The Pennsylvania State University

The Graduate School

College of the Liberal Arts

HOW SECOND LANGUAGE LEARNERS PROCESS ARGUMENT STRUCTURE: THE EFFECTS OF FIRST LANGUAGE AND INDIVIDUAL DIFFERENCES

A Dissertation in

Spanish and Language Science

by

Aroline E. Seibert Hanson

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2012

The dissertation of Aroline E. Seibert Hanson was reviewed and approved* by the following:

Chip Gerfen
Head of the Department of Spanish, Italian, and Portuguese
Associate Professor of Linguistics and Spanish Linguistics
Dissertation Adviser
Chair of Committee

Paola E. Dussias
Associate Professor of Spanish, Linguistics, and Psychology

Carol A. Miller
Associate Professor of Communication Sciences and Disorders and Linguistics

Karen Miller
Assistant Professor of Spanish Linguistics

Matthew T. Carlson
Assistant Professor of Psychology
Special Member

*Signatures are on file in the Graduate School.

ABSTRACT

Research suggests that adult second language learners have difficulty processing certain argument structures in their L2. In particular, it has been shown that L1 English learners of Spanish are not at first successful in processing preverbal clitic structures in Spanish such as in (a), (e.g. Liceras, 1985; VanPatten, 1984).

(a)	*Lo*	*besa*	*la niña*.
	[3RD SINGMASCACC	kiss-3RD SINGPRE	the girl]
	The girl kisses him.		

These structures are relatively common in Spanish, and being able to process them properly is essential to becoming proficient in the language. The present dissertation aimed to determine the factors involved in how L2 learners process Spanish structures with preverbal object clitics and postverbal subjects. I considered three areas that possibly contribute: (1) the characteristics of the L1 (i.e. transfer), (2) the proficiency of the learners (previous studies typically tested only low proficiency learners), and (3) individual differences in cognitive abilities such as working memory capacity (WMC) and inhibitory control/executive attention.

Testing (1), I compared the performance of L1 English speakers to L1 speakers of a language that is typologically like Spanish, Romanian, on two sentence processing tasks (one listening and one reading) in Spanish. Romanian has a similar system of pronominal clitics to Spanish, allowing for structures with preverbal clitics and postverbal subjects, as in (b).

(b)	*O*	*caută*	*băiatul*
	[3RD SINGMASCACC	kiss-3RD SINGPRE	girl-the]
	The girl kisses him.		

Testing (2), I included participants of a wide range of L2 proficiency levels, which I assessed using an independent measure of proficiency (adapted from the DELE exam). Testing (3), each participant performed the Letter-Number Sequencing test of working memory

(Wechsler, 1997) and the Flanker Task, which assesses executive attentional abilities (Luk, 2008). In all, 71 L1 Romanian and 65 L1 English L2 Spanish learners and 36 L1 Spanish monolingual controls participated in the two sentence processing experiments.

With the use of logistic regression with mixed-effects models, I was able to include both discrete (e.g. first language) and continuous (e.g. proficiency level) variables, include all trials and participants with missing trials, and thus, provide a more accurate picture than other more traditional methods (see Jaeger, 2008 for discussion) of the relative contributions of these factors on learners' L2 processing of this challenging structure. Specifically, I found main effects on accuracy for proficiency and condition in both tasks, indicating the strength of the role of these two factors, the importance of testing different proficiency levels, and the difficulty of the clitic structures for all L2 learners. Additionally, there was a main effect for L1 in the listening task, but not in the reading task, and a main effect for working memory in the reading task, but not in the listening task. These results suggest that the two modalities place differing demands on the participant, and that neither L1 nor working memory play a strong determining role in successful processing of the present structure. In addition, there were significant interactions between proficiency and condition in both tasks, such that the more proficient L2 learners were more accurate on the OVS condition, and significant interactions between working memory and condition, such that the L2 learners with higher WMC were less accurate on the OVS condition, a vexing result that requires further research. Most provocative was the result that there was no interaction between L1 and condition, indicating that although the L1 Romanians were more accurate than the L1 English participants on the sentence processing tasks on the whole, this improved accuracy was not specific to clitic structures. Based on these results, transfer appears to play only a small role in L2 processing, and when it occurs, may not be structure-specific.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Researching and writing this dissertation has most certainly been a marathon and not a sprint. There have been many uphills and downhills. There have been many water stops and cheering sections that have kept me moving, and potholes and timing mats to navigate. When I began the race, I had a time and performance goal set in my mind. In the beginning, as I was still warming up, there were times that I was forced to sprint to the next mile marker, trying to achieve a particular split that was not well-planned. Even though I had friends and family handing me water and cheering me on, I was tiring out too quickly. At about half way, I realized that that pace was not for me. I had to switch coaches, ultimately finding one that has provided me with steady encouragement and knowledgeable support. I am extremely thankful for that. It has helped me near the finish line in not only a timely manner, but also with the gained experience that I needed to be a strong and successful lifelong runner.

I would like to thank in particular my parents for their support, love, and encouragement through the years. When things looked grim, I could always count on you to cheer me up and cheer me on. Thank you for not pressuring me, for believing in me. Thank you to Grandma as well, for listening patiently to my stories of how my studies were progressing, and offering sage advice. You were a wonderful friend. I will miss you very much. Also, I would like to thank my oldest and dearest friend, Lauren, for commiserating with me on writing a thesis, and for being caring and fun when I needed it. I would like to thank my newest and very dear friend, Melisa, for being supportive and sharing her knowledge and time with me as we both have tried to make our way through the trials and triumphs of graduate school, and for sharing common interests with me both academic and athletic. I thank my advisor, Chip Gerfen, for helping me rescue my dissertation and really make it flourish. I appreciate all your time and dedication. I thank Matt Carlson for patiently walking me through how to use R and how to build mixed models. Your patience is unequaled. I also thank Nuria Sagarra, Jacqueline Gauthier, and Caroline Hauser, who helped gather some of the data, Mila Crespo, for welcoming me and allowing me to collect data at the Instituto Cervantes, and April Jacobs for her help recording the auditory stimuli. I thank all my professors for teaching me what they know and love, and opening my eyes to new worlds. I am forever changed. Most important, I thank my husband, Eric. I would not be here if it were not for you. You encouraged me to seek my dream even when it meant moving far away, and then you made sacrifices to join me here. You provide me with daily reinforcement and unfaltering love. Thank you.

CHAPTER 1: Introduction and Literature Review

1.0 Introduction

Adults who wish to learn a second language face a daunting task. Regardless of the method, adult second language learning requires the acquisition and integration of an entirely new system of phonology, morphology, syntax, and an entirely new lexicon into an already developed system of language. Second language learners (L2 learners) must learn the nuances of semantics and pragmatics as well in order to successfully communicate in the new language. This process is so challenging that the majority of late L2 learners never fully accomplish the feat.

There has been great debate about why the challenges of learning a second language appear to be greater than the challenge of learning one's first language, something that all normally developing children achieve with a high level of success. It has been stated that it is the mere presence of the first language that interferes with the acquisition of the second; once the first language has been established, one will never be able to fully acquire a second language. Certainly there are limits to human cognitive capacity. The acquisition of a second language pushes our minds to these limits by providing multiple and often, conflicting cues to those we have already learned from our first language (see Ellis, 2006(b)). Not only has the research shown that we become attuned to our particular language in many ways, but it has also shown that this begins to happen very early on in human development (e.g. Werker and Tees, 1984). On the other hand, the fact that all late second language learners already have an understanding of the world and language, albeit through the lens of their first language, has been used as an argument in favor of our ability to acquire a second language. Furthermore, it has been reasoned that the first language has more specific influence, possibly helpful and possibly harmful, on certain areas of second language development.

While there may be some overlap between one's first and one's second language, such as the presence of cognate words with both phonological and semantic overlap or the same basic word order, the challenge of acquisition is still great. Often, the apparent similarities between two languages are not perfectly matched, such as when two words have the same form but different meaning in each language, which leads to possible misunderstanding and further confusion on the part of the learner. When there are definite similarities between two languages, however, such as when both languages have the same argument structure, it is unclear whether this truly aids acquisition. This dissertation aims to illuminate what factors affect second language learners' ability to successfully acquire and process a second language (L2) in regards to the processing of argument structure. I address the issues of whether a learner's first language (L1) plays a determining role, and what role other individual differences play in L2 learners' processing of argument structure.

There have been various models put forth to explain how L2 learners process argument structure in their L2. First, VanPatten's Input Processing Theory argues that second language learners at first may be processing their L2 based on universal parsing principles. In contrast, MacWhinney and colleagues (Bates and MacWhinney, 1979, 1982; MacWhinney, 1987, 1989, 1992) in their Competition Model have stated that learners transfer their processing strategies from their L1 to their L2. However, neither of these theories encompasses the entire realm of issues involved in second language acquisition. Ellis (2006a) has proposed an Associative-Cognitive Model that provides a more comprehensive language-experience based account for second language acquisition. This dissertation adopts the Associative-Cognitive framework with the purpose of delineating to what extent the L1 and other individual difference factors determine how learners are processing a particular structure.

In order to test the role of the L1 in L2 processing of argument structure, I sought a structure that was present in Spanish, but not in English. A good structure to test the strength of the L1's influence on L2 learners' ability to process L2 argument structure is the preverbal direct object clitic, licit in Spanish and in Romanian, but not in English. In the literature, this structure has been overwhelmingly shown to be challenging for L1 English learners of Spanish to process (e.g. Liceras, 1985; VanPatten, 1984).

Anticipating the background discussion below, if second language learners do transfer structures from their L1, then I hypothesize that L1 Romanian learners of Spanish will be more successful (i.e. accurate) in processing preverbal clitics than L1 English learners of Spanish due to their increased language experience with clitics in their L1. This follows the predictions of the Associative-Cognitive theory in that the increased exposure that the L1 Romanians have to a particular structure (preverbal clitics) will aid in their acquisition of that structure in their L2, while the L1 English speakers will require more experience over time with the structure to overcome L1 interference and entrenchment biases. Based on previous findings from Havik, Roberts, van Hout, Schreuder, and Haverkort (2009), among others, I also hypothesize that factors such as proficiency level and working memory abilities (capacity and executive attention) will further modulate the learners' ability to process these structures. In order to test these hypotheses, in the present chapter, I present the relevant background literature on the current theories of how second language learners process argument structure, including the Associative-Cognitive Model, and recent studies that have tested these theories. I also outline the major research on the role of individual difference factors such as working memory and the executive attention component of working memory, identifying the gaps in this work that still remain in the field of second language acquisition and processing. In chapter 2, I describe my method and procedure for the two sentence processing experiments that I conducted to test the aforementioned hypotheses and to possibly fill the gaps in the research, and provide some background on clitic structures in both Spanish and Romanian. I also provide the descriptive statistics for the assessment and individual difference measures. In chapter 3, I present and discuss the results from the listening sentence processing experiment. In chapter 4, I present and discuss the results from the reading sentence processing experiment. Finally, in chapter 5, I compare and draw conclusions based on the results, relating them back to the literature, and disclosing possible limitations that will be addressed in future work.

1.1 Background

The goal of this chapter is to provide an in-depth literature review of the theoretical underpinnings of this dissertation. I begin with a background on the research leading up to and including VanPatten's first noun strategy from Input Processing, which is a universal processing theory, and review studies that have tested it and their relevant results (section 1.1). This is followed by a review of studies on transfer that have pertinent results, in which I focus on work done in the Competition Model framework (section 1.2). Then I present an overview of the Associate-Cognitive framework (section 1.3), the framework adopted for the present dissertation, which offers a compromise between the universal and transfer accounts that is efficient and comprehensive. In section 1.4, I discuss studies that have found no effects of transfer. In the final two sections, I provide background on the cognitive individual differences of working memory and inhibitory control/executive attention and how they have been examined in the literature in relation to L2 argument structure processing (section 1.5). I conclude with a summary of the literature presented in the chapter, highlighting some of the open questions of past research and addressing what the current dissertation attempts to resolve (section 1.6).

1.2 Universal Processing Strategies

In this section, first I review the work leading up to VanPatten's (1984) seminal study. VanPatten based his experiments on prior research by Bever (1970) and Slobin and Bever (1982), who wrote about what they considered to be universal cognitive strategies that had been found in monolingual children, and LoCoco (1982, 1987), who investigated differing word order strategies in adult L2 learners. Next, I describe VanPatten's work that led to the development of Input Processing and his coining of the concept of the *First Noun Principle*. It is important to keep in mind that the research reviewed in this section aimed to identify strategies that are universal to all language users in the processing of argument structure, which is one of the ways that I consider that may be used to account for L2 learners' processing outcomes in the data found in the present dissertation.

