (2011). All fillers were grammatical sentences of English in which subject number and type of subordinate clause (embedded statement or question) were manipulated (see 10a-10b). The study's stimuli included a proportionally larger number of grammatical sentences compared to ungrammatical sentences in order to prevent participants from expecting or searching for grammatical errors instead of reading for comprehension.

10a) The waitress(es) said that the coffee for us was ready.

10b) The customer(s) asked if the hostess in the restaurant was serving the appetizers.

Stimuli within experimental blocks were randomized for each participant. In order to ensure that participants were reading for meaning as naturally as possible, comprehension questions occurred after all trials (11). Participants indicated their response by clicking on a "Yes" box or a "No" box presented on the same screen as the comprehension question.

11) **Sample Sentence:** The author from cold eastern Canada sometimes hikes in the morning.

**Comprehension Question:** Was this sentence about an author from eastern Canada?

For participants in the stressed native speaker group, a concurrent processing load of six digits was added to the self-paced reading task. The digit load was included in 100% of the sentence trials in order to consistently require participants to hold six digits in memory while reading sentences for comprehension. A digit memory check was included after each comprehension question. This task will be described in detail below.

All participants also completed an additional working memory measure, which will not be discussed in the present paper.

### 5.3 Procedure

Participants were tested individually in a quiet room with a computer. Before beginning the experiment, all participants were presented with and signed consent forms. Both native speakers and learners of English completed a language background questionnaire, which asked primarily for their language learning history and environments.

All participants completed the moving window self-paced reading task first. Instructions were presented in English for all groups, and each participant received seven practice trials. They were encouraged to read as naturally as possible and for meaning. Participants were also offered five short breaks throughout the task. For the L2 English group and native English control group, each moving window trial began with a series of dashes across the middle of an otherwise blank screen. Participants clicked the left mouse button to progress through a sentence one word at a time, uncovering one word while covering the previous word with each click of the mouse. After the last word of each sentence, the computer screen automatically presented the comprehension question. Participants chose each answer by clicking with the mouse on boxes labeled “Yes” and “No” below the comprehension question. After a choice was made, the computer screen immediately moved to the next trial with no feedback provided.

Participants in the stressed native speaker group performed the same procedure with a few additions. Each trial began with a fixation cross presented at the center of the screen (1500ms), after which the screen automatically progressed to the digit presentation screen. A string of six digits was presented simultaneously in the center of the screen for a short period of time (3000ms). The screen automatically progressed to the moving window screen, and participants progressed through the sentence-reading and comprehension question just as the other two groups did. After choosing each answer to the comprehension question, participants



were presented with a similar screen in which a string of digits was again simultaneously presented. Participants were to determine whether or not the newly presented string was identical to the original string. Choices were indicated by clicking with the mouse on boxes labeled “Yes” and “No” below the string. The string of digits was identical to the original string in 50% of trials. In the 50% of trials in which the string was not identical, two digits from the original string were transposed. Transpositions for these stimuli were balanced across all six locations in the strings. Once participants indicated their answer to the digit accuracy check, the screen progressed to the next trial.

All participants next completed the counting span task. Native speaker participants were then compensated, and participants in the L2 group completed the proficiency task as their final task. The entire experiment lasted 45-60 minutes for native speakers of English and 75-90 minutes for the L1 Spanish/L2 English group. Stimuli for the self-paced reading task and counting span task were presented with Paradigm (Tagliaferri, 2005), and the proficiency test was administered with pencil and paper.

## **6. Results**

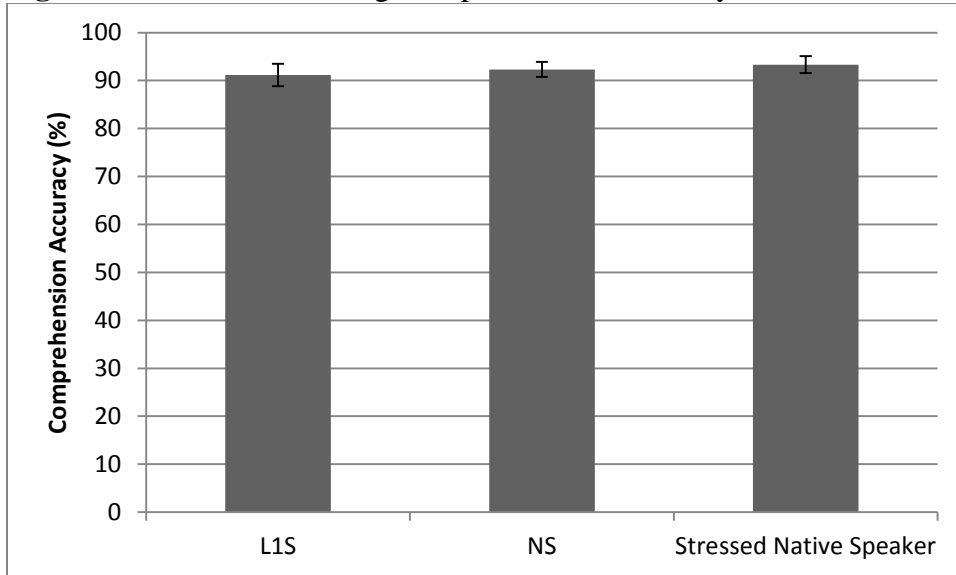
### **6.1 Descriptive Statistics**

#### **6.1.1 Comprehension Accuracy**

Comprehension accuracy rates were calculated for all participants in the three participant groups. All participants scored above 75% on the comprehension questions, so no participant was removed from analysis based on this criterion. The average accuracy rate for each of the three groups was above 90% (see Figure 1), suggesting that participants were reading for

meaning and that participants in the L2 English group were capable of comprehending sentences similarly to native speakers of English.

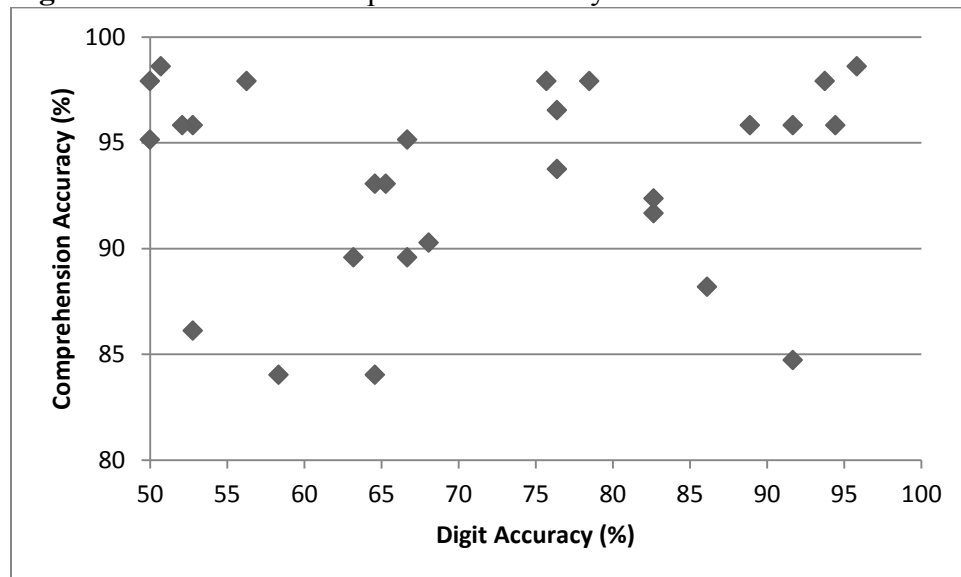
**Figure 1.** Self-Paced Reading Comprehension Accuracy Rates



### 6.1.2 Digit Accuracy

Within the native speaker group under the concurrent processing burden, digit accuracy was far more varied than comprehension scores. The average digit accuracy rate for the group was 71.3% ( $SD=15.4$ ), with a minimum score of 50% and a maximum score of 95.83%. In order to measure potential trade-off effects, a correlation analysis was run on the digit accuracy and comprehension accuracy scores of the NSM group, with no substantial correlation ( $r=.07, p=.72$ ) between participants' performance on the digit task and their ability to comprehend sentences (see Figure 2). One participant's data was excluded from analysis due to the fact that the participant selected "Yes" for all digit accuracy probes.