1.2.1 Studies Prior to VanPatten (1984)

One of the first studies that asked the question about whether (first) language learners use universal strategies to process argument structure was Bever (1970). His study dealt with children's ability to comprehend passive vs. active structures in their native language, and suggested that in the absence of semantic constraints passive constructions are more difficult universally to comprehend than active constructions. Applying his *Strategy D*, as stated in (1), Bever posited that the presumed explanation for the perceptual difficulty of the passive is that a passive construction reverses the assumptions of *Strategy D*.

(1) *Any Noun-Verb-Noun (NVN) sequence within a potential internal unit in the surface structure corresponds to 'actor-action-object'.* (Bever, 1970, p. 298)

In addition, he hypothesized that the increased length of passive structures over the active structures may contribute to how complex they are to process. In his experiment, as well as in others, Bever argued that the errors children made were in assuming that the first noun in an NVN sequence was the grammatical subject in the external structure as well as the actor in the internal structure. He concluded that children understood surface NVN sequences as the constituent order of *actor-action-object* (Bever, 1970, p. 299). He added that this preference to interpret the first noun as the actor might be due to the higher statistical frequency of active structures in actual utterances as opposed to passives. Yet, he conceded, there was little known about the frequencies of these constructions at that time.

Building on their previous work, Slobin and Bever (1982) proposed that children construct a canonical sentence schema as a preliminary organizing structure for language. The researchers investigated 48 monolingual children between the ages of 2;0 and 4;4, who were native speakers of English (rigid word order, uninflectional), Italian (weakly ordered, weakly inflectional), Serbo-Croatian (weakly ordered, inflectional), and Turkish (minimally ordered, inflectional), and their sensitivity to word order and inflections in simple transitive sentences in their respective languages. Slobin and Bever hypothesized that if word order strategies were fundamental then children of all languages should initially posit a strict word order strategy regardless of the fixedness of their language's word order. Although across the age groups they found strong evidence for a universal word order strategy in all four language groups that averaged 3;6 in age, the age at which overgeneralizations of word order had been previously reported in English-speaking subjects, they also found that the children were prepared to develop word order and inflectional strategies according to the properties and regularities of their language. Hence, the authors concluded that there could be some additional maturational factors involved in determining when children take into account the presence of a second noun, thus eliminating the need to assume the first noun is the subject, but that these would only occur with further development. When children can process utterances with three major phrases, such as when they have sufficient working memory to hold three constituents while processing them, may be the point at which they begin to process argument structure in a more language-specific way.

Hoping to circumvent the problem of underdeveloped cognitive resources present when studying children, LoCoco (1982) investigated college-aged students learning a second language on their interpretations of sentences with various word orders. She tested L1 English learners of Spanish and L1 English learners of German on matching pictures to sentences that they heard. She found that the learners she tested made many errors in interpreting the subject of the sentence when it was postverbal such as in (2).

(2)	*A*	*la chica*	*le*	*trae*	*cerveza*	*el chico.*
	[To	the girl	3^{RD}SINGDAT	bring-3^{RD}SINGPRE	beer	the boy]
	The boy brings the girl beer.					

LoCoco concluded that this type of error was due to the same universal strategy that Bever (1970) had reported children used in their L1.

Continuing along the same lines of her previous work, LoCoco (1987) compared comprehension of aural and written sentences by L1 English first-year learners of German and L1 English first-year learners of Spanish. Learners were asked to mark the drawing that was best described by the sentence they heard. The same method was then repeated, except that the participants also saw the sentences as they heard them. LoCoco found that both groups of learners relied on word order in most instances when reading and listening to sentences of varying word order. Only certain prepositions and inanimate nouns (which participants eliminated as possible subjects due to their lack of animacy) helped learners disambiguate the subject and object when the word order was OVS. The learners of German were slightly more aided by the presence of cues for case, whereas the learners of Spanish continually relied on word order cues to comprehend the sentences. LoCoco stated that this increased tendency for the learners of Spanish to rely on word order (as opposed to the learners of German) may be due to the fact that basic Spanish word order is similar to English word order. She concluded that the abandonment of the reliance on word order strategies sooner by the learners of German suggested that the structure of the second language influences comprehension strategies, although seemingly not in the beginning. This research suggests the possibility of a trajectory of language learning that spans universal and language-specific strategies as a learner increases in her proficiency.

1.2.2 The First Noun Principle

The aforementioned research focused on finding universal strategies present during L2 processing of argument structure, specifically in identifying the subject of a sentence. Following along these lines and adopting the same method as LoCoco (1982), VanPatten (1984) focused on the preverbal object clitic structure in Spanish. He tested college-aged L1 English learners of Spanish on two aural comprehension tests involving a picture-matching task. Participants saw the same four pictures while they heard twelve sentences in Spanish. Based on the sentence they heard, they were to circle the letter of the picture that was described. In the first test, VanPatten used sentences in which only direct object clitics and not the full noun phrase were present, such as in (3).

(3)	*Lo*	*invitan*	*ellos*	*al cine.*
	[3RD SINGMASCACC	invite-3RD PLPRE	they	to the movies]
	They invite him to the movies.			

In experimental sentences, VanPatten varied the gender and the number of the subject and the object. During the testing, the author made sure that the participants were aware of who was performing the action in each of the pictures to ensure that they could choose the correct picture based on their interpretation of the sentences. The second test involved the use of indirect object clitics in sentences such as in (4).

(4) *Le* *dan* *dinero ellos.*
[3RD SINGDAT give-3RD PLPRE money they]
They give him/her money.

VanPatten found that both first and second semester learners did not perform significantly differently. They definitively interpreted the clitics as agents of the action; in some cases incorrect picture selection was as high as 70%. VanPatten also found that the learners made significantly more errors interpreting the direct object clitics than the indirect object clitics. Moreover, he found that participants were not paying any attention to the object marker *a* when it was present before a direct object noun. VanPatten explained that this might be due to its low acoustic salience or to the fact that it has no semantic or syntactic equivalent in English, and therefore the learners skip over it and rely on word order to interpret the sentence (incorrectly, in this case). The same is not true for *a* as a dative or indirect object marker, since this exists and is translated into English as *to.* This could explain why the learners made fewer errors on interpreting the indirect objects. Another possible reason that VanPatten posited for the learners performing better on the indirect object clitic sentences is that the indirect object clitics in Spanish are unique in form and function, whereas the singular feminine and plural forms of direct object clitics are the same form as the singular feminine and plural determiners in Spanish, perhaps rendering them more challenging to acquire due to this many-to-many form-meaning mapping, which is why I have chosen to examine further how second language learners process direct object clitics and not indirect object clitics in the present dissertation. Based on the evidence that these learners of Spanish interpreted sentences of the form: clitic-pronoun-verb-subject as subject-verb-object, VanPatten concluded that this learner strategy may be part of the "back up system" for processing input. Later, as part of his Input Processing Theory, VanPatten articulated this as the *First Noun Principle*, as shown in (5).

(5) *Learners tend to process the first noun or pronoun they encounter in a sentence as the subject or agent.*

However, since he had only tested L1 English (rigid SVO word order) learners of Spanish, VanPatten's results were not sufficient to conclude that a universal strategy, and not transfer from the L1, is at play.

Also looking at subject identification in sentences with varied word order (SVO, SOV, and VSO), Gass (1989) tested L1 English learners of Italian as well as L1 Italian learners of English. Since English has a rigid word order, yet Italian has a more flexible word order, Gass proposed that the L1 Italians would use animacy cues over word order cues to determine subjecthood. She found that the L1 English learners of Italian indeed had a stronger preference for identifying the first noun in all of the word order sequences as the subject. By contrast, the L1 Italian learners of English took into account the animacy of the nouns in the structures presented, and relied less on the word order to determine subjecthood. With the inclusion of learners from an L1 other than English, this suggests, instead of a universal default for processing strategies as VanPatten (1984) had proposed, that other factors such as the structure of the L1 may influence argument structure processing strategies in the L2. However, since the two groups were learning distinct languages, it is hard to compare their success rates. Another way to account for the data found in the studies above and in the present dissertation may be to consider a transfer account, which I illuminate in the following section.

1.3 Studies on L1 Transfer

If a language learner uses processing strategies that derive from his experience in the L1 in order to process argument structure in the L2, then it is said that he has transferred these strategies from his L1 to his L2. The term "transfer" here refers to the influence of native language structures and their processing on L2 processing. In the present section, I review the Competition Model (Bates and MacWhinney, 1979, 1982; MacWhinney, 1987, 1989, 1992), a major framework that takes into account the concept of transfer, as well as pertinent studies conducted by researchers who account for their results by employing this framework. In addition, I review similar studies whose authors employ the concept of transfer without adopting the specifics of the Competition Model to explain their findings.

1.3.1 The Competition Model and Studies within this Framework

In addition to the study done by Gass (1989), there is a large body of research that has suggested that processing strategies are not universal, but rather directly transferred from the L1. Working from a more functionalist perspective, MacWhinney and colleagues have posited the Competition Model (Bates and MacWhinney, 1979, 1982; MacWhinney, 1987, 1989, 1992). Not unlike VanPatten's principles, the Competition Model is a language-processing model based on many-to-many form-function mappings. Some research along these lines thus far has indicated that L2 learners initially transfer their L1 strategies to their L2, adapting them as is appropriate to the L2 with increased exposure. In the case of assigning agent status in a sentence in English, MacWhinney and colleagues' connectionist model included cues such as preverbal positioning, animacy, verbal agreement morphology, sentence initial positioning, nominative case-marking for pronouns, and use of the article *the*. For example, in (6), these cues lead the reader to believe that *the boy* is the subject/topic/agent/actor of the sentence.

(6) The boy shatters the window.

Use of these cues that mark function, however, is probabilistic in that in another example in the same language, (7), the only cue to subjecthood is the preverbal position of *the wind*.

(7) The wind shatters the window.

Consider also (8), in which the cue of preverbal position would mark *the window* as the topic/subject, while noun animacy and the preposition *by* point to *the boy* as the actor/agent, hence, the probabilistic nature of the form-function mapping.

(8) The window was shattered by the boy.

With all of the possible cues available in a given language, the Competition Model posits that sentence interpretation occurs by way of the competition and cooperation of these cues, and by their strength, validity and reliability. Since form-function mappings differ cross-linguistically, the strength of the cues will also differ across languages. In addition, the cues that are the most frequent and reliable (i.e. valid) for a particular language have larger cue weights. For example, in Spanish, the object marker *a* serves as a cue to the object (although not a salient one for L2 learners as VanPatten (1984) found), and the subject would then be the noun that is not the object. Such a cue does not exist in English, in which the dominant cue is preverbal positioning for determining subjecthood. When learning a second language, the process includes starting with L2 cue weights that are close to those for the L1. Slowly, the cue weights are transferred from the L1 and then must be "retuned," according to MacWhinney (1997). Where languages have no corresponding form-function mappings, L2 learners can only transfer the meanings, and then start from the beginning constructing new mappings from functions to forms (MacWhinney, 1987). Transfer occurs at all levels of linguistic representation (lexical, syntactic, and phonological) in this model. In the early stages of acquisition, there are considerable amounts of transfer, according to the Competition Model.

The process by which second language learners incorporate or abandon cues from their L1 in order to process the L2 is testable within the framework of the Competition Model. Working from MacWhinney's original Competition Model framework, Harrington (1987) compared L1 English learners of Japanese and L1 Japanese learners of English based on their processing strategies. He found that the L1 Japanese learners' strategies were more influenced by animacy cues, whereas the L1 English learners were more sensitive to word order cues, supporting the idea that learners transfer cue weights from their L1 to their L2. There is a host of other work that has looked at cue weights and cue validity in determining the agent of a sentence that supports this model (see for example, for German, MacWhinney, Bates, and Kliegl, 1984; for Hungarian, MacWhinney, Pléh and Bates, 1985; for Dutch, De Bot and Van Montfort, 1988; for Italian, MacWhinney and Bates, 1978; for Chinese, Li, Bates, Liu, and MacWhinney, 1992; for Japanese, Kilborn and Ito, 1989; for Hebrew, Sokolov, 1988).

Another example of this is Kilborn (1989), who found evidence for transfer during online processing of L2 English by L1 German participants. The author predicted that, due to the allowance of more variations in word order in their L1, German-speakers would rely less on word order cues and more on verbal morphology. Presented with stimuli of varying conditions (word order, animacy, and verb agreement permutations), the participants indicated the actor or subject of the sentence. The English-speaker controls patterned with previous results (e.g. Bates, MacWhinney, Caselli, Devescovi, Natale, and Venza, 1984), in that they relied on word order cues for sentence interpretation. The L1 German learners of

English took longer overall in both their L1 and L2 to respond to the items, which the author argued was an indication that the participants were waiting for all of the potential cues to be provided before they made their decision. Specifically, in their L2 English, the German-speakers employed agreement and animacy cues over word order cues. Although the German participants showed some use of word order cues in their L2 English, they did not fully rely on them as a native speaker would. According to Kilborn, the results from this experiment support a theory of transfer of L1 strategies to L2 processing.

Also working from a Competition Model framework, Heilenman and MacDonald (1993) tested less advanced L1 English learners of French on their development of the use of word order versus clitic pronoun type and agreement cues in sentence interpretation in the L2. According to the authors, French native speakers tend to rely on information provided by clitic pronouns and verb agreement more than on information from word order, which is not the case for English native speakers as shown previously. Heilenman and MacDonald found that all four of their proficiency groups of L2 learners used word order cues, although all but the least experienced group did not differ significantly from the native French speakers in this use. Furthermore, only the most advanced learner group incorporated clitic pronoun agreement cues in their sentence interpretation, which did not approach the level of use shown by the French native speakers, and were also the only group to not treat object clitic pronouns as subjects. The authors explained that the learners' rapid decline in use of word order cues was a response to the number of errors that this type of strategy would induce based on French input. However, the L2 learners did not quickly adapt to L2 cue strategies even though they had discounted their L1 strategy. Heilenman and MacDonald concluded that the apparent transfer of English word order strategies into French by the lower three proficiency groups was due to their direct translation of the French stimuli into English under the erroneous assumption that there is a one-to-one mapping between each French item and its English translation. Thus, Heilenman and MacDonald stated that their results suggested, along with others done within the Competition Model framework, that there is clear evidence that L2 learners make an effort to transfer their L1 strategies at first. When these strategies fit in the L2, then learners continue with these strategies. When the strategies prove to be of no use, then they are abandoned rapidly, perhaps indicating a variable intermediate stage between the transfer of L1 strategies and the complete acquisition of L2 strategies. In short, Heilenman and MacDonald posited that based on language experience, L2 learners will choose to keep or jettison L1 processing strategies when processing in the L2. Defining language experience in terms of experience with interpreting specific cues in the L2 in order to process argument structure properly, the Competition Model provides a viable account for the data presented above and in the present dissertation.

1.3.2. Additional Studies on Transfer

Juffs (1998) also attributed his results to possible transfer of L1 parsing strategies, without naming the Competition Model specifically as his framework. Asking if advanced second language learners transfer lexical representations from the L1 to the L2, Juffs tested L1 Romance, L1 Chinese and L1 Japanese/Korean learners of English on verbs that alternate between causative and inchoative structures, which differ crosslinguistically. Example structures are shown in (9).

(9) a. First of all the cook melted the chocolate on the cake.
b. *First of all the chocolate melted itself on the cake.
c. First of all the chocolate melted slowly on the cake.
d. First of all the cook made the chocolate melt on the cake.

Since Japanese and Korean are head-final languages, in which bound causative morphology is required for verbs that alternate between being causative and inchoative, speakers of these languages were grouped together. Chinese, on the other hand, requires overt free, preverbal causative morphology for verbs that alternate between causative and inchoative. Romance languages, which are basically SVO, require the anticausative morpheme 'se' to express this meaning. In a moving window experiment, Juffs asked participants to make plausibility judgments at the end of each sentence. The judgment accuracy results revealed that the L1 Chinese learners, the least accurate, preferred the form that was most like their L1 form. Online measures showed a discrepancy between the knowledge of argument structure that participants displayed on offline measures and their ability to parse these structures online. The L1 Japanese/Korean learners were the most sensitive to argument-structure differences in the causative/inchoative verb class. Juffs speculated that this may be due to L1 transfer.

Fernández (1998) considered the effects of transfer on both early and late second language learners. She summarized the results from an offline questionnaire about sentences with syntactic ambiguities based on relative clause attachment that she had conducted both with L1 Spanish (Fernández, 1995) and with L1 Japanese (Fernández and Hirose, 1997) early and late learners of English, which showed evidence for what she named *forward transfer* for both groups of late learners of English. Although Fernández (1998) defined two types of transfer, *forward transfer*, in which bilinguals transfer their L1 strategies into their L2, and *backward transfer*, in which bilinguals transfer their L2 strategies into their L1, I only consider *forward transfer* in the present dissertation. In the first study (Fernández, 1995), the L1 Spanish late learners preferred high attachment of relative clauses (attaching the subordinate clause to the first noun phrase instead of the second) over low attachment in the English sentences, which indicated transfer of parsing strategies from their L1. The English monolinguals preferred low attachment, and the early L1 Spanish learners showed a range of preferences in between, the heterogeneity of which surprised the researcher, and perhaps indicated a variable intermediate stage between the transfer of L1 strategies and the independent use of L2 strategies. In a separate study with the same materials (Fernández and Hirose, 1997), L1 Japanese participants showed no difference between early and late learners (although the author addressed this issue later and revealed that they were a rather homogeneous group), but were significantly different in their preferences from the monolingual English speakers, again showing evidence of transfer from their L1. Based on the results from her studies, Fernández deduced that the influence that the L1 exerts on L2 parsing strategies is great not only in the case of late second language learners, but in the case of some early second language learners also (Fernández, 1998).

Frenck-Mestre (2002) examined the results from various studies on adult bilinguals' syntactic analysis abilities, including Fernández (1998), in order to determine the role of *forward transfer*. She first described a study from Frenck-Mestre (1997), in which low proficiency English-French bilinguals read French sentences that included syntactic ambiguities and exhibited a very different pattern than the native French speaker controls. In sentences such as in (10), where number agreement between the first or the second noun phrase and the verb of the relative clause disambiguated the structure, the English-French group preferred low attachment (attaching the subordinate clause to the second noun phrase instead of the first), whereas the French native-speakers preferred high attachment.

(10)	*Aline telephone*	*aux filles*	*de la gardienne*	*qui*	*reviennent/revient*	*de Paris.*
	Aline calls	the daughters	of the nanny	who	are/is returning	to Paris.

Since it has been shown that in English there is a low attachment preference, Frenck-Mestre argued that it is most likely that the pattern shown by the English-dominant learner group can be attributed to the influence of their L1 rather than to some general (or universal) strategy of attaching a clause or phrase to the most recent constituent processed. Frenck-Mestre continued on to report results from a higher proficiency learner group, which patterned very closely to the native-speaker group, but had also had immersion experience in the second language environment. Frenck-Mestre concluded that her data revealed a clear evolution of processing strategies, from performance in low proficiency learners that resembles their L1 to performance in high proficiency learners that closely resembles native speakers' performance in the language. However, comparing the data from these language learner groups as if they were longitudinal data is problematic due to the discrepancy among their language histories, with the high proficiency group having spent time abroad (as noted in Dussias and Sagarra, 2007). Language acquisition in a study-abroad setting involves a different language experience based on type and amount of input than language learning in a classroom in the L1 environment, thus complicating the comparability of these two groups.

In addition, without a learner group whose L1 also preferred high attachment, Frenck-Mestre's conclusion that her results are due irrefutably to L1 influence is untenable. To address this issue, Frenck-Mestre (2002) then compared her results to Fernández' (1998) offline results for Spanish-English and Japanese-English bilinguals, which showed that level of proficiency, based on a self-rating, was predictive of native-like performance on the L2 syntactic task in that the Japanese group claimed that they were of lower proficiency and showed transfer effects from their L1. In short, Frenck-Mestre (2002) provided evidence of transfer from learners of different L1's. She showed that less proficient language learners process their L2's like they process their L1's in the case of attachment preferences, while more proficient learners appeared to process the L2 like native speakers.

1.4 The Associate-Cognitive CREED

Simply positing the idea of transfer of strategies from the L1 to the L2 as the answer to the question of how second language learners process argument structure in their L2 leaves many questions unanswered regarding how and at what point in a learner's development this transfer occurs. The Associative-Cognitive Model, centering on language experience and use, and going beyond the idea of salient cues, not only has offered an attractive compromise between the universal and the transfer accounts, but also has provided some answers about second language acquisition that previous models have not offered. The present dissertation will adopt this theory as its frame of reference. The Associative-Cognitive model or CREED (Ellis 2006a) states that second language acquisition occurs based on the same cognitive and associative principles as other types of learning. Ellis' framework has five main principles, which hold that second language acquisition is: 1) Construction-based, 2) Exemplar-driven, 3) Rational, 4) Emergent, and 5) Dialectic. According to Ellis, second language learning is construction-based in that learners acquire constructions that map linguistic form and function. It is exemplar-driven due to the idea that learners tune in to the frequency of memorized exemplars of use of these constructions. Second language learning is rational in that prior L1 usage is taken into account first and to its utmost capacity when processing in the L2. A second language system is emergent because it develops over time in complex and adaptive ways. Finally, in this framework, second language learning is dialectic due to the often non-native-like endstate reached that Ellis (2006a), citing Klein (1998), describes as having the following characteristics: adequate for basic communication, containing little or no functional inflection, and lacking most if not all closed-class items such as determiners, subordinating elements, and prepositions (2006a: 110). These characteristics are pervasive in

L2 learner language such that many researchers believe them to be so permanent and consistent that it is as if all L2 learners shared their own dialect.

Ellis (2006b) further delineated second language acquisition as being rational by addressing its apparent irrationalities, in other words when language learners receive the necessary input but fail to acquire certain features of the second language. Ellis defined important concepts from his theory that successfully explain these apparent irrationalities. He divides the concept of 'learned attention' into phenomena such as contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning.

First, Ellis discussed the factors of contingency, competition between multiple cues, and salience, as they affect both L1 and L2 acquisition. Contingency learning, according to Ellis, is when a language learner is faced with cues that point to particular outcomes and can learn to use these cues to acquire the new language. For example, in English, the morpheme –s can serve as a cue to plurality, as in 'book-s'. If a learner notices this relationship, a relatively frequent one, then he makes the association between the cue and the outcome. However, the –s morpheme can also indicate the third person singular present of a verb such as in 'he walk-s'. Hence, the learner is faced with a cue that can have multiple outcomes. Further confounding the issue, in English the notion of plurality can be expressed without the use of –s, as in 'fish' or 'data'. Thus, learners not only must grapple with instances of a cue being present without it predicting a particular outcome but also must understand that the outcome may occur in the absence of this cue as well. According to Ellis, a contingency learning paradigm predicts that these cue-outcome associations are not easily learnable. In respect to the present dissertation, L2 learners could confuse preverbal object clitics with cues for definite articles since the forms are identical in the singular feminine (*la*) and the plural forms (*los*) and (*las*), thus rendering the preverbal object clitics more difficult to acquire.

Next, Ellis (2006b) related his theory back to work done by MacWhinney (1997, 1987) within the Competition Model, concurring that learners deal with the presence of multiple cues and the competition among them by considering the cues' validity and strength, factors which MacWhinney had described as being based on frequency of occurrence. Ellis also cited the shifting of cue preferences that MacWhinney predicted for second language learners would occur with sufficient exposure to the new language (see MacWhinney, 1987). The Competition Model, however, Ellis stated, does not account for how learners acquire cues that have low salience (i.e. are overshadowed by more obvious cues). To explain this case, Ellis drew from the Rescorla-Wagner (1972) model of conditioning that incorporated the concept of salience (or perceived strength) of cues, which plays a determining role in how much learning results from the input containing such cues. Ellis noted that the majority of the form-meaning mappings that are challenging for second language learners to acquire are those that have low salience in the speech stream, such as the third person singular present grammatical morpheme '–s' in English. These bound inflectional morphemes, as well as grammatical function words, such as clitics, are typically short and unstressed, and thus difficult to perceive. This phenomenon is particularly relevant for the present dissertation since it focuses on the processing of clitics, low salience structures. Ellis concluded that the factors of frequency, salience, and contingency serve as significant predictors for the acquisition difficulty that second language learners have with particular morphemes.