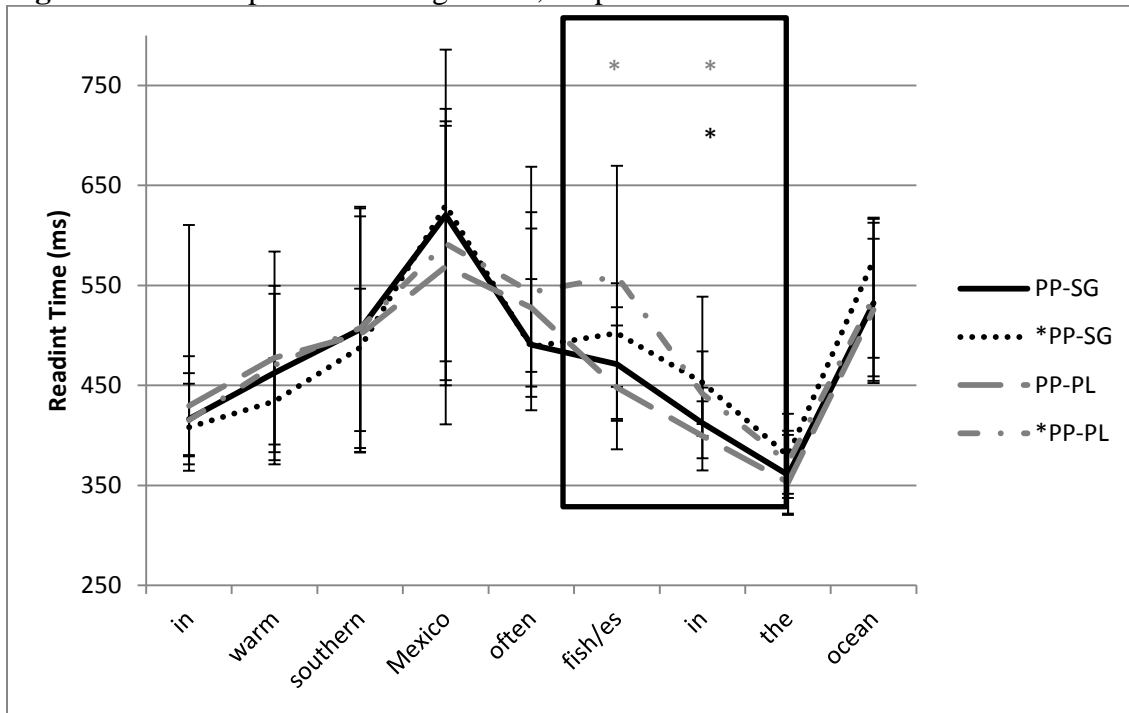
**Figure 2.** Stressed Native Speakers' Accuracy Scores



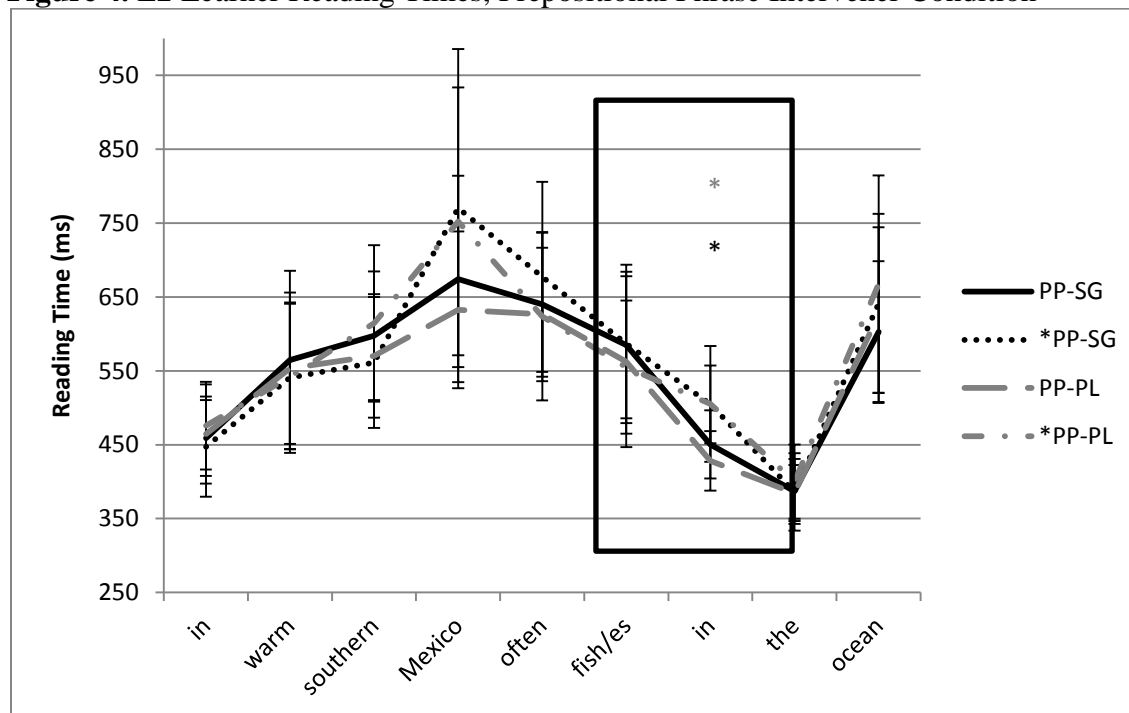
### 6.1.3 Reading Times

In the conducted analyses, any reading time that was two standard deviations above the individual participant's mean reading time for a given region was excluded. This restriction resulted in the removal of 5.3% of the reading time data. Figures 3 through 8 display reading times for each participant group for the prepositional phrase intervener condition (Figs. 3-5) and the relative clause intervener condition (Figs. 6-8). The verb (*fish/es*) is the critical region, and the two following regions (*in, the*) were also examined in analyses.

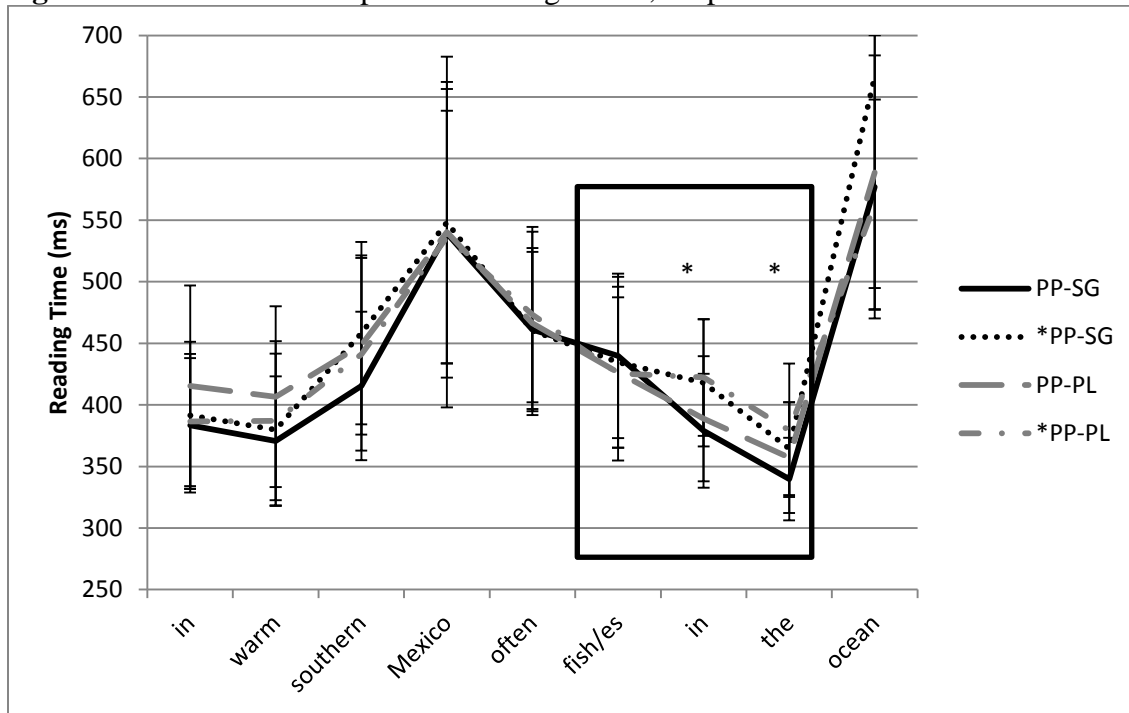
**Figure 3.** Native Speaker Reading Times, Prepositional Phrase Intervener Condition



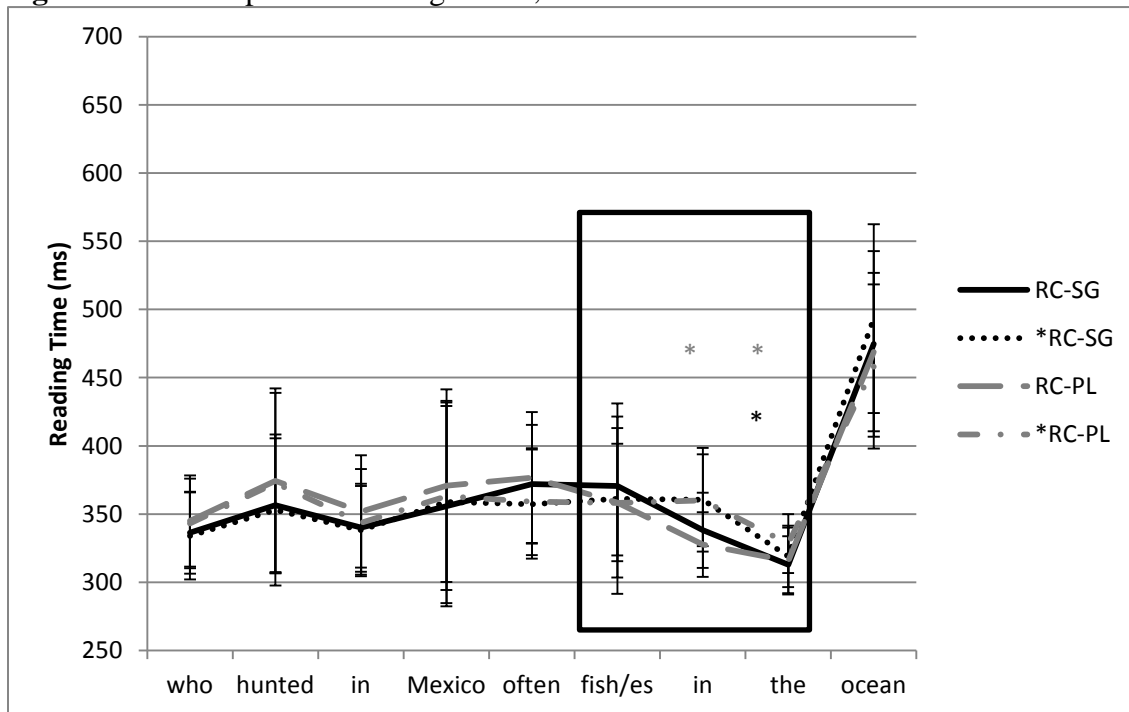
**Figure 4.** L2 Learner Reading Times, Prepositional Phrase Intervener Condition