Finally, Ellis (2006b) elaborated upon the concepts of L1 interference, overshadowing and blocking, and perceptual tuning, all of which offer an explanation of why

L2 acquisition is not fully realized without drawing on an age-related account. When analyzing the concept of L1 interference, Ellis took into account both Proactive Inhibition (PI), the notion that prior learning inhibits future learning, and the Contrastive Analysis Hypothesis (CAH) (Lado, 1957), which stated that the elements similar to one's first language will be easy to acquire and the elements that are different will be difficult to acquire. This is basically the concept of transfer. In addition, Ellis linked these accounts to mutual exclusivity in child language acquisition, saying that mutual exclusivity and PI are essentially the same as well (2006b: 176). By synthesizing the critical and widespread accounts of mutual exclusivity, the CAH, the Competition Model, and PI, Ellis justified the inclusion of L1 interference in his theory. In the present dissertation, L1 interference would come into play in that the L1 English learners of Spanish would have the L1 bias of subjects surfacing preverbally and objects surfacing postverbally, and may therefore interpret the preverbal clitics and postverbal subjects incorrectly as opposed to the L1 Romanians, who may have less difficulty because the target structure is present in their L1.

In the case of overshadowing and blocking, Ellis again adopted concepts from associative learning research. Both phenomena shape our attention to language, according to Ellis. He described overshadowing as the situation in which two cues presented together predict the same outcome, yet one cue is more salient than the other and therefore becomes associated with the outcome. Hence, the less salient cue is overshadowed. As overshadowing plays out over time, Ellis wrote, it produces a learned selective attention called blocking. Cues that are experienced in conjunction with known highly salient cues are blocked from being predictive of the outcome. Ellis cited work done by Chapman and Robbins (1990) and by Kruschke and Blair (2000) that supports the presence and importance of these concepts in associative learning. He also relates these concepts to second language learning, in particular to the acquisition of the form-meaning relationship of inflectional morphemes marking tense that often co-occur with temporal adverbs. The co-presence of temporal adverbs, which are typically acquired early by second language learners as opposed to late in child L1 acquisition, renders the inflectional morphemes redundant in their predictability of tense. Ellis remarked this redundancy has more impact in second language acquisition than in first language acquisition (2006b: 179). Under the above conditions of low salience and blocking, Ellis highlighted that no amount of extra input would be sufficient to save the learner from a "fossilized" interlanguage that bespeaks incomplete acquisition. Just as in the situation outlined previously with the co-occurrence of adverbs and inflectional morphemes that both predict tense, it is possible for the structure considered here to be redundant in normal discourse as well, which would be problematic for learners according to this theory. In Spanish, preverbal clitics can be doubled such that object information is present both in the preverbal clitic as well as in a postverbal full object NP. Thus, following the Associative-Cognitive CREED, this doubled version of the present structure would provide direct object information to the learner in two ways, which might make the less salient preverbal clitic less likely to be acquired. Experimentally, I did not include doubled structures in the present dissertation, although they are present and frequent in a learner's input. The present dissertation aims to demystify the apparent challenge that L1 English speakers have when interpreting preverbal clitics and postverbal subjects in Spanish.

The final concept that Ellis discussed as part of his theory was perceptual tuning. Ellis explained this concept from the perspective of perceptual systems that change their structure over the course of their history. Due simply to usage, Ellis wrote, perceptual systems become more sensitive to those characteristics of the stimuli that indicate psychologically significant dimensions of variation, and become less sensitive to those characteristics that are redundant, and therefore useless in aiding in classification. Perceptual learning, or tuning, results from experience with exemplars. As learners become more familiar with particular exemplars, they are better able to tune in to the features that distinguish an exemplar of one type from another. Werker and colleagues (Werker and Lalonde, 1988; Werker and Tees, 1984) have documented this tuning, or development of sensitivity to detect a particular feature, during the process by which infants acquire the phonemic inventories of their native language. Initially, infants' brains are highly plastic, allowing them to perceive the phoneme contrasts of every possible language. However, by the end of their first year they are less sensitive to phonemic contrasts that are not present in their own language. Ellis pointed out that second language learners, on the other hand, do not begin learning the L2 with the same neural plasticity as newborns, but with a brain already tuned to the L1. Ellis cited work done by Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, and Kettermann (2003), who looked at how early language experience changes low-level perceptual processing such as in the case of L1 Japanese speakers learning English. Since in Japanese the phonemes /r/ and /l/ are part of the same category, they have more difficulty with them in English, in which /r/ and /l/ are separate phonemes, than L1 speakers of German, who also have separate phonemic categories for these sounds. Connectionist simulations have corroborated this work (e.g. McClelland, 2001) and extended it to differentiate between early and late L2 learners, showing that early, balanced simultaneous bilinguals successfully kept the two languages separate whereas less balanced or later L2 learners showed little L1-L2 separation with interference and transfer at a maximum (e.g. Li, Farkas, and MacWhinney, 2004). This research on perceptual tuning has all pointed to the idea that adult L2 acquisition difficulties are related to the effects of entrenchment of the first language. Again, the present dissertation separates out the variable of L1 in order to determine its role in the processing of structures with preverbal clitics and postverbal subjects in Spanish.

Ellis (2006b) concluded that second language acquisition is a multi-dimensional continuum affected by a multitude of factors including frequency, contingency, salience, interference, overshadowing and blocking, and perceptual learning. Learners are continually biased by these factors as they attempt to selectively attend to the cues in their L2. In sum, the Associative-Cognitive CREED put forth by Ellis (2006a, 2006b) is a broad framework, spanning psychological, linguistic, neurological, social, and educational phenomena. It incorporates both cognitive and social factors as important elements in the development of a second language system, basing its assumptions on a myriad of empirical evidence and making it one of the more viable accounts of second language acquisition. The CREED also adopts the view that regardless of whether one subscribes to a language universals-type theory or a more connectionist Competition Model-type theory, second language acquisition, and thus second language processing of argument structure, is a dynamic process from which strategies and regularities emerge based on the interaction between the people, societies, and cultures that use that language. Given the elements of Ellis' theory, the CREED provides a particularly useful lens for generating predictions for the present dissertation.

Ellis and Sagarra (2010) have tested the specific claims of the Associative-Cognitive CREED with two experiments in which they taught native English-speaking participants a condensed version of a Latin verb paradigm in order to determine whether learners' acquisition of one set of temporal reference cues (verbal inflections) was affected by previous knowledge of another set of reliable temporal reference cues (adverbs), and vice versa. They pretrained one group on verbal inflections and a second group on adverbs in Latin, exposed all participants to sentences combining the different verb forms and their appropriate adverbs, and asked them to judge the temporal reference of all of the verb/adverb combinations and translate from English to Latin the various verbs and adverbs they had learned. Ellis and Sagarra found that the adverb pretraining group judged sentences that had a mismatch in temporal reference by the temporal reference of the adverb, whereas the verbal inflection pretraining group chose the temporal reference of mismatched strings as the temporal reference of the verb. Hence, their data revealed significant and large effects of the blocking of later learned cues by previously learned cues in the early acquisition of a language. To further this line of research, Ellis and Sagarra conducted a second experiment in which they investigated long-term effects of learned attention on cue acquisition through the testing of Chinese-English bilinguals (Chinese speakers rely more heavily on lexical cues to temporal reference since there is no morphology in Chinese that corresponds to tense) on the same materials by the same procedure. The only exception was that there was no pretraining phase since Ellis and Sagarra figured that a life of learned attention directed towards lexical cues such as adverbs and prepositions would more strongly bias this particular participant group than, for example, the brief adverb pretraining had biased the participants in the first experiment. In short, the L1 Chinese participants in the second experiment, even with their exposure to English morphology, patterned like the adverb pretraining group in the first experiment on their perception of temporal reference when mismatched cues were present in the stimuli. Ellis and Sagarra thus confirmed that their experiments showed clear short- and long-term learned attention effects and the resulting cue blocking in early second (and third) language acquisition, demonstrating the strong influence of L1 background on language learning in general.

1.5 Mixed Account Studies

1.5.1 Studies Showing No Evidence of Transfer

Although the argument and evidence presented above for L1 transfer seem robust, some recent studies have failed to find transfer effects in second language learners. It is important to consider these studies in order to perhaps determine why transfer effects were not found. Another factor involved in the processing of L2 argument structure that Ellis in his account of second language acquisition did not specifically mention, but incorporated under the umbrella of language experience, is proficiency. As researchers have begun to include learners of more varied L1's and L2's, the effects of proficiency have become decidedly more relevant when testing the question of transfer or universal strategies. Studies that have shown little effect for transfer, such as Marinis, Roberts, Felser, and Clahsen (2005) and Cuervo (2007), have only included participants at one level of proficiency. Studies that have included proficiency as a factor, such as Hopp (2006), have provided more robust evidence for transfer.

Although Marinis, Roberts, Felser, and Clahsen (2005) tested L2 learners of English from L1's as diverse as Chinese, Japanese, German, and Greek, they found no effects of transfer in their online reading time experiment. All participants, however, were at the same level of proficiency based on the results from a language proficiency test. While reading sentences with long-distance *wh*-dependencies in a moving window task, the participants answered comprehension questions. The accuracy results showed no differences between language groups nor between L2 learners and the monolingual controls. In terms of reading times, the authors expected no statistical differences between groups if the learners were processing long-distance *wh*-dependencies in the same way as the native speakers. However, if the L1 did have an effect, then the authors predicted that the L1 Chinese and L1 Japanese (these languages lack successive-cyclic *wh*-movement) participants would differ from the L1 German and L1 Greek participants (these languages have successive-cyclic *wh*-movement). Finally, if L2 processing differed from L1 processing but was not influenced by L1 grammar, then Marinis et al. expected to find differences between native speakers and learners, but not among individual language groups. The results revealed that the learner groups behaved significantly differently from the native speakers in that they did not posit intermediate gaps while processing long-distance dependencies. More interesting, the results indicated that all of the L2 groups processed the experimental sentences in essentially the same way but differently from the native speakers even though this study tested four groups of L2 learners with similar levels of proficiency in the L2, two of which were from *wh*-movement backgrounds and two of which were from *wh*-in-situ backgrounds. Marinis et al. explained this absence of transfer effects by stating that L2 learners' sensitivity to syntactic information during real-time processing is more limited than that of native speakers, and that L2 learners seem to rely more on lexical-semantic and other nonsyntactic cues for sentence interpretation. It is possible, however, that these participants were at a level of proficiency in which they no longer transferred L1 syntactic strategies directly, had begun to approximate the level of native speakers, but not yet fully, and were still relying on lexical-semantic cues for comprehension. With participants of a range of proficiency levels, Marinis et al. may have obtained more clear results. The lack of transfer effects in this study may stem from the researchers testing participants who were in flux between an L1 transfer and an L2 nativelike state, when their processing strategies may have been universal for second language learners regardless of L1.

Considering adult L1 English learners of Spanish also at an early stage of development, Cuervo (2007) reported an experimental study on their acquisition of the double-object construction in Spanish. On a grammaticality judgment task, the participants in her study rated the acceptability of sentences that involved double-object constructions such as in (11a), and constructions that were ungrammatical in Spanish such as in (11b) or (11c), but that followed an English word order.

(11) a. *Le mandé una carta a Juan.*
[3rd SingDat send-1st SingPa a letter to Juan]
I sent a letter to Juan to him.

b. **Te dejé los libros a Juan.*
[2nd SingDat leave-1st SingPa the books to Juan]
I left the books to Juan to you.

c. **Compré mi amigo una planta.*
[Buy-1st SingPa my friend a plant]
I bought my friend a plant.

The author predicted that if there were transfer of L1 word order involved in the acquisition of double-object constructions, then the participants would accept the above examples that showed no agreement between the dative and the clitic or had the wrong case and wrong word order. However, she found that the learners correctly rejected these constructions just as the native speaker controls had, suggesting that there was no transfer of L1 syntactic processing to the L2 at this stage of development. Although Cuervo stated that she found no conclusive evidence for transfer, her task was an offline metalinguistic task that did not tap into the participants' online processing abilities. With an online task, which allows participants less time to analyze these structures, thus making them more cognitively taxing, it is possible that learners may rely on their L1 more and show transfer effects.