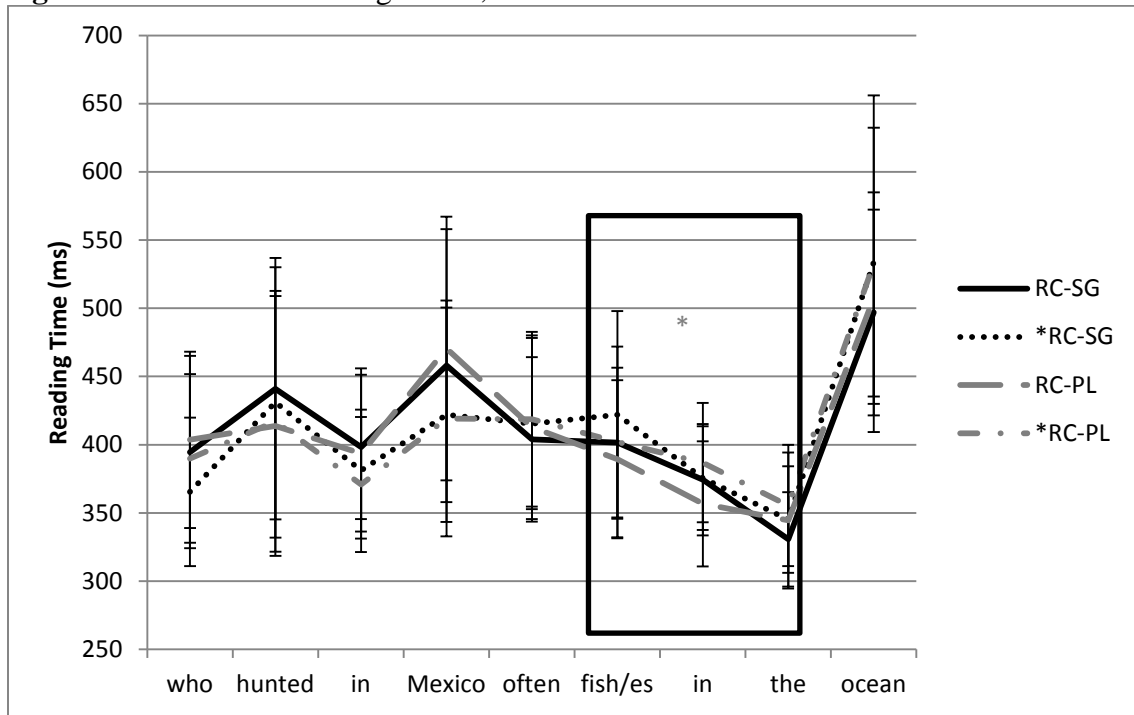
**Figure 5.** Stressed Native Speaker Reading Times, Prepositional Phrase Intervener Condition



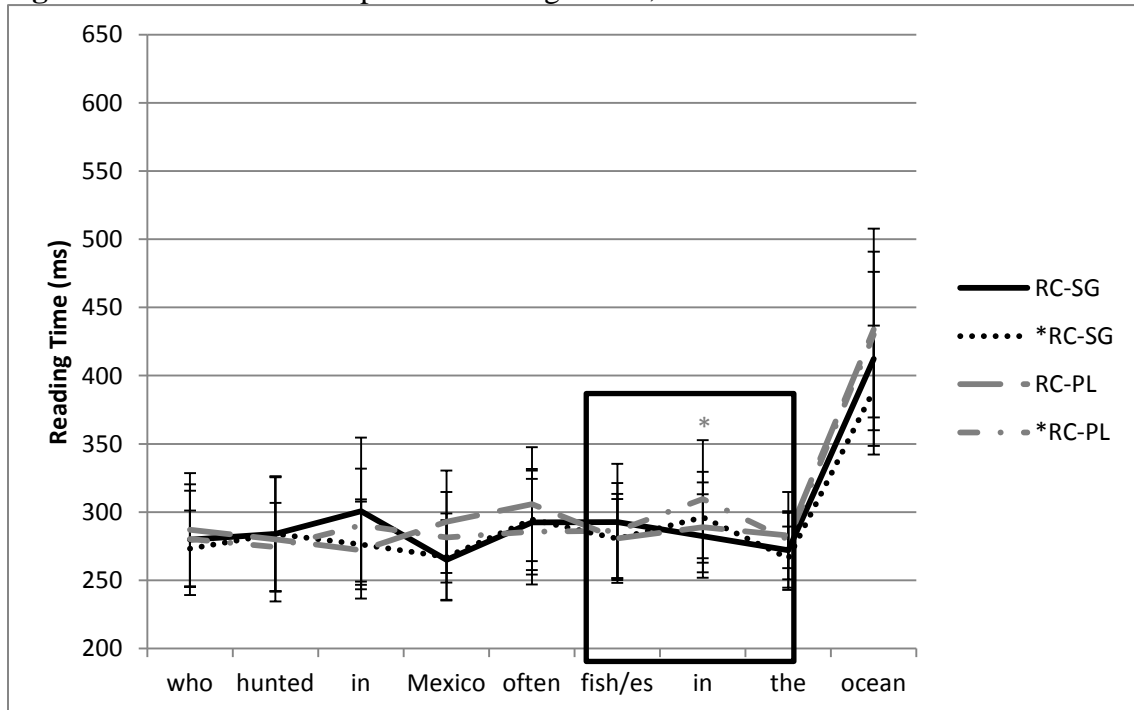
**Figure 6.** Native Speaker Reading Times, Relative Clause Intervener Condition



**Figure 7.** L2 Learner Reading Times, Relative Clause Intervener Condition



**Figure 8.** Stressed Native Speaker Reading Times, Relative Clause Intervener Condition



## 6.2 ANOVA Results

In order to investigate linguistic factors of L2 agreement and effects of processing limitations on native speaker agreement, a series of repeated-measures analyses of variance (ANOVAs) were conducted separately for each participant group<sup>1</sup>. 2x2 ANOVAs were conducted for prepositional phrase intervener conditions and relative clause intervener conditions within each group with subject noun number (singular, plural) and grammaticality as within-subjects factors. Each analysis was carried out at the critical region—the grammatically or ungrammatically marked verb--and the two following sentence regions. These segments were selected for analysis since previous research (see Jiang, 2011 for review) has indicated that sensitivity to grammatical errors may not become evident until the parser reaches the segments shortly after the error is encountered, termed the “spillover regions.” Table 3 illustrates the effects and interactions found at each region (beginning with Region 8, the critical region) for each participant group.

---

<sup>1</sup> Initially, a 2x2x2 repeated measures analysis of variance was conducted for each participant group, with intervener type (prepositional phrase, relative clause), subject noun number (singular, plural), and grammaticality (grammatical, ungrammatical) as factors. Results of the ANOVAs revealed interactions between intervener and grammaticality for all participant groups.

**Table 3.** Repeated Measures ANOVAs on L2 Learner, Native Speaker, and Stressed Native Speaker Reading Times for Subject Number and Grammaticality

		L2 English			Native Speakers			Stressed Native Speakers		
		Reg. 8, CR	Reg. 9	Reg. 10	Reg. 8, CR	Reg. 9	Reg. 10	Reg. 8, CR	Reg. 9	Reg. 10
<b>Prepositional Phrase</b>										
Num	<b>F</b>	1.53	0.642	0.56	0.822	2.206	0.489	1.068	0.358	2.213
	<b>p</b>	(.231)	(.433)	(.463)	(.373)	(.149)	(.491)	(.311)	(.555)	(.148)
Gram	<b>F</b>	0.016	10.554	0.736	8.1	20.295	6.914	0.076	8.945	4.251
	<b>p</b>	(.901)	(<.01)	(.402)	(<.01)	(<.001)	(<.05)	(.785)	(<.01)	(<.05)
			**		**	***	*		**	*
Num x Gram	<b>F</b>	0.038	0.565	1.247	5.885	0.011	0.002	0.015	0.056	0.009
	<b>p</b>	(.847)	(.462)	(.278)	(<.05)	(.919)	(.961)	(.903)	(.814)	(.924)
					*					
<b>Relative Clause</b>										
Num	<b>F</b>	2.116	.121	1.12	0.532	0.766	3.456	0.185	1.875	5.516
	<b>p</b>	(.162)	(.732)	(.303)	(.472)	(.389)	(.074)	(.67)	(.182)	(<.05)
							†			*
Gram	<b>F</b>	.871	2.466	1.873	0.083	9.881	4.044	0.258	5.205	0.6
	<b>p</b>	(.362)	(.133)	(.187)	(.776)	(<.01)	(.054)	(.616)	(<.05)	(.445)
						**	†		*	
Num x Gram	<b>F</b>	.044	2.808	.012	0.281	1.177	0.709	0.592	0.175	0.159
	<b>p</b>	(.835)	(.11)	(.914)	(.6)	(.288)	(.407)	(.448)	(.679)	(.693)

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

\*\* Effect is significant at the  $p < .01$  level

\*\*\* Effect is significant at the  $p < .001$  level

† Effect is marginally significant

### 6.2.1 L2 Learner Results: Prepositional Phrase Intervener

Results of the 2x2 ANOVA for L2 English learners' reading times within the prepositional phrase intervener conditions revealed a main effect of grammaticality (by participants ( $F_1$ ) and items ( $F_2$ )) ( $F_1(1,19)=10.554, p < .01$  (Reg. 9);  $F_2(1,23)=9.503, p < .01$  (Reg. 9)). The effect indicated that the L2 learners were reading the ungrammatical conditions more slowly than the grammatical conditions at that region.



### 6.2.2 L2 Learner Results: Relative Clause Intervener

Results of the 2x2 repeated measures ANOVA for L2 learners' times for the sentences with relative clause interveners revealed an effect of number ( $F_2(1,23)=4.587, p<.05$  (Reg. 8)) in the by items analysis only, due to overall slower reading times at that region in the plural subject conditions compared to the singular subject conditions. A marginal interaction between subject number and grammaticality was found in the by items analysis as well ( $F_2(1,23)=3.151, p=.089$  (Reg. 9)), indicating greater sensitivity to grammatical errors in the plural subject conditions compared to the singular subject conditions. However, no effects of grammaticality were found at any region within the relative clause intervener conditions.

### 6.2.3 Native Speaker Results: Prepositional Phrase Intervener

Results of the 2x2 ANOVA for native speakers' reading times within the prepositional phrase intervener conditions showed a main effect of grammaticality ( $F_1(1,27)=8.1, p<.01$  (Reg. 8);  $F_2(1,23)=32.588, p\leq.000$  (Reg. 8);  $F_1(1,27)=20.295, p<.001$  (Reg. 9);  $F_2(1,23)=10.211, p<.01$  (Reg. 9);  $F_1(1,27)=6.914, p<.05$  (Reg. 10);  $F_2(1,23)=3.86, p=.062$  (Reg. 10)). The grammaticality effect indicates that native speakers read ungrammatical critical regions more slowly than grammatical critical regions. The analysis also revealed an interaction between number and grammaticality ( $F_1(1,27)=5.885, p<.05$ ;  $F_2(1,23)=9.11, p<.01$ ) in Region 8 of the native speaker results. Due to this interaction, two post-hoc paired-samples t-tests were conducted. The first test compared reading times at the critical region (Region 8, the verb) for grammatical sentences within the singular subject condition (*The tourist in warm southern Mexico fishes...*) with those for ungrammatical sentences within the singular subject condition (*The tourist in warm southern Mexico fish...*). The results of this comparison indicated that native speakers' reading times were

not significantly slower at Region 8 of ungrammatical sentences compared to grammatical sentences when the subject noun was singular ( $t(27)=2.05, p=.248$ ). The second t-test compared reading times at the critical region (Region 8) for grammatical sentences within the plural subject condition (*The tourists in warm southern Mexico fish...*) to ungrammatical sentences from the same condition (*The tourists in warm southern Mexico fishes...*). The results of the test revealed that reading times at Region 8 of the ungrammatical sentences were significantly slower than grammatical counterparts when the subject noun was plural ( $t(27)=2.05, p<.01$ ).

#### **6.2.4. Native Speaker Results: Relative Clause Intervener**

Within the relative clause intervener conditions, native speakers' 2x2 ANOVA results showed a marginal effect of number by participants only ( $F_1(1,27)=3.456, p=.074$ ) at Region 10, due to slower reading times in the plural subject condition than the singular subject condition. A main effect of grammaticality was found at Region 9 ( $F_1(1,27)=9.881, p<.01$ ;  $F_2(1,23)=17.352, p\leq.000$ ), and a marginal effect of grammaticality was found at Region 10 ( $F_1(1,27)=4.044, p=.054$ ) in the by participants analysis only. Both of these effects emerged due to slower reading times for ungrammatical conditions compared to grammatical conditions.

#### **6.2.5. Stressed Native Speaker Results: Prepositional Phrase Intervener**

The stressed native speakers' results also demonstrated a main effect of grammaticality within the prepositional phrase-intervener conditions ( $F_1(1,27)=8.945, p<.01$  (Reg. 9);  $F_2(1,23)=8.47, p<.01$  (Reg. 9);  $F_1(1,27)=4.251, p<.05$  (Reg. 10);  $F_2(1,23)=7.209, p<.05$  (Reg. 10)). The effect revealed that reading times for those two regions were slower within the ungrammatical conditions than the grammatical conditions.

### 6.2.6 Stressed Native Speaker Results: Relative Clause Intervener

Within the relative clause intervener conditions, the stressed native speakers' results revealed a main effect of number in Region 10 ( $F_1(1,27)=5.516, p<.05$ ;  $F_2(1,23)=7.741, p<.05$ ). The effect was due to slower reading times in the plural subject conditions compared to the singular subject conditions. A main effect of grammaticality occurred in Region 9 ( $F_1(1,27)=5.205, p<.05$ ;  $F_2(1,23)=5.459, p<.05$ ), indicating slower reading times for the ungrammatical conditions than the grammatical conditions.

### 6.3 Pairwise Comparisons

Due to relatively small participant group sizes and a lack of robust effects or interactions in ANOVA results (particularly where number was concerned), exploratory paired samples t-tests were conducted in order to investigate any potential weaker effects of number and intervener type. These t-tests compared grammatical and ungrammatical reading times at the critical region and two spillover regions for the following sentence types: singular subject/prepositional phrase (*The tourist in warm southern Mexico often fishes/\*fish in the ocean.*), plural subject/prepositional phrase (*The tourists in warm southern Mexico often fish/\*fishes in the ocean.*), singular subject/relative clause (*The tourist who hunted in Mexico often fishes/\*fish in the ocean.*), and plural subject/relative clause (*The tourists who hunted in Mexico often fish/\*fishes in the ocean.*). Results of the pairwise comparison are presented in Table 4 below.

**Table 4.** Results of paired samples t-tests comparing reading times for grammatical and ungrammatical sentences, separated by intervener and subject number for L2 English learners, Native Speakers, and Stressed Native Speakers

		L2 English				Native Speakers				Stressed Native Speakers			
Prepositional Phrase		Reg. 8, CR		Reg. 9		Reg. 10		Reg. 8, CR		Reg. 9		Reg. 10	
Singular Sub.		Reg. 8, CR		Reg. 9		Reg. 10		Reg. 8, CR		Reg. 9		Reg. 10	
By Participants	<i>t</i>	-0.07	-3.18	0.07	-1.18	-3.21	-1.99	0.26	-2.73	-2.23			
	<i>p</i>	.94	<.01**	0.94	.25	<.01*	.06†	.08†	<.05*	<.05*			
By Items	<i>t</i>	-0.23	-2.81	-0.25	-1.25	-2.03	-1.49	0.19	-2.72	-3.08			
	<i>p</i>	0.82	<.01**	.81	.22	.053†	.15	.85	<.05*	<.01**			
Plural Sub.		Reg. 8, CR		Reg. 9		Reg. 10		Reg. 8, CR		Reg. 9		Reg. 10	
By Participants	<i>t</i>	0.19	-2.46	-1.28	-3.35	-3.50	-1.99	0.04	-1.79	-1.01			
	<i>p</i>	.85	<.05*	.21	<.01**	<.01**	.06†	.97	.08†	.32			
By Items	<i>t</i>	-0.16	-2.46	-1.34	-5.91	-3.29	-1.86	-1.51	-1.88	-1.34			
	<i>p</i>	.87	<.05*	.19	<.001***	<.01**	.08†	.14	.07†	.19			
Relative Clause		Reg. 8, CR		Reg. 9		Reg. 10		Reg. 8, CR		Reg. 9		Reg. 10	
By Participants	<i>t</i>	-0.58	0.11	-1.27	0.41	-2.14	-0.91	0.78	-1.38	0.86			
	<i>p</i>	.57	.92	.21	.69	<.05*	.37	.44	.18	.40			
By Items	<i>t</i>	-0.83	0.35	-1.42	1.60	-1.98	-0.83	1.21	-0.85	0.97			
	<i>p</i>	.42	.73	.17	.12	.06†	.41	.24	.40	.34			
Plural Sub.		Reg. 8, CR		Reg. 9		Reg. 10		Reg. 8, CR		Reg. 9		Reg. 10	
By Participants	<i>t</i>	-1.06	-2.27	-0.83	0.02	-3.46	-2.29	-0.53	-1.68	0.44			
	<i>p</i>	.30	<.05*	0.42	.99	<.01**	<.05*	.60	.11	.66			
By Items	<i>t</i>	-0.64	-2.54	-0.84	-0.26	-3.72	-1.39	0.17	-2.19	0.67			
	<i>p</i>	.53	<.05*	.41	.79	<.01**	<.18	.87	<.05*	.51			