Grüter and Crago (2010) conducted a study that considered the roles of L1 transfer and processing limitations in L2 French object clitic acquisition. Just as with learners of Spanish, learners of French have shown to have difficulty in mastering object clitic constructions (e.g. Herschensohn, 2004). To disentangle transfer effects from a more universal default processing strategy, Grüter and Crago tested L1 Spanish and L1 Chinese learners of French on an elicitation task and a truth-value judgment task (TVJ) involving null objects. Since Chinese allows for null objects, the authors predicted that if transfer were involved then the L1 Chinese learners would be more likely to omit objects and to accept structures with null objects in French, while L1 speakers of Spanish, which contains a similar clitic construction to French, should perform more like the French native speakers on such constructions. The participants were all children between the ages of five years ten months and ten years in the same French school system in Montreal, showing no significant differences in age, length of exposure to French, or age on arrival in Quebec. However, the authors never mentioned the children's level of proficiency in French nor their daily usage of the language, which could differ significantly. The results from the elicitation task revealed that the children in the Spanish group produced more clitics than those in the Chinese group, and omitted them significantly less than their Chinese counterparts. Moreover, the distribution of omission was clustered between two out of 11 of the L1 Spanish participants, whereas it was more spread out among the L1 Chinese participants in that 13 out of 15 of them omitted at least one clitic. In the TVJ task, both groups were at ceiling in rejecting the null object condition, and overall showed no statistically significant differences in their responses. Thus, Grüter and Crago concluded that the transfer-based prediction that the L1 Spanish participants would produce more clitics in the elicitation task as opposed to the L1

Chinese participants was supported based on their equal length of exposure to French, again leaving out any information on their relative proficiency levels. Furthermore, the authors argued that the lack of statistically significant difference between the two groups' performance on the TVJ task might be due to either a lack of statistical power in the task or to a differential impact of positive transfer on production as compared to comprehension. Also, the fact that the L1 Chinese participants correctly rejected null object constructions in French could be a product of their proficiency, the authors suggested, in that they had already restructured their interlanguage from its initial state to a more developed state. Of course, the authors noted, Chinese allows for both overt and null objects, which could confound the situation by creating a potential superset-subset relation between the L1 and the L2 and might be evidence against a transfer account. Hence, Grüter and Crago (2010) proposed a mixed hypothesis that L1 transfer is limited to overt material (156). This would account for the L1 Spanish learners showing more production of clitics than the L1 Chinese learners and for the L1 Chinese learners not showing negative transfer of a null object condition in the TVJ task.

Taking proficiency into account, Hopp (2006), like Marinis et al. (2005), argued that a certain level of proficiency is necessary to show transfer effects on complex structures. He tested L1 English and L1 Russian speakers of German of different proficiency levels on two online tasks involving subject-object ambiguities in L2 German. In German and Russian, objects can optionally precede subjects, which is overtly expressed either by case marking on definite determiners as in German (12a) or by suffixes on the nouns as in Russian (12b), while in English none of this is possible.

(12) a. *Maria glaubt, dass [den Onkel] der Vater schlägt.*
[Maria thinks that [the-ACC uncle] the-NOM father beats.
Maria thinks that the father beats the uncle.

b. *Utrom* *djadj-u* *udaril* *dedushk-a.*
[In the morning, the uncle-ACC hits the grandfather-NOM.
In the morning, the grandfather hits the uncle.

Stating that native speakers of German and Russian had previously shown a strong subject-initial preference in relative clause processing, Hopp cited the Minimal Chain Principle by De Vincenzi (1991), as shown in (13), as the universal parsing principle underlying this preference.

(13) *Avoid postulating unnecessary chain members at S-structure, but do not delay required chain members* (De Vincenzi, 1991, p. 13).

According to this principle, the parser would first construct a default subject-initial phrase structure and then have to revise it to a more elaborate and cognitively taxing object-first order upon encountering a garden-path. In order to determine if the driving force for L2 learners' parsing strategies is a universal principle as suggested above or transfer of parsing strategies from the L1, Hopp asked participants to complete a self-paced reading task, which involved reading and answering comprehension questions on sentences with subject-object and object-subject word order that contained syntactic ambiguities. The accuracy results on this task revealed that constituent order and how the ambiguity was resolved had significant effects for the native and near-native participants, whereas first language did not. Also, although all participants were slower in their reading times for object-subject ordered constituents than in their reading times for subject-object ordered constituents, there was no significant main effect of order or ambiguity for the advanced participants. In the speeded grammaticality judgment task, again there was no difference for language on accuracy scores, with all groups demonstrating a preference for subject-object constructions over object-subject constructions. Thus, L2 learners' processing behavior was found to differ more according to level of proficiency than according to L1. The near-natives differed according to their L1 in that only the L1 Russian learners showed native-like use of case information during the speeded grammaticality judgment task and the L1 English learners continued to have difficulty integrating case features and syntactic function. Hopp dubbed this a "delayed L1 effect" that is based on the idea that the increased processing demands of the use of case-marking information in online tasks persist to an advanced level (2006: 290). He surmised that the lack of evidence for online syntactic revision of ambiguous relative clauses by the advanced learners, on the other hand, may be due to use of different strategies, such that the advanced learners resorted to a canonical word order strategy when processing sentences in the L2.

To account for his results, Hopp proposed that the development of L2 processing might follow a trajectory such that less proficient learners might first resort to default processing strategies based on linear order or other non-syntactic information instead of employing the syntactic cues and strategies from their L1's. Once the L2 has been further acquired, L2 learners should be able to tap into the processing routines of their L1's, which might explain the overall differences between near-native L2 learners. Since the experiments in this study were online, perhaps the effects of proficiency were even more robust than in previous studies. Hopp added that the development of L2 processing may not mirror L2 grammar development with respect to transfer effects. He concluded that it should be further investigated whether sensitivity to other types of grammatical information in L2 processing varies according to proficiency level and L1 in the same way as shown in his data. These results are in line with Marinis et al.'s lack of transfer effects in that less proficient learners might first resort to default strategies rather than transfer L1 strategies before they have attained L2 strategies.

1.5.2 Studies Incorporating the *Transfer Principle*

Much of the research reviewed above focused more on the time course of L1 transfer during L2 development, rather than on whether or not transfer was occurring. Not to be left out of the discussion, VanPatten (2007) also addressed the possibility of a transfer account. When subsequent studies such as Gass (1989), testing learners of L1's other than English, began to reveal evidence for the possible transfer of L1 parsing strategies to the L2, VanPatten created the *Transfer Principle*, as shown in (14), to account for the new data.

(14) Learners begin acquisition with L1 parsing procedures. (VanPatten, 2007, p. 122)

If transfer were to be the case, VanPatten explained, then any parsing problems in the L2 would be specific to the learner's L1. For example, an Italian speaker learning Spanish would not have any trouble correctly interpreting OV and OVS structures in Spanish because they are present in Italian, whereas an English speaker would have difficulty because English has a rigid SVO word order, and thus lack a parsing mechanism to handle structures that are not SVO. VanPatten added that a combination of the universal position and a transfer position is possible such that learners start with universal parsing strategies, and then the L1 parser is "triggered" at some point to aid or inhibit comprehension, which would explain the proficiency differences found in studies such as those reviewed in Frenck-Mestre (2002) and the study in Fernandez (1998), and would be in line with the conclusions reached in Hopp (2006). VanPatten argued that evidence from the Competition Model research as well supports his proposed time course in which there is first a default first noun strategy, and later L1 transfer. He concluded that more research is needed to determine how L2 learners use a combination of the two principles, the *First Noun Principle* and the *Transfer Principle*, when processing sentences in their L2.

With the specific aim of testing VanPatten's new transfer principle, Isabelli (2008) examined the results of L1 Italian and L1 English learners of Spanish on an aural sentence interpretation task, in which participants chose between two pictures the one that was best described by the sentence they heard. Isabelli found that the L1 English first-year learners of Spanish had a high error rate on OVS (15a) and OOVS (15b) structures (20% and 5.6%).

(15a) *La* *escucha* *el chico.*
[3^{RD}SINGFEMACC listen to-3^{RD}SINGPRE the boy]
The boy listens to her.

(15b) *A* *la chica* *la* *saluda* *el chico.*
[To the girl 3^{RD}SINGFEMACC greet-3^{RD}SINGPRE the boy]
The boy greets the girl.

However, the L1 English participants' data alone did not provide support for either the transfer or the universal processing strategies position. Thus, Isabelli compared them to the first-year L1 Italian learners of Spanish, who showed a high rate of correct subject identification on the OVS and OOVS structures, which she stated suggested that transfer is at play. Yet, the L1 Italian learners of Spanish did not score as well on the OVS and OOVS structures as on the SVO structures, which may indicate that there was still a preference for a universal first noun strategy for these early learners. In line with VanPatten's Transfer Principle, Isabelli argued that there is a possibility that both universal strategies and transfer are involved, that the L1 parser does not get transferred until the learner has reached a certain level of proficiency. However, Isabelli (2008) did not test her participants for Spanish proficiency, and therefore cannot state unequivocally whether her participants from the two groups were operating at the same level of L2 proficiency during the time of testing. In fact, based on her accuracy results, it appears that the L1 Italian participants were more highly proficient than the L1 English participants, which obviates a clear comparison between the two groups, rendering her results inconclusive in regard to the transfer debate.

While the studies reviewed previously have provided a strong base for a mixed account, there are still many limitations that need to be overcome in future research if a more definitive answer is to be found. For example, the use of offline measures to test processing may not be a reliable way to tap into what is actually happening when L2 learners are faced with constructions that are unique to the L2 (e.g. Cuervo, 2007). Arguably, such tasks are limited in that they only reveal what metalinguistic knowledge the participants have gained, and not their processing abilities, which could be vastly distinct (Hopp, 2006). Moreover, as many of the studies have shown, proficiency appears to be an important variable in how learners process L2 morphosyntax, yet many studies have used self-rated, teacher-rated, or no proficiency measures instead of providing an independent measure of proficiency (e.g. Grüter & Crago, 2010), which could seriously confound their results due to their unreliability. In addition, comparing populations from two different language experience backgrounds such as classroom learners and immersed learners could cause further misinterpretation of results, as in Frenck-Mestre (2002). The inclusion of multiple proficiency level groups within each language group, the individuals of which have been taught the L2 in a similar environment, is a crucial element to include in order to tease apart these variables. In short, taking all of the different terminologies together, the literature shows an apparent consensus that transfer effects are readily evidenced by a variety of tasks, but that proficiency also plays a role in modulating whether these transfer effects are found. Keeping this in mind, the studies reviewed here are designed both to examine the possible effects of transfer driven by distinct L1 learners of Spanish, and for some, also to incorporate the factor of relative proficiency. One other important area of investigation regards not simply the L1 of the learner and his level of proficiency, but rather cognitive individual differences in learners.

The following section reviews the literature on cognitive individual differences and their impact on adult L2 acquisition.

1.6 Cognitive Individual Differences

Researchers in the field of cognitive psychology and more recently in the fields of language acquisition and language processing have identified a construct of memory that appears to differ among individuals. The broad idea of a working memory that is characterized by either a limited capacity or limited resources has been theorized by various scholars, who have subsequently created models of working memory and tests that tap into and measure this construct and its subcomponents. The research reviewed in the following section addresses the configuration and significance of the relationship between one's working memory abilities and one's ability to process language. In particular, working memory is examined as an important factor in determining the success of L2 development. Finally, I include research on the role of an executive attention component of working memory in L2 research.