### 6.3.1 L2 Learner Results

Results of the t-tests conducted for the prepositional phrase intervener conditions revealed that L2 learners were sensitive to agreement violations when the subject was singular at the first spillover region ( $t(19)=-3.18, p<.01$ , by participants;  $t(19)=-2.81, p<.01$ , by items). Learners were also sensitive to errors within the prepositional phrase when the subject noun was plural ( $t(19)=-2.46, p<.05$ , by participants;  $t(19)=-2.46, p<.05$ , by items). However, results of t-tests conducted in the relative clause conditions did not reveal any significant differences in reading times when the subject noun was singular. The learners did show sensitivity at the first spillover region in the relative clause conditions when the subject noun was plural ( $t(19)=-2.27, p<.05$ , by participants;  $t(19)=-2.54, p<.05$ , by items). These results confirm the results found in the ANOVAs, indicating that learners show decreased sensitivity to errors within relative clause intervener conditions compared to the prepositional phrase intervener conditions.

### 6.3.2. Native Speaker Results

As expected, results of the pairwise comparisons revealed that unstressed native speakers were sensitive to agreement errors in all conditions. Significant (or marginal, in the items analysis) differences in reading times emerged in the first spillover region of the singular subject-prepositional phrase intervener conditions ( $t(27)=-3.21, p<.01$ , by participants;  $t(23)=-2.03, p=.053$ , by items). Significant differences in reading times emerged within the plural subject-prepositional phrase conditions at Region 8 ( $t(27)=-3.35, p<.01$ , by participants;  $t(23)=-5.91, p<.001$ , by items) and Region 9 ( $t(27)=-3.5, p<.01$ , by participants;  $t(23)=-3.29, p<.01$ , by items); the effect was marginal in Region 10 ( $t(27)=-1.99, p=.06$ , by participants;  $t(23)=-1.86, p=.08$ , by items). In the singular subject-relative clause intervener conditions, reading time

differences emerged in the first spillover region, Region 9 ( $t(27)=-2.14, p<.05$ , by participants;  $t(23)=-1.98, p=.06$ , by items). Significant differences in reading times for the plural subject-relative clause intervener conditions emerged at the first spillover region ( $t(27)=-3.46, p<.01$ , by participants;  $t(23)=-3.27, p<.01$ , by items) and the second spillover region ( $t(27)=-2.29, p<.05$ , by participants).

### 6.3.3 Stressed Native Speaker Results

Results for this group indicated that stressed native speakers were sensitive to agreement violations in the singular subject-prepositional phrase, made evident by significant differences in reading times at the first spillover region ( $t(27)=-2.73, p<.05$ , by participants;  $t(23)=-2.72, p<.05$ , by items) and the second spillover region ( $t(27)=-2.23, p<.05$ , by participants;  $t(23)=-3.08, p<.01$ , by items).

A marginal difference in reading times was also found within the same conditions at the critical region ( $t(27)=0.26, p=.08$ , by participants). In the plural subject-prepositional phrase conditions, on the other hand, only a marginal difference in reading times emerged at the first spillover region ( $t(27)=-1.79, p=.08$ , by participants;  $t(23)=-1.88, p=.07$ ). These results differ from the results of the ANOVAs, since the pairwise comparisons indicate that the stressed native speakers showed decreased sensitivity to errors in the plural condition compared to the singular condition when a prepositional phrase intervened, and the ANOVAs revealed no interactions between number and grammaticality for the group.

Within the relative clause intervener conditions, the stressed native speakers showed no significant differences in reading times when the subject was singular. When the subject was plural, significant differences in reading times emerged in the items analysis only at the first

spillover region ( $t(23)=-2.19, p<.05$ ), indicating that participants were sensitive to errors when in the relative clause intervener condition when the subject was plural but not when it was singular. These results also differ from the ANOVA for the relative clause, which revealed an effect of grammaticality and no interactions between number and grammaticality. Instead, these paired comparison results are more similar to those for the L2 learner group, which also indicated that learners treated the two intervener conditions differently and that the plural subject number facilitated agreement establishment in the more complex relative clause conditions.

#### **6.4 Summary of Results**

In summary, results of 2x2 ANOVAs indicated that the L2 learners showed decreased sensitivity to agreement errors when a relative clause intervened between subject and noun, compared to the prepositional phrase intervener conditions. Results from the paired comparisons confirmed these findings, since learners showed significant differences in reading times for both singular and plural subject in the prepositional phrase intervener conditions, but were only sensitive to errors in the relative clause conditions when the subject was plural.

The 2x2 ANOVAs conducted for unstressed native speaker reading times revealed that native speakers were ultimately sensitive to agreement errors regardless of intervener. The ANOVA for prepositional phrase intervener and post-hoc t-tests indicated that native speakers became sensitive to agreement errors one region earlier (the critical region) when the subject was plural instead of singular. Again, however, the native speakers were ultimately sensitive to agreement errors regardless of subject number. This overall sensitivity to agreement errors was confirmed in the pairwise comparisons conducted for native speakers' reading times.

Similar to the unstressed native speaker results, the 2x2 ANOVAs for stressed native speakers indicated that these participants were ultimately sensitive to agreement violations regardless of intervener. Planned pairwise comparisons of the stressed native speakers' reading times, however, showed differences for the sentences with the prepositional phrase intervener and sentences with the relative clause intervener. T-tests comparing grammatical and ungrammatical sentence reading times in the singular subject-prepositional phrase intervener conditions (*The tourist in warm southern Mexico often fishes/\*fish in the ocean*) indicated that participants were sensitive to agreement violations, but comparisons of grammatical and ungrammatical sentences in the plural subject-prepositional phrase intervener conditions (*The tourists in warm southern Mexico often fish/\*fishes in the ocean*) revealed only a marginal difference between reading times at one spillover region. These results are not predicted by the plural markedness proposal (Wagers et al., 2009), and are not consistent with results from the L2 learner group or unstressed native speaker control group. Comparisons in the singular subject-relative clause conditions (*The tourist who hunted in Mexico often fishes/\*fish in the ocean*) failed to reveal significant differences in reading times, similar to the L2 learners. Comparisons in the plural subject-relative clause conditions (*The tourists who hunted in Mexico often fish/\*fishes in the ocean*), however, did indicate that stressed native speakers were sensitive to agreement errors. The pattern of results in the relative clause pairwise comparisons for stressed native speakers is consistent with that in the L2 learner group, and suggests that native speakers under a processing burden also perform differently across intervener conditions. These results for all three participant groups and their implications towards effects of intervener, plural markedness, and processing limitation proposals will be further discussed in the following section.



## 7. Discussion

In order to further investigate the L2 processing of agreement, the current study examined the effects of two linguistic factors on agreement establishment: structural distance and plural markedness. Structural distance was manipulated between a subject and verb by inserting either a prepositional phrase (*from warm southern Mexico*) or a relative clause (*who hunted in Mexico*) between the agreeing elements. Plural markedness was examined by comparing L2 learner sensitivity to agreement errors when the subject of a sentence was singular to sensitivity when the subject was plural. So that potential effects of proficiency or L1 transfer would not influence results, advanced learners of Spanish made up the L2 learner group. As was discussed above, the Shallow Structure Hypothesis does not predict that the L1 is a factor of morphosyntactic variability (Clahsen & Felser, 2006). However, since other studies (Coughlin & Tremblay, in press; Gabriele et al., in press) have found evidence of L1 transfer effects, we chose to examine L1 Spanish/L2 English learners because both languages require subjects and verbs to agree in number.

The current study also aimed to investigate the effects of a concurrent processing load on native speaker processing of long-distance subject-verb agreement by placing an additional concurrent digit load on native speakers as they completed an online self-paced reading task examining sensitivity to agreement violations.

In this section, I will discuss the results found in light of the research questions posited regarding structural distance effects, plural markedness effects, and processing limitation proposals.

## 7.1 Effects of Structural Distance

The Shallow Structure Hypothesis proposes that even advanced L2 learners are incapable of processing non-local agreement in a native-like way. This hypothesis also predicts that L2 learners are also insensitive to hierarchical structure (Clahsen & Felser, 2006). However, recent studies have found that L2 learners are capable of establishing agreement when the agreeing elements are separated linearly (Coughlin & Tremblay, in press) and that learners and native speakers are similarly affected by structural distance (Gabriele et al., in press). The current study, therefore, asks the following research question:

RQ1: Will L2 learners of English be less sensitive to subject-verb agreement errors when agreement is established across a structurally more complex intervening phrase?

This question was investigated by separating subjects and verbs by one of two types of interveners: A prepositional phrase (*in warm southern Mexico*) or a structurally more complex relative clause (*who hunted in Mexico*). Based on the predictions of the Shallow Structure Hypothesis, one pattern of results would provide evidence in support of the hypothesis: Learners will not be sensitive to agreement errors in either distance condition, indicating an inability to establish agreement across phrases as well as a lack of sensitivity to hierarchical structure. In this section, the ANOVA and pairwise comparison results of the L2 learner and unstressed native speaker control group will be discussed in light of the predictions of the Shallow Structure Hypothesis.

The native speaker control group's ANOVAs revealed an effect of grammaticality one region earlier (the critical region; the verb) in the prepositional phrase intervener condition than the relative clause intervener condition (first spillover region). Despite this slight time advantage

for the structurally less complex prepositional phrase intervener condition, native speakers were ultimately sensitive to agreement violations regardless of intervener type.

Unlike the native speaker group, results of the 2x2 ANOVAs for learners indicated that learners were sensitive to errors only when the intervener was a prepositional phrase. Exploratory pairwise comparisons of ungrammatical and grammatical sentences confirm these results: learners showed significant differences in reading times within the singular subject-prepositional phrase intervener condition (*The tourist from warm southern Mexico often fish/es in the ocean*) and the plural subject-prepositional phrase intervener condition (*The tourists from warm southern Mexico often fish/es in the ocean*). Within the relative clause intervener condition, learners showed sensitivity to errors only when the subject was plural (*The tourists who hunted in Mexico often fish/es in the ocean*).

Although the L2 learner results show some patterns of sensitivity that differ from native speaker patterns, the fact that learners showed sensitivity to errors in the prepositional phrase intervener conditions presents two problems for the Shallow Structure Hypothesis. First, though the prepositional phrase intervener was less syntactically complex than the relative clause intervener, it still required learners to establish agreement across phrases. Since the Shallow Structure Hypothesis predicts that L2 learners will be insensitive to any non-local agreement errors (such as contexts in which the agreeing elements are located in different phrases), these results cannot be presented as support for the hypothesis. Secondly, according to the proposal that learners are insensitive to hierarchical structure, the L2 learners in the current study should have shown equal insensitivity to errors in both the prepositional phrase and relative clause intervener conditions as we controlled for linear distance. The L2 learners, however, showed native-like sensitivity to errors in the prepositional phrase intervener condition while they

showed a decrease in sensitivity to errors in the structurally more complex relative clause intervener condition. These results do not support the predictions of the Shallow Structure Hypothesis, since they imply that learners are, in fact, sensitive to hierarchical structure. Instead, the L2 learner results of the current study are more similar to those found in Gabriele et al. (in press). The researchers found that, for both native speakers and advanced learners, sensitivity to errors decreased when distance was established across a phrase rather than within-phrase. This decrease in sensitivity was found only in the intermediate and advanced learner groups, suggesting that hierarchical structure does not influence L2 processing at lower levels of proficiency.

An alternative explanation for the learners' nonnative-like sensitivity to agreement errors lies in processing limitation proposals. This possibility will be examined below in the discussion of processing limitation effects on the stressed native speaker group (Section 7.3).

## **7.2 Plural Markedness Effects**

In a study conducted on native speakers of English, Wagers et al. (2009) investigates whether the marked plural number feature facilitates agreement establishment even as linear distance increases between a demonstrative and noun. The results of a speed-accuracy trade-off task revealed that participants were more quickly able to respond to agreement violations when the demonstrative was plural, regardless of distance. The researchers, therefore, conclude that the plural advantage of the markedness effect is able to withstand linear distance. Based on these findings, the current study posited the research question below:

RQ2: Will the effects of plural markedness facilitate successful subject-verb agreement resolution for L2 learners as structural distance between the two agreeing elements is increased?

This question was investigated by systematically manipulating the number of the subject noun (singular or plural) across two intervener types: prepositional phrases and structurally more complex relative clauses. In order to fully support the evidence found in Wagers et al. (2009), results of the current study needed to indicate that the plural subject noun facilitated agreement error detection particularly as the distance between agreeing elements increased (i.e., in the relative clause intervener conditions). Results of the native speaker control group and L2 learner group and their application to plural markedness effects will be discussed below.

Results of the 2x2 ANOVAs for unstressed native speakers revealed an interaction between number and grammaticality only within the prepositional phrase intervener condition, not the relative clause intervener condition. Post-hoc t-tests comparing grammatical and ungrammatical versions of sentences revealed that, when the subject noun was singular and a prepositional phrase intervened (i.e., *The tourist from warm southern Mexico often fish/es in the ocean*), no significant differences in reading times emerged at the critical region. However, a significant difference in reading times was found at the critical region in the plural subject-prepositional phrase intervener condition (*The tourists from warm southern Mexico often fish/es in the ocean*). These t-test results suggest that plural subject noun did give the native speakers a slight time advantage, since they displayed sensitivity to agreement violations within the prepositional phrase intervener condition earlier (at the critical region, Region 8) in the plural subject condition compared to the singular subject condition. Though these results provide support for the possible presence of a plural markedness effect, the fact that the native speakers ultimately showed sensitivity to agreement errors regardless of subject number suggests that the

effect is weak. Furthermore, no advantage for plural subject emerged in the relative clause condition, as native speakers showed similar sensitivity to errors in both number manipulations.

In the ANOVA results for L2 learners, only a marginal interaction between number and grammaticality emerged. This result does not provide strong evidence for the presence of a plural markedness effect, since plural number seemed to provide no significant facilitation of agreement resolution for learners in the prepositional phrase or relative clause intervener conditions. The planned pairwise comparisons confirmed these weak results. T-tests comparing ungrammatical and grammatical sentences for the singular subject-prepositional phrase intervener condition (*The tourist in warm southern Mexico often fish/es in the ocean*) and for the plural subject-prepositional phrase intervener condition (*The tourists in warm southern Mexico often fish/es in the ocean*) both revealed significant reading time differences, indicating that the learners were sensitive to errors in both of those conditions regardless of subject number. However, t-tests for within the relative clause intervener condition revealed significant differences in reading times only when the subject was plural (*The tourists who hunted in Mexico often fish/es in the ocean*); no significant differences emerged when the subject was singular (*The tourist who hunted in Mexico often fish/es in the ocean*). These results do suggest that a plural markedness effect is present for L2 learners as structural distance increases, even though it was not strong enough to produce a significant interaction between number and grammaticality in the repeated measures analyses of variance.

Taking the results of the control group and L2 learner group into consideration, it seems to be the case that a weak plural markedness effect may be present for both groups. One limitation of the current study that should be taken into consideration at this point is the size of the participant groups. Both groups are relatively small, and the absence of strong effects for

plural subjects could potentially be caused by a lack of statistical power. This alternative explanation is supported by the fact that mean reading times at and after the critical region for both participant groups show the same numerical trend: differences between ungrammatical and grammatical reading times are consistently larger for plural subject conditions than they are for singular subject conditions.

### **7.3 Processing Limitations**

Several recent studies (MacDonald, 2006; Hopp, 2010; Lopez-Prego, 2012) have placed native speakers under a concurrent processing burden while testing for sensitivity to morphosyntactic errors. The researchers then compared results of the native speakers to results of L2 learners, and found that the processing burden causes native speakers to show similar patterns of sensitivity seen in the learner results. These patterns have led to a proposal that morphological variability in L2 learners may be a result of processing limitations in the second language; however, this method of taxing native speakers to elicit learner-like behavior is relatively new in the field of second language acquisition. The current study, therefore, aimed to use this testing paradigm to investigate a construction not yet examined: long-distance subject-verb agreement. The following research question was posed:

RQ3: Will native speakers placed under a processing burden show variability similar to the L2 learners in their processing of agreement?

In order to investigate this question, a group of native English speakers completed the same reading comprehension task in which sensitivity to subject-verb agreement errors was investigated while structural distance and subject number were manipulated. This group, as

opposed to the learner or native speaker control groups, completed the task while holding a concurrent digit load in memory. Finding that the stressed native speaker group produced similar patterns of sensitivity compared to the learner group would corroborate evidence from the studies discussed above, suggesting that L2 learner variability is tied to processing limitations in the second language and not to a lack of grammatical knowledge. In this section, the results of the stressed native speaker group will be compared to those of the two other participant groups.