1.6.1 The Construct of Working Memory

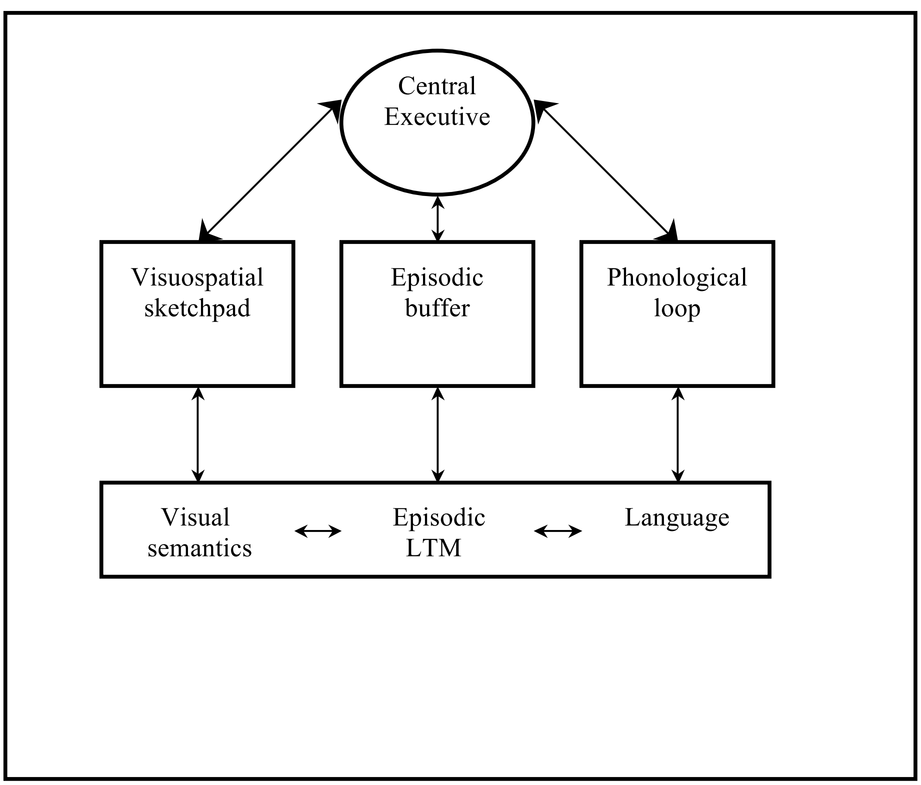
The term "working memory" (WM) has been used to stand for a range of theoretical models, first becoming popular in 1974 with Baddeley and Hitch's adoption of the term to emphasize the differences between their new multicomponent model and previous models of short-term memory. Since then, there have been a variety of theories that have adopted the term. In general, the theories in cognitive psychology have concurred on the definition of WM as a limited capacity system that underlies the storage and processing of task-relevant information during the performance of a cognitive task (Baddeley & Hitch, 1974; Daneman and Carpenter, 1980). Although most researchers agree on its basic definition, they do not necessarily see eye-to-eye on what constitutes its components or how to best test it. This section reveals the general components of the most popular models, showing the points that they have in common that are most relevant to the purposes of this dissertation. I outline the history of the most prominent models of working memory (WM) and discuss the tests used to measure WM. In addition, I present multiple studies that examine the role of WM in L2 learning, focusing on adult processing of L2 morphosyntax. Finally, I review the literature on the cognitive demands of processing OVS structures as opposed to SVO structures, and how individual differences in working memory capacity have highlighted this discrepancy.

1.6.1.1 Working Memory Models

1.6.1.1.1 Baddeley and Hitch's Model

Baddeley and Hitch (1974) characterized working memory as a three-component system that included a 'central executive', which controlled attention, and two domain-specific 'slave' systems, the 'articulatory loop', which later was named the 'phonological loop', and the 'visuo-spatial scratch-pad', which was later named the 'visuospatial sketchpad', as recreated in Figure 1 from Baddeley (2000).

Figure 1. *Baddeley's Model of Working Memory*



Each of the slave systems held information relevant to its domain such that the phonological loop held auditory information and the visuospatial sketchpad held visual, spatial, and possibly kinesthetic information. Baddeley and Hitch suggested that the core of WM consisted of a ‘work space’ that has limited capacity, and can be divided to some degree into storage and control processing components. They included a caveat that it is unclear if allocation of capacity is flexible or if there are set parts designated to each component. They demonstrated in their experiments that concurrent memory load had very little impact on reasoning, sentence comprehension and learning, and therefore, provided evidence for these separate components. Baddeley and Hitch hypothesized that there may be one central executive, which they described as a complex component that is responsible for chunking material and interpreting phonemic traces in the verbal modality, shared by comparable systems (such as a possibly separate visual memory system) or there may be multiple central processors, one for each comparable system. In all, Baddeley and Hitch posited an account that replaced the Short Term Storage-Long Term Storage viewpoint, and encouraged the acceptance of a multicomponent model.

Baddeley and Hitch’s model of WM in one form or another continued to be used as one of the prominent models of memory in the cognitive sciences for twenty-five years. In Baddeley (2000), after summarizing the literature that showed phenomena not explained by the 1974 model and its various forms, the author proposed an additional component to account for these phenomena, the episodic buffer. Baddeley suggested this fourth component, the episodic buffer, in order to satisfy the need illuminated by the data from patients with short-term memory deficits for a mechanism that stores serial recall and possibly integrates phonological, visual, and other types of information.

Baddeley defined the episodic buffer as a short-term limited-capacity storage system controlled by the central executive that had the ability to hold episodes and integrate information across space and time from a variety of sources. Its addition reinforced the idea of a multi-component model in light of other models of WM that had been proposed in the field.

1.6.1.1.2 Unitary Models

Beyond the work of Baddeley and colleagues, there was another line of research that followed the idea that WM was a subset of long-term memory. Agreeing with this line of researchers that believed WM was an essentially unitary mechanism, Schneider and Detweiler (1987) proposed a WM architecture within a connectionist framework. Although Schneider and Detweiler accepted and incorporated the idea of bufferlike processes from the modal model, the basic structure that they described assumed that processing occurred in a set of modules that were organized into levels and regions that communicated with each other on an innerloop of connections. Information transfer occurred on this loop and was managed by a control processing system in their model. Accepting the definition of WM that Baddeley posited (Baddeley, 1986), Schneider and Detweiler stressed that their model better accounted for phenomena such as how information was recovered after an interruption "flushes the buffers" (1987:112). They noted that their model, as opposed to previous models, accounted for numerous WM phenomena as emergent properties of the system. As an individual learned a particular item, the item's connection weights would change. They depicted their model in three levels: micro-, macro-, and system-level.

Also promoting the idea of WM as a subset of long-term memory, Cowan (1988) proposed an "embedded processes" model. Not wanting to lose the temporal aspects of processing, Cowan eschewed a schematic, linear model in favor of a more functional approach. His model showed the time since the reception of the stimulus as an arrow along the x-axis, as in Figure 2.

Figure 2. *The Cowan (1988) Embedded Processes Model of Working Memory*

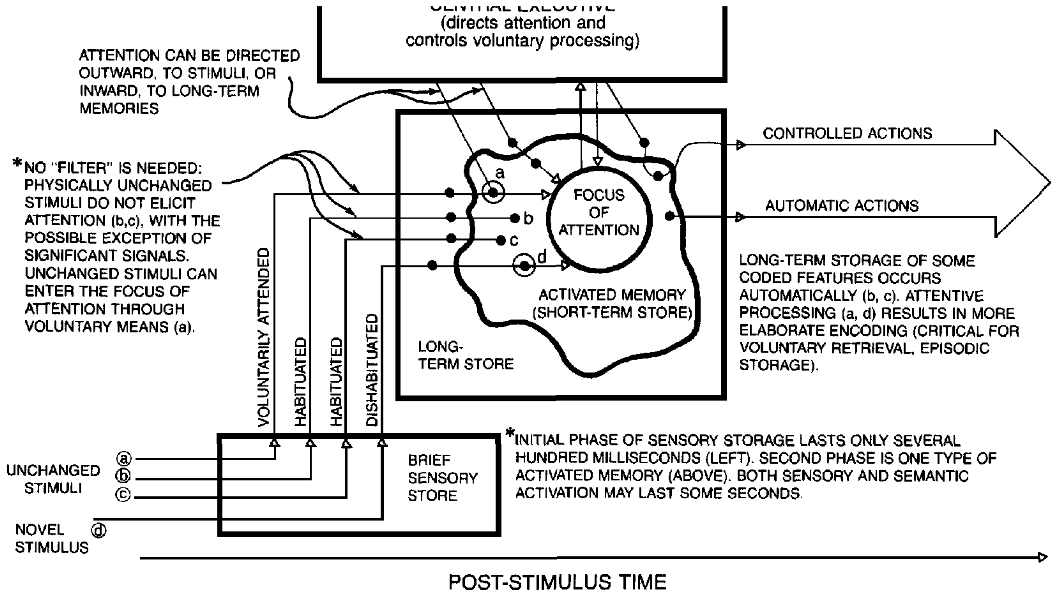


Figure 1. A revised model of the information-processing system. The time since stimulus reception is represented ordinally along the *x* axis. The components are arranged in real time, and stimulus information can be present in more than one component at the same time. Short-term storage is represented as an activated

Cowan arranged the components of what he called the information-processing system in real time, depicting how stimulus information may be found in more than one component at a given time. He denoted short-term storage as a subset of long-term storage that is activated, and the focus of attention as a subset of the activated short-term storage component. The involvement of the central executive over the time course of processing was flexible, according to this model, but a crucial component. In addition, Cowan (1995), as cited in Cowan (1999), found biological underpinnings for the majority of the model's components such as evidence that the inferior parietal areas appeared to represent the focus of attention.

In sum, although there are some specific differences in the implementation of WM models, researchers generally accept that there is a working memory system, that it serves a buffering purpose, and regarding this dissertation, that there is an important role to be played by the central executive component. After a discussion of whether individual differences in WM stem from its capacity, its efficiency, or the ability to focus the attention of the central executive, I address WM and L2 processing in greater detail in section 1.6.1.3.

1.6.1.1.3 Individual Differences in WM

Although the majority of researchers working along these lines seemed to agree that WM was a subset of long term memory, in that it was the activated portion of memory and was not structurally separable, there was still disagreement about whether it was WM capacity itself or the efficiency of processing that resulted in individual differences in WM span. Daneman and Carpenter (1980), which I discuss in greater detail in the following section, set the stage for the debate when they developed a reading span measure that correlated well with reading comprehension performance unlike the previous measures thus far used in the field (e.g. traditional digit span and word span). In response to Daneman and colleagues' claim about the specificity of the reading span task and how it can predict reading comprehension ability based on a fixed-capacity model of WM (Daneman & Carpenter, 1980; Daneman and Green, 1986), Turner and Engle (1986, 1989) tested participants on four WM span measures and found that those measures other than reading that were embedded in a processing task (i.e. were more complex than the simple digit and word span tasks) correlated just as well as the reading span with reading comprehension. Since the presence of a background task in these complex span tasks made it more difficult for participants to use rehearsal strategies to avoid capacity limitations, as they had been shown to do during simple memory-span tasks, Turner and Engle argued that their results were a better indicator of the participants' 'true' capacity. More interesting, WM appeared to have a general storage capacity that was independent of the task that was being used, although the authors failed to mention the possibility that linguistic skills may have been used in the processing of arithmetic equations (Richardson, 1996). They concluded that it was not processing efficiency that allowed for a higher span score but a larger WM capacity, which varied with each individual.

On the other hand, Just and Carpenter (1992) proposed a capacity theory of a different kind. Basing their work on the Just and Carpenter (1980) process model of reading comprehension, which included a WM component, Just and Carpenter's (1992) theory focused on a WM for language that roughly corresponded to the part of the central executive that dealt with language comprehension in Baddeley's (1986) theory. Thus, there were no modality-specific buffers in Just and Carpenter (1992). They expressed that total WM capacity was determined by the maximum amount of activation available in WM to support storage and processing. Each element, in this theory, had an activation level associated with it. In this way, the theory more closely resembled a connectionist view of WM. Just and Carpenter also viewed individual differences as deriving from the amount of activation that individuals have available for computation and storage during language processing. Individuals differ, according to this model, in their speed and accuracy, and thus, in their efficiency, in comprehending language. In order to lend support to their theory, Just and Carpenter provided a computer simulation model of WM that demonstrated how the same comprehension strategy produced different types of performance based on the resources that were available to the system. However, it is important to note that since Just and Carpenter could not claim that the WM capacity used for language comprehension was the sole cognitive capacity, they challenged future researchers to show that this version of a capacity theory functions within a domain other than language.

Arguing that the empirical evidence provided as support of the capacity model in Just and Carpenter (1992) was unconvincing, Waters and Caplan (1996a) critiqued the use of Daneman and Carpenter's reading span measure based on the grounds that it required conscious retrieval of items held in memory, which is more demanding than the practice of processing sentence structure while reading. Waters and Caplan provided an alternative theory that stated that there are specializations within the verbal-processing resource system for different verbally mediated tasks. They posited two resource pools, one for on-line psycholinguistic operations such as syntactic parsing, and one for controlled, verbally mediated tasks. They named their theory the theory of separate language-processing resources. Their predictions based on this theory included that performance on general verbal working-memory tasks would not predict language-processing efficiency, that comprehension-external and comprehension-internal factors draw on different resource pools (predicting main effects of each but no interference), and that low-span participants should perform less well under conditions of high memory load, but that this effect is not greater for syntactically complex sentences. They concluded that the research has favored a model in which the processing resources used in unconscious, on-line language comprehension are partially separate from those used in controlled, verbally mediated functions.