Results of the stressed native speaker 2x2 ANOVAs reveal an effect of grammaticality in both the prepositional phrase intervener and relative clause intervener conditions, indicating that stressed native speakers were sensitive to agreement errors regardless of structural distance. Here, the stressed native speakers and native speaker control group show similar patterns of sensitivity, while learners show no sensitivity to errors in the relative clause intervener condition. The planned pairwise comparisons, however, present different results. Recall that in pairwise comparisons of L2 learner reading times for grammatical and ungrammatical sentences, significant differences emerge for both prepositional phrase intervener conditions (singular subject-prepositional phrase intervener, plural subject-prepositional phrase intervener). Comparisons of learner reading times within the relative clause intervener conditions, however, reveal significant differences only when the subject is plural (*The tourists who hunted in Mexico often fish/\*fishes in the ocean.*). While unstressed native speaker results indicated that these participants showed sensitivity to agreement violations in all conditions, t-tests conducted for the stressed native speaker group revealed results almost identical to the L2 learners. The results of the ANOVAs indicate that stressed native speakers were sensitive to agreement errors regardless of intervener type, but the paired comparisons suggest that—like L2 learners—native speakers under a processing burden perform differently in detecting agreement errors based on the



structural distance separating agreeing elements. Paired comparisons conducted on stressed native speaker reading times also revealed a highly unexpected result that was not evident in either of the two other participant groups: within the singular subject-prepositional phrase intervener conditions, no significant slowdown for ungrammatical critical regions was found. Future studies may consider testing larger participant groups in order to see if this weak unexpected result is replicated.

MacDonald (2006) argues that L2 learners' lack of sensitivity to grammatical violations within syntactically complex structures (such as long-distance agreement) in real-time processing may not be indicative of a deficit in grammatical knowledge, as the Shallow Structure Hypothesis suggests. Instead, if syntactic processing capabilities and general processing abilities are both used in the processing of sentences, more complex syntactic structures may be more easily affected by a processing stress (Miyake, Carpenter, & Just, 1996). Recall that, in the current study, a prepositional phrase intervener separates the subject and verb phrase by only two syntactic nodes, whereas a relative clause intervener separates the subject and verb phrase by five syntactic nodes. Therefore, the relative clause intervener condition in the current study is syntactically more complex than the prepositional phrase intervener condition. The fact that both the L2 learners and the stressed native speakers in the current study failed to show native-like sensitivity to agreement errors in the relative clause intervener condition provide tentative support for a processing limitation proposal such as that laid out in MacDonald (2006).

Furthermore, although no interactions between number and grammaticality were found in the ANOVA results for stressed native speakers or L2 learners, pairwise comparisons for both groups revealed sensitivity to agreement errors in the relative clause condition only when the subject was plural. This similarity suggests that, like L2 learners, stressed native speakers may

benefit from potential plural markedness effects as structural distance between subject and verb increases.

In summary, results of the ANOVAs conducted for L2 learner and stressed native speakers do not reveal substantial similarities between the patterns of variability for each group. However, exploratory pairwise comparisons suggest that both groups show a decrease of sensitivity to grammatical errors when a relative clause intervenes. The comparisons also suggest a weak plural markedness effect may facilitate long-distance agreement, since both groups showed sensitivity to errors in the relative clause condition only when the subject was plural.

## **8. Conclusion**

The current study aimed to investigate the effects of two linguistic factors on the establishment of subject-verb agreement in the L2: structural distance and plural markedness. The study also examined processing limitation proposals by placing native speakers under a concurrent digit load and comparing their performance to that of L2 learners. Results for L2 learners suggest that structural distance does affect morphological variability in learners, since learners were less sensitive to errors in the structurally more complex relative clause intervener condition. The fact that learners are sensitive to hierarchical structure and the fact that stressed native speakers show similar, albeit weaker, patterns of variability in the pairwise comparisons provide some support for a processing limitation proposal of variability instead of the Shallow Structure Hypothesis.

Based on the L2 learner and unstressed native speaker results in the current study, it is difficult to argue for a strong plural markedness effect such as that seen in Wagers et al (2009).

Although a weak advantage may have been present for the plural subject condition, future studies should further examine this possible effect by examining other agreement constructions and by constructing larger participant groups. The fact that pairwise comparisons revealed a lack of sensitivity in the unmarked singular condition when a relative clause intervened for both L2 learner and stressed native speaker groups adds further support to a processing limitation proposal of morphological variability.

One direction of research that has become relevant in L2 processing literature is the consideration of individual differences (such as working memory capacity) as factors influencing L2 variability (Coughlin & Tremblay, in press). Although the current study essentially taxes native speaker working memory by placing participants under a concurrent digit load, it does not correlate the participants' abilities to establish agreement with individual working memory capacity. Future research should take individual participants' cognitive skills into consideration when comparing learners and stressed native speakers, in order to determine whether successful agreement establishment is aided for both groups by non-linguistic factors.

## References

- Birdsong, D., & Mollis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, *44*, 235-249
- Blackwell, A., & Bates, E. (1995). Inducing agrammatic profiles in normals: Evidence for the selective vulnerability of morphology under cognitive resource limitation. *Journal of Cognitive Neuroscience*, *7*, 228-257.
- Bock, J.K., & Cutting, J.C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99-127.
- Bock, J.K., & Miller, C.A. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45-93.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3-42.
- Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*, 21-43.
- Coughlin, K., & Tremblay, A. (in press). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics*, 1-32.
- Dick, F., Bates, E., Wulfeck, B., Utman, J.A., Dronkers, N., & Gernsbacher, M.A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review*, *108*, 759-788.
- Eberhard, K. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language*, *36*, 147-164.
- Eberhard, K.M, Cutting, J.C., & Bock, (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*, 531-559.
- Gabriele, A., Bond, K., Canales, A., Coughlan, S.J. Examining sensitivity to agreement anomalies during online processing. Poster presented at the Second Language Research Forum 2010. University of Maryland. October 14-17, 2010.
- Gabriele, A. Fiorentino, R., and Alemán-Bañón, J. (in press-2013). Examining second language development using event-related potentials: a cross-sectional study on the processing of gender and number agreement. *Linguistic Approaches to Bilingualism*.
- Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*, 901-931.

- 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.
- Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning*, 61, 940-967.
- 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
- Keating, G.D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning*, 59, 503-535.
- Keating, G.D. (2010). The effects of linear distance and working memory on the processing of gender agreement in Spanish. In B. VanPatten & J. Jegerski (Eds.), *Research in Second Language Processing and Parsing* (pp. 112-134). Philadelphia, PA: John Benjamins Publishing Company.
- Lardiere, D. (1998). Case and tense in the 'fossilized' steady state. *Second Language Research*, 14, 1-26.
- Lopez-Prego, B. (2012). Examining the nature of variability in gender and number agreement in native and non-native Spanish. MA Thesis, University of Kansas.
- McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, 24, 469-486.
- 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.
- Pearlmutter, N.J., Garnsey, S.M., & Bock, J.K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427-456.
- Prevost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 103-133.

- Solomon, E., & Pearlmutter, N. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, *49*, 1-46.
- Tagliaferri, B. (2005). *Paradigm*. Perception Research Systems, Inc.  
www.perceptionresearchsystems.com.
- Wagers, M., Lau, E., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*, 206,237.
- Wagers, M., Lau, E., Stroud, C., McElree, B., & Phillips, C. (2009). Encoding syntactic predictions: evidence from the dynamics of agreement. In Proceedings from the 22<sup>nd</sup> CUNY Conference on Sentence Processing (pages). University of California, Davis, CA.
- Waters, G., Caplan, D., & Rochon, E. (1995). Processing capacity and sentence processing in patients with Alzheimer's disease. *Cognitive Neuropsychology*, *12*, 1-30.
- Waters, G., Caplan, D., & Yampolsky, S. (2003). On-line syntactic processing under concurrent memory load. *Psychological Bulletin and Review*, *10*, 88-95.
- White, L. (2003). Fossilization in steady state grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition*, *6*, 128-141.

## Appendix. Target Stimuli

1. The singer from hot northern Africa never practices before a performance.  
The singer from hot northern Africa never practice before a performance.  
The singers from hot northern Africa never practice before a performance.  
The singers from hot northern Africa never practices before a performance.  
The singer who asked about Africa never practices before a performance.  
The singer who asked about Africa never practice before a performance.  
The singers who asked about Africa never practice before a performance.  
The singers who asked about Africa never practices before a performance.
  
2. The author from cold eastern Canada sometimes hikes in the morning.  
The author from cold eastern Canada sometimes hike in the morning.  
The authors from cold eastern Canada sometimes hike in the morning.  
The authors from cold eastern Canada sometimes hikes in the morning.  
The author who camped in Canada sometimes hikes in the morning.  
The author who camped in Canada sometimes hike in the morning.  
The authors who camped in Canada sometimes hike in the morning.  
The authors who camped in Canada sometimes hikes in the morning.
  
3. The orphan from crowded southern Brazil sometimes begs at the market.  
The orphan from crowded southern Brazil sometimes beg at the market.  
The orphans from crowded southern Brazil sometimes beg at the market.  
The orphans from crowded southern Brazil sometimes begs at the market.  
The orphan who moved from Brazil sometimes begs at the market.  
The orphan who moved from Brazil sometimes beg at the market.  
The orphans who moved from Brazil sometimes beg at the market.  
The orphans who moved from Brazil sometimes begs at the market.
  
4. The prisoner in dangerous western Russia never shouts at the guards.  
The prisoner in dangerous western Russia never shout at the guards.  
The prisoners in dangerous western Russia never shout at the guards.  
The prisoners in dangerous western Russia never shouts at the guards.  
The prisoner who confessed in Russia never shouts at the guards.  
The prisoner who confessed in Russia never shout at the guards.  
The prisoners who confessed in Russia never shout at the guards.  
The prisoners who confessed in Russia never shouts at the guards.

5. The pilot from rural northern England sometimes smiles at the passengers.  
The pilot from rural northern England sometimes smile at the passengers.  
The pilots from rural northern England sometimes smile at the passengers.  
The pilots from rural northern England sometimes smiles at the passengers.  
The pilot who landed in England sometimes smiles at the passengers.  
The pilot who landed in England sometimes smile at the passengers.  
The pilots who landed in England sometimes smile at the passengers.  
The pilots who landed in England sometimes smiles at the passengers.
  
6. The secretary from sunny northern California always laughs in the office.  
The secretary from sunny northern California always laugh in the office.  
The secretaries from sunny northern California always laugh in the office.  
The secretaries from sunny northern California always laughs in the office.  
The secretary who complained about California always laughs in the office.  
The secretary who complained about California always laugh in the office.  
The secretaries who complained about California always laugh in the office.  
The secretaries who complained about California always laughs in the office.
  
7. The runner in humid eastern Florida always showers after a workout.  
The runner in humid eastern Florida always shower after a workout.  
The runners in humid eastern Florida always shower after a workout.  
The runners in humid eastern Florida always showers after a workout.  
The runner who raced in Florida always showers after a workout.  
The runner who raced in Florida always shower after a workout.  
The runners who raced in Florida always shower after a workout.  
The runners who raced in Florida always showers after a workout.
  
8. The athlete from charming southern Italy sometimes argues during a game.  
The athlete from charming southern Italy sometimes argue during a game.  
The athletes from charming southern Italy sometimes argue during a game.  
The athletes from charming southern Italy sometimes argues during a game.  
The athlete who competed in Italy sometimes argues during a game.  
The athlete who competed in Italy sometimes argue during a game.  
The athletes who competed in Italy sometimes argue during a game.  
The athletes who competed in Italy sometimes argues during a game.
  
9. The politician from dry western Libya often works in the evening.  
The politician from dry western Libya often work in the evening.  
The politicians from dry western Libya often work in the evening.  
The politicians from dry western Libya often works in the evening.



The politician who feared for Libya often works in the evening.  
The politician who feared for Libya often work in the evening.  
The politicians who feared for Libya often work in the evening.  
The politicians who feared for Libya often works in the evening.

10. The chef in glamorous eastern Hawaii always talks to the waiter.  
The chef in glamorous eastern Hawaii always talk to the waiter.  
The chefs in glamorous eastern Hawaii always talk to the waiter.  
The chefs in glamorous eastern Hawaii always talks to the waiter.  
The chef who cooked in Hawaii always talks to the waiter.  
The chef who cooked in Hawaii always talk to the waiter.  
The chefs who cooked in Hawaii always talk to the waiter.  
The chefs who cooked in Hawaii always talks to the waiter.
  
11. The firefighter in friendly southern Oklahoma sometimes volunteers at the shelter.  
The firefighter in friendly southern Oklahoma sometimes volunteer at the shelter.  
The firefighters in friendly southern Oklahoma sometimes volunteer at the shelter.  
The firefighters in friendly southern Oklahoma sometimes volunteers at the shelter.  
The firefighter who trained in Oklahoma sometimes volunteers at the shelter.  
The firefighter who trained in Oklahoma sometimes volunteer at the shelter.  
The firefighters who trained in Oklahoma sometimes volunteer at the shelter.  
The firefighters who trained in Oklahoma sometimes volunteers at the shelter.
  
12. The actor from windy western Peru always performs at the theater.  
The actor from windy western Peru always perform at the theater.  
The actors from windy western Peru always perform at the theater.  
The actors from windy western Peru always performs at the theater.  
The actor who entertained in Peru always performs at the theater.  
The actor who entertained in Peru always perform at the theater.  
The actors who entertained in Peru always perform at the theater.  
The actors who entertained in Peru always performs at the theater.
  
13. The dancer from warm northern Texas never travels to the beach.  
The dancer from warm northern Texas never travel to the beach.  
The dancers from warm northern Texas never travel to the beach.  
The dancers from warm northern Texas never travels to the beach.  
The dancer who called from Texas never travels to the beach.  
The dancer who called from Texas never travel to the beach.  
The dancers who called from Texas never travel to the beach.  
The dancers who called from Texas never travels to the beach.

14. The journalist from stormy eastern Uruguay often reports in the afternoon.  
The journalist from stormy eastern Uruguay often report in the afternoon.  
The journalists from stormy eastern Uruguay often report in the afternoon.  
The journalists from stormy eastern Uruguay often reports in the afternoon.  
The journalist who vacationed in Uruguay often reports in the afternoon.  
The journalist who vacationed in Uruguay often report in the afternoon.  
The journalists who vacationed in Uruguay often report in the afternoon.  
The journalists who vacationed in Uruguay often reports in the afternoon.
  
15. The cowboy from dusty southern Australia sometimes rests under a tree.  
The cowboy from dusty southern Australia sometimes rest under a tree.  
The cowboys from dusty southern Australia sometimes rest under a tree.  
The cowboys from dusty southern Australia sometimes rests under a tree.  
The cowboy who lived in Australia sometimes rests under a tree.  
The cowboy who lived in Australia sometimes rest under a tree.  
The cowboys who lived in Australia sometimes rest under a tree.  
The cowboys who lived in Australia sometimes rests under a tree.
  
16. The student from scenic western France always waves to the professors.  
The student from scenic western France always wave to the professors.  
The students from scenic western France always wave to the professors.  
The students from scenic western France always waves to the professors.  
The student who cheated in France always waves to the professors.  
The student who cheated in France always wave to the professors.  
The students who cheated in France always wave to the professors.  
The students who cheated in France always waves to the professors.
  
17. The teacher from modern northern Europe never jogs at the park.  
The teacher from modern northern Europe never jog at the park.  
The teachers from modern northern Europe never jog at the park.  
The teachers from modern northern Europe never jogs at the park.  
The teacher who stayed in Europe never jogs at the park.  
The teacher who stayed in Europe never jog at the park.  
The teachers who stayed in Europe never jog at the park.  
The teachers who stayed in Europe never jogs at the park.
  
18. The nurse in noisy eastern Greece often walks to the store.  
The nurse in noisy eastern Greece often walk to the store.  
The nurses in noisy eastern Greece often walk to the store.  
The nurses in noisy eastern Greece often walks to the store.

The nurse who relaxed in Greece often walks to the store.  
The nurse who relaxed in Greece often walk to the store.  
The nurses who relaxed in Greece often walk to the store.  
The nurses who relaxed in Greece often walks to the store.

19. The robber from violent western Philadelphia often smokes on the train.  
The robber from violent western Philadelphia often smoke on the train.  
The robbers from violent western Philadelphia often smoke on the train.  
The robbers from violent western Philadelphia often smokes on the train.  
The robber who returned to Philadelphia often smokes on the train.  
The robber who returned to Philadelphia often smoke on the train.  
The robbers who returned to Philadelphia often smoke on the train.  
The robbers who returned to Philadelphia often smokes on the train.
20. The coach from quiet southern Kansas always exercises at the gym.  
The coach from quiet southern Kansas always exercise at the gym.  
The coaches from quiet southern Kansas always exercise at the gym.  
The coaches from quiet southern Kansas always exercises at the gym.  
The coach who celebrated in Kansas always exercises at the gym.  
The coach who celebrated in Kansas always exercise at the gym.  
The coaches who celebrated in Kansas always exercise at the gym.  
The coaches who celebrated in Kansas always exercises at the gym.
21. The kid in beautiful northern Spain never plays on the playground.  
The kid in beautiful northern Spain never play on the playground.  
The kids in beautiful northern Spain never play on the playground.  
The kids in beautiful northern Spain never plays on the playground.  
The kid who learned about Spain never plays on the playground.  
The kid who learned about Spain never play on the playground.  
The kids who learned about Spain never play on the playground.  
The kids who learned about Spain never plays on the playground.
22. The teenager in urban western India never shops at the mall.  
The teenager in urban western India never shop at the mall.  
The teenagers in urban western India never shop at the mall.  
The teenagers in urban western India never shops at the mall.  
The teenager who talked about India never shops at the mall.  
The teenager who talked about India never shop at the mall.  
The teenagers who talked about India never shop at the mall.  
The teenagers who talked about India never shops at the mall.

23. The tourist in warm southern Mexico often fishes in the ocean.  
The tourist in warm southern Mexico often fish in the ocean.  
The tourists in warm southern Mexico often fish in the ocean.  
The tourists in warm southern Mexico often fishes in the ocean.  
The tourist who hunted in Mexico often fishes in the ocean.  
The tourist who hunted in Mexico often fish in the ocean.  
The tourists who hunted in Mexico often fish in the ocean.  
The tourists who hunted in Mexico often fishes in the ocean.

24. The artist from busy eastern China often paints at the lake.  
The artist from busy eastern China often paint at the lake.  
The artists from busy eastern China often paint at the lake.  
The artists from busy eastern China often paints at the lake.  
The artist who backpacked in China often paints at the lake.  
The artist who backpacked in China often paint at the lake.  
The artists who backpacked in China often paint at the lake.  
The artists who backpacked in China often paints at the lake.