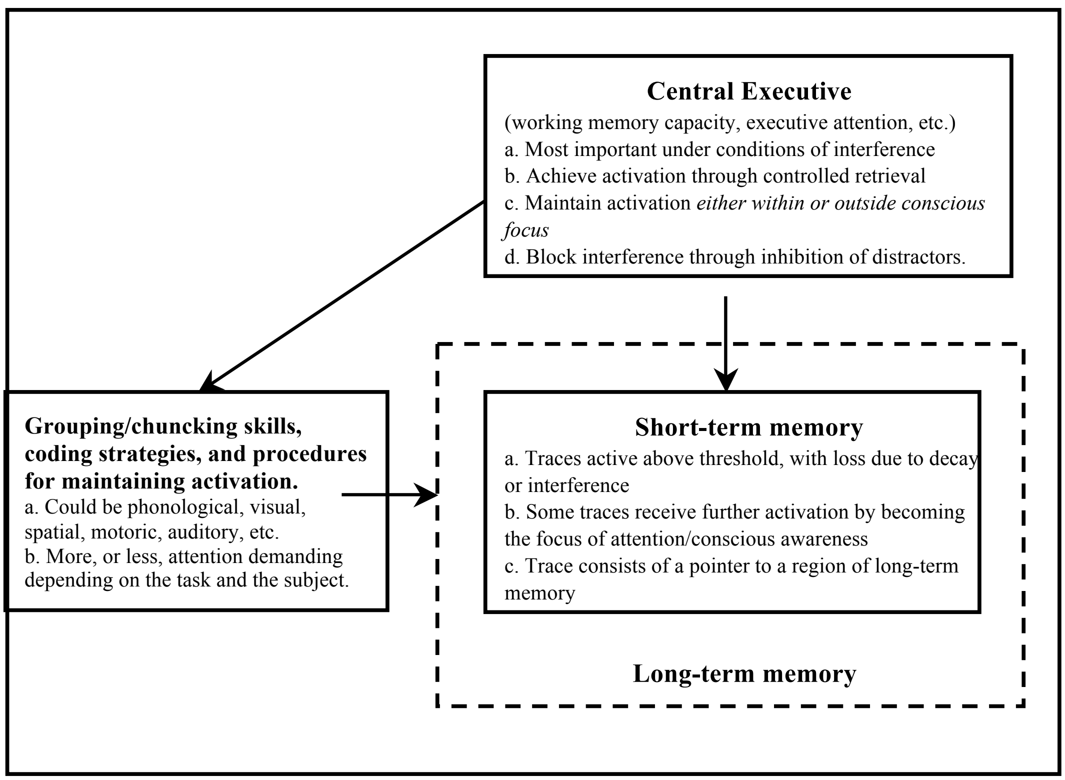
Building on Turner and Engle's work, the general capacity theory of WM was further developed in Engle, Cantor and Carullo (1992). Stating that their ideas about the structure of WM paralleled theories proposed by Schneider and Detweiler (1987), Cowan (1988), and Just and Carpenter (1992), among others, Engle et al. assumed that knowledge units in the memory system vary in their activation levels and that WM consisted of those knowledge units that had recently been activated. In addition, activation differences create individual differences, according to this model. Instead of providing support for a domain-specific system like in Just and Carpenter (1992) or for a task-specific processing hypothesis like in Daneman and Carpenter (1980), however, Engle et al. reported evidence for a domain-general WM storage capacity that was independent of task. They found that the high and low span participants did not differ on their processing time of the elements of the task except that high span participants spent more time on the words that had to be recalled. As in Turner and Engle (1989), they also used complex span tasks to eliminate the possibility of the participants using rehearsal strategies, and reported that the results from the span tasks correlated with the participants' verbal SAT scores, thus supporting the general capacity model of WM.

Caplan and Waters (1999), building on their work from Waters and Caplan (1996b) and Caplan and Waters (1996), focused their study on the finer-grain distinction between proposed subsystems of verbal WM. Depending on the type of verbal task, they proposed a division between "interpretive processing", which is the extraction of meaning from a linguistic signal such as recognizing words and constructing syntactic representations, and "post-interpretive processing", which is the use of that meaning to accomplish other tasks. Comparing the SR (single-resource) approach of researchers such as Just and Carpenter to their theory of a separate-sentence-interpretation-resource (SSIR), Caplan and Waters presented evidence of no significant differences in effect size between high- and low-span participants' on-line processing of object-relativized and subject-relativized sentences. They cited their work from Caplan and Waters (1995), in which they found that participants with differences in WM capacities did not perform differently as a function of the syntactic complexity of the stimuli. In addition, they cited data on aging (e.g. Zurif, Swinney, Prather, Wingfield, and Brownell, 1995) that indicated that elderly participants had no problems with post-interpretive processing, yet showed interpretive processing declines, providing indirect support to Caplan and Waters' hypothesis of the division of verbal WM into subsystems. Finally, they summarized the results from studies with aphasic patients as being consistent with their theory since they found that the resources used in syntactic processing were not reduced in patients with reduced WM capacity. Proposing a new hypothesis, the separate language interpretation resource hypothesis (SLIR), Caplan and Waters (1999) concluded that the proposed separate resource in WM allocated to syntactic processing may also serve other integrated operations involved in computing a coherent discourse meaning such as determination of topic, focus, and causality, among others. In order to test this hypothesis, they suggested studies that include tasks that emphasize controlled and conscious manipulation of verbal information coupled with tasks of processing efficiency, and they predicted that there would not be any correlations between the results of these two types of tasks.

Caplan and Waters' view that individual differences in WM capacity are based on the ability to retain in memory information about the propositional content of a sentence and not in speed or accuracy of syntactic processing was contested by MacDonald and Christiansen (2002), among many others, who responded with a commentary on the usefulness of the traditional concept of WM. They claimed that all of the tasks that have been included under the headings of language-processing tasks and verbal WM tasks are simply different measures of the same language processing skill. They argued that individual differences are due to variations in language exposure (such as accumulation of time spent reading) and to biological differences (such as differences in phonological representation precision). Instead of basing their views on a “symbolic processing architecture” like Just and Carpenter and Waters and Caplan, MacDonald and Christiansen made their claims based on a connectionist approach to language comprehension. They stated that although Just and Carpenter and Waters and Caplan strongly disagreed on how WM is structured, the four researchers concurred that there is a crucial separation between linguistic knowledge and WM, which is MacDonald and Christiansen’s biggest point of contention. MacDonald and Christiansen argued that these are inseparable, emerging from the interaction between the network architecture and the individual’s experience. Their model depicted the processing of input as the passing of activation through a multilayer network, which is essentially WM. According to their model, the ability of the network to process information varied based on the input, the properties of the network itself, and the interaction of these properties. Individual differences arise in this model as variations in the amount of training, variations in the efficiency of the network in passing information, and variations in the amount of units that the network has available to learn and process information. In short, MacDonald and Christiansen maintained that their view accommodates both the domain specificity and the biological data that Caplan and Waters asserted were central to a model of WM without the inclusion of multiple separate working memories.

Refining the idea that activation differences in task-relevant knowledge create individual differences, Engle (2002) stated that WM capacity is not directly about memory but about the focusing of attention on task-relevant knowledge. Engle wrote that although greater WM capacity means that there is a greater number of items that can be maintained active, this is due to a greater ability to control attention, not to a larger memory store. Citing work done by Kane and Engle (2000), he posited that inhibiting distraction and dealing with proactive interference, which he defined as the difficulty experienced when a new behavior is associated with a context associated with other behaviors, is within the domain of WM. Results from Kane and Engle (2000) showed that low WM capacity participants showed a greater loss of recall of words from word lists than high WM span participants as the number of trials (and thus the number of word lists) increased. That is, as participants experienced greater proactive interference, differences in WM ability were reflected by the differences in their ability to recall the words. Engle and Kane (2004) and Kane, Conway, Hambrick, and Engle (2007) further developed these ideas in their measurement model of WM, version 1.2 of which is adapted and reproduced in Figure 3.

Figure 3. *Kane et al. (2007) Measurement Model of WM*



Engle (2002) concluded that based on his research and that of others in the field, the tasks used to measure WM capacity are in fact, measuring a construct that is fundamental to higher-order cognition.

I have given this review of the diverse WM models and their interpretations to show the complexity of the issue at hand. Although a large portion of the field continues to incorporate a Baddeley and Hitch-type model in their work with the idea of domain-specific parts of WM, there is a whole other vibrant line of research that follows the idea that WM is a subset of activated long-term memory, such as depicted in the models by Schneider and Detweiler (1987), Cowan (1988), and Engle and colleagues. Furthermore, there is a continuing debate between followers of a unitary model about whether individual differences in WM are based on differences in capacity or in efficiency of processing, which leads to increased capacity, as seen in Just and Carpenter (1992), and contested in Waters and Caplan (1996a) and Caplan and Waters (1999), as well as in MacDonald and Christiansen (2002). The idea that it is activation levels of knowledge units and the ability to shift and focus attention on the relevant aspects of a task has also become popular with Engle and colleagues' work. For the purposes of this dissertation, I adopt a domain-general framework of working memory and the terminology proposed by Engle and colleagues in regards to the executive attention component of WM (the relevance of which I elaborate upon in section 1.6.2), with the understanding that regardless of the model that a researcher ultimately chooses there is widespread agreement about the importance of WM capacity as an individual difference in language learning. In the following section, I discuss the various tests used to assess WM capacity and how they tap into WM capacity (or efficiency or executive attention) differences that can be used to help account for individual variability in adult L2 acquisition.

1.6.1.2 Working Memory Tests

It is crucial to examine the various tests of WM, in hopes of shedding light on how best to assess the WM constructs most applicable to language learning, the focus of this dissertation. Prior to the cognitive revolution and proliferation of research on WM, there were various tests used to assess memory capacity. These tests typically were span-type tests such as digit span or word span tests. The digit span test requires participants to remember a sequence of digits from one to nine. Participants hear a sequence of digits and must recall them exactly as they were presented. The task begins with sets of two digits and proceeds with sets of increasing numbers of digits until the participant cannot correctly recall the digits at least three times out of ten for a given set size. According to Miller's (1955) seminal paper, which based its assumptions on this type of task, human short-term memory capacity is limited to seven items, plus or minus two. Another traditional measure of short-term memory that has had similar results is the word span task, which involves the participant recalling sets of individual words. The words are typically grouped into sets from two to seven words, and within each set the words have to be recalled in order of presentation. Again, the participants' span is based on the number of words that they can consistently recall at one time.

Tasks such as these span tests were criticized by Daneman and Carpenter (1980) for their use as measures of WM in linguistic studies based on evidence that they only correlated weakly or not at all with reading ability. Based on Daneman and Carpenter's perspective that within WM processing and storage competed for a shared limited capacity, they stated that digit span and word span tests only taxed the storage component and not the processing component of WM. To rectify the situation, they proposed a new measure that would better tax the processing component. They predicted that this measure would correlate with reading comprehension performance. Their reading span test required participants to read a set of sentences aloud and attempt to recall the final word of each sentence at the end of the set. The number of sentences in a set increased from trial to trial until the participant could not maintain perfect recall of the final words. The maximum number of sentence-final words that the participant could read would be his reading span. Daneman and Carpenter considered this reading span an index of the participants' WM capacity (keeping in mind that Daneman and Carpenter defined WM as specific to reading comprehension). Then they tested their measure to see if it correlated with reading comprehension measures such as the verbal SAT. Their results showed that the reading span test highly correlated with the traditional reading comprehension assessments, whereas it did not correlate as strongly with the traditional word span measure. The authors concluded that their reading span task was successful in taxing both the processing and the storage components of WM.

While many researchers have used the Daneman and Carpenter reading span task, it is not the only task of this type. Turner and Engle (1989) created a different version of the reading span task. In their version, there were fewer items overall and the participants had to respond to whether or not the sentences were semantically or syntactically correct. Turner and Engle presented the stimuli both auditorially and visually while participants read them aloud. Other versions of the reading span task have variations such as the sentences may be simpler syntactically with a grammaticality judgment after each one (see Harrington and Sawyer, 1992), the to-be-remembered word may be different from the last word, or any word, in the sentences (see Engle, Tuholski, Laughlin, and Conway, 1999), or participants may be required to remember isolated letters following each sentence (see Kane, Hambrick, Tuholski, Wilhelm, Payne, and Engle, 2004). While they still agreed that this task provided a measure of some kind of WM, Waters and Caplan (1996b) found that the test re-test reliability of the Daneman and Carpenter test was very low and proposed that the measure include a plausibility judgment following each sentence rather than requiring the participants to read aloud the sentences to ensure their processing of them. MacDonald and Christiansen (2002), based on their connectionist approach to WM, viewed the reading span task as a measure of language processing skills in general and not of some unitary construct.

The operation span and counting span, as documented in Engle (2002), are two other popularly used measures of WM. First used for language experiments in Turner and Engle (1989), the operation span, in particular, resembles the reading span tasks of Daneman and Carpenter (1980) and Turner and Engle (1989) in that participants read aloud sets of arithmetic operations such as, “is 4/2 + 3 = 6?” and must first say if the operation is correct or not, and then read aloud a word at the end of the operation. These operation-word strings are grouped into sets of two to seven. After the set, the participant must recall as many of the words as possible. The number of words recalled is the operation-span score. The third span test of this type is the counting span task (e.g. Engle, Tuholski, et al., 1999), in which participants must count the number of target items on a card while ignoring distractor items. Then they must recall in order the digits that correspond to the numbers of target items on the cards. The score is the number of digits recalled correctly.

Conway, Kane, Bunting, Hambrick, Wilhelm and Engle (2005) analyzed these three widely used measures (the counting span, the operation span, and the reading span tasks) in their user’s guide, warning that the construct of WM had been successfully translated into other disciplines, whereas the tasks may not have been. They provided a thorough review of the tasks that included how best to administer and score them as well as an assessment of their reliability and validity. Conway et al. first clarified that their view of WM was domain-general, and that their review would incorporate the idea that the tasks tap into complex cognitive behavior across domains. They confirmed that all three tasks share a common structure in that they force WM storage during processing to engage executive attention processes. In addition, they asserted that regardless of what WM span tasks are believed to be measuring, they are measuring what they are actually measuring very reliably in that statistical tests of reliability such as coefficient alphas and split-half correlations have confirmed that they are reliable across participants and over time in test-retest situations. Although they cannot confirm what the span tasks measure based on reliability tests alone, Conway et al. showed that the span scores are being influenced by a stable construct. To determine what this construct was exactly, they considered numerous correlational studies and found that WM span performance correlated with a variety of cognitive tasks (e.g. reading and listening comprehension (Daneman and Carpenter, 1980), language comprehension (King and Just, 1991), vocabulary learning (Daneman & Green, 1986), reasoning (Kyllonen and Christal, 1990), and complex-task learning (Kyllonen and Stephens, 1990). Insofar as WM involves control of attention and thought, these measures of WM capacity showed both convergent and discriminant construct validity, according to Conway et al. They also stated that basic short-term memory span tasks correlated slightly with WM tasks, in that they both test the construct of storage, but that the inclusion of a secondary task that competes with the storage of information is necessary to more precisely measure WM capacity. Conway et al. also reviewed other tasks such as the running span (participants are presented with a list of stimuli of unknown length and must recall only the last *n* items), *n*-back tasks (participants must continuously state whether an item matches the one that they saw *n* times previously), and the alpha span task (participants must recall target words in alphabetic order), concluding that these tasks are more dynamic and lack sufficient research as to their validity and reliability. They ended by highlighting that a particular researcher's goals must drive the decision on which tasks to use to measure WM capacity. For example, if a researcher wishes to test hypotheses about storage, then she should focus on WM tasks that tap into storage and pay little attention to those that assess attentional aspects of WM.

An important inference that can be drawn from Conway et al.'s paper that is of particular relevance for this dissertation is that if a researcher is dealing with multiple groups that speak various languages, it may behoove her to include WM tasks that allow participants to use their first language (e.g. tasks that do not include full words or sentences) instead of creating versions in each of the different languages. One such test would be the number-letter sequencing (also called the letter-number sequencing or LNS) subtest from the revised version of the Wechsler Adult Intelligence Scale (WAIS), which measures WM by presenting the participant with a mixed series of letters and numbers and asking him to recall them, numbers first in numerical order, then letters in alphabetical order (Wechsler, 1997). This task's first reported use was in the study conducted by Gold, Carpenter, Randolph, Goldberg, and Weinberger (1997), who examined auditory WM and executive processes in schizophrenia. In addition, Haut, Kuwabara, Leach, and Arias (2000) found evidence that an auditory version of the LNS task involved verbal WM abilities through the use of [O^{15}] water PET methodology. They saw activation during the LNS task that was typical for WM tasks: right-sided activation in the posterior parietal cortex (temporary storage of information in verbal and visual format), the dorsolateral prefrontal cortex (organization and maintenance of temporarily stored information), and the premotor cortex (storage of visual information and subvocal rehearsal). Finally, Crowe (2000) found significant correlations between reading ability, digit spans and spatial tasks, and the letter-number sequencing task in his data. Thus, he concluded that this task assesses auditory (verbal) and possibly spatial WM.

Research in L2 learning and WM has employed the various WM tests outlined in this section. Osaka and Osaka's findings (1992) that WM capacity may be language independent, which resulted from their creation and testing of a reading span task equivalent to Daneman and Carpenter (1980) in Japanese, has caused some researchers to ask participants to take the tests in both their L1 and their L2, whereas others, when feasible, have found it sufficient to test participants' WM solely in their L1. One of the dangers of testing WM in the L2 is that proficiency could be potentially confounded with WM. When taking certain tests of WM, if administered in the L2, a sufficient level of proficiency is required for a participant to complete the task. Otherwise, the construct of WM will not be tapped into, and instead of a test of WM the test will serve as a test of the participant's L2 proficiency level. In short, a researcher's choice of WM measure or measures should be justifiable based on the aspects she wishes to test and the resources available to her as a researcher. For this dissertation, I incorporate a letter-number sequencing task to test language learners' processing and storage abilities, which is a relatively language independent task. Although participants may encode the numbers and letters in language for purposes of storage, and differences in length and number of syllables may exist among the various languages involved, this task is less language dependent than a reading span or operation span task, which involve words and full sentences. Since the participants in the experiments in the present dissertation hail from varying language backgrounds, one of which has no known versions of the more commonly used WM reading span or operation span tests, the choice of a numeral and letter-based task is preferable in that it is at least as neutral to language dependence as it is feasible.

1.6.1.3 Working Memory and L2 Processing

1.6.1.3.1 Introduction

Thus far, I have reviewed the models of processing strategies, both universal and those that have incorporated transfer, and I have reviewed the concept of WM and the tests that have been employed. At this point, what remains is to put WM into the context of L2 learning in order to understand the design and goals of the studies reported here, which examine both the question of transfer and the potential contributions of individual differences in adult L2 acquisition. This section reviews the literature on WM and L2 processing.

In general, researchers have sought to identify why there are such great individual differences in L2 attainment. Research in the last twenty years has aimed to decipher if and to what extent individual differences in WM abilities underlie the differences found in second language acquisition. Paralleling the WM and L1 studies, many of the L2 studies that have found positive correlations between WM and L2 performance have focused on reading comprehension and processing. In this section, I discuss studies that look at the influence of WM as an individual difference factor in L2 processing of argument structure. First, I summarize the literature on WM and processing of the L2, in particular L2 processing of morphology, morphosyntax, and syntax. Next, I synthesize the evidence presented for the relationship between WM and L2 processing. Then, I summarize the research on processing object-first constructions vs. subject-first constructions in the L2 and in relation to the construct of WM.

1.6.1.3.2 Studies of WM and L2 Lexical Processing

Studies of L2 processing have shown correlations between WM and L2 abilities, while being relatively consistent in their use of measures of WM. At the word level, Kroll, Michael, Tokowicz, and Dufour (2002) found correlations between WM capacity and the translation of cognates. As part of a larger study, Kroll et al. tested a group of L2 learners in their L1 English using a version of the reading span task from Waters and Caplan (1996c), and divided them into a high span group and a low span group. Kroll et al. also tested the participants on reading and translating words from their L1 to their L2 and vice versa, and found a correlation between their WM span and the magnitude of a cognate advantage (i.e. lower response times for words that are cognates in both the L1 and the L2). In particular, the low span participants showed a larger cognate advantage than the high span participants, while the reverse was true for noncognates. The authors stated that this suggested that high span learners are less likely to rely on cues from the forms of the words than low span learners. In other words, the high span learners might focus their mental resources on generating strategies to improve processing, but sometimes these strategies incur a processing cost.

In addition, Tokowicz, Michael, and Kroll (2004) found effects for WM and single-word translation performance for L1 English-L2 Spanish and L1 Spanish-L2 English learners in a study abroad context. The participants with higher WM capacity made proportionally more meaning errors to non-response errors than those participants with lower WM capacity with the same amount of study abroad experience because they had more capacity or better allocation of resources to be able to use communicative strategies (such as naming something rather than remaining quiet), which may lead to more success in a study-abroad context. Suggesting that learners who make more response errors than non-response errors would be better understood by their interlocutors, the data from Tokowicz et al. showed support for the advantages of higher WM capacity for language learning and for the importance of including this variable in L2 research.

1.6.1.3.3 Studies of WM and Processing of L2 Morphology and Morphosyntax

1.6.1.3.3.1 Visual Processing

In a study focusing on visual processing of morphological features, Sagarra (2007a) tested L1 English low proficiency learners of Spanish on their reading ability of Spanish sentences with noun and adjective combinations that showed morphological agreement or disagreement and on their accuracy on comprehension questions within a moving window paradigm. Sagarra found that those participants with high WM, as measured by the Waters and Caplan (1996b) reading span test, were more sensitive to gender agreement violations than those with low WM. This suggested that learners with higher WM capacity may be able to acquire a second language more rapidly and accurately than learners with lower WM capacity.

In the same vein, Sagarra and Herschensohn (2010) sought to find out whether individual differences in not only WM but also in proficiency level affected the online and offline processing of L2 gender and number agreement with both animate and inanimate nouns. Sagarra and Herschensohn (2010) also employed the Waters and Caplan (1996b) reading span test as their measure of WM, which was given in the participants' L1 based on the view that WM is language-independent (Osaka & Osaka, 1992) and that L2 proficiency could influence the results of a test in the L2. The correlation between accuracy on the grammaticality judgment items with gender agreement violations and WM span was positive, showing that intermediate learners with higher WM capacity were more sensitive to gender disagreement than those with lower WM capacity, while the beginning learners showed no significant effects. The authors argued that this might have been due to the overall low WM mean of the beginning learner group. In conclusion, Sagarra and Herschensohn found evidence that higher WM capacity can facilitate acquiring sensitivity to L2 morphology, specifically to morphosyntax that is cognitively taxing such as gender agreement.

In a separate study, Sagarra (2008) investigated L2 learners of low proficiency on their processing of redundant grammatical forms in visual stimuli using a moving-window paradigm. For example, participants' processing of sentences of the form, *Ayer el estudiante miró una película de terror en el cine* (Yesterday the student watched a horror movie at the cinema) and ungrammatical sentences of the form, **Ayer el estudiante mira una película de terror en el cine* (Yesterday the student watches a horror film at the cinema) was compared to their WM capacity to see if WM capacity acted as a mediator. Sagarra found that when she divided the participants into high and low WM capacity groups, as assessed by the Waters and Caplan (1996b) reading span task, the high span group processed the redundant grammatical forms more often than the low span group. These results suggest that even at a low proficiency level, WM capacity mediates morphological processing in the L2. These studies of visual processing of the L2 at the word level and of L2 morphology have all used the same standard measure of WM capacity. Thus, their findings can arguably be compared, resulting in strong support for a positive relationship between WM capacity and L2 visual processing ability.

1.6.1.3.3.2 Aural Processing

In addition to studies of visual processing, there have been a few studies looking at aural processing of L2 morphology and WM. Ando, Fukunaga, Kurachachi, Suto, Nakano, and Kage (1992) tested L1 Japanese fifth graders' WM capacity on both L1 and L2 reading and listening span tasks. Then, after nine hours of form-focused aural instruction in L2 English, Ando et al. asked participants to complete a delayed posttest, and found that those participants who were most successful in L2 learning as assessed by their posttest scores were those with higher WM capacity, suggesting that those with higher WM capacity could better integrate the input presented during the instruction into their developing L2 systems. Hence, differences in WM capacity correlated with differences in L2 acquisition, according to the data from this study.

Along similar lines, Mackey, Philp, Egi, Fujii, and Tatsumi (2002) examined the effect of learners' WM capacity on their noticing of interactional feedback, and subsequently on their L2 learning. Mackey et al. investigated adult Japanese learners of English, testing them both in their L1 and L2 on WM measures. The authors found significant benefits from instruction and feedback on delayed posttests for learners with high WM capacity as compared to learners with low WM capacity. The authors stated that this indicated that the participants with higher WM capacity had made stronger cognitive connections between the form and meaning of the input, and thus, were better able to retain these connections over time as opposed to the participants with lower WM capacity. These results strongly suggested that there are significant consequences in L2 learning in the aural mode as well based on WM capacity.

Also working with feedback processing, Sagarra (2007b) found that WM capacity, as assessed by the Waters and Caplan (1996b) reading span test, was a predictor of linguistic accuracy on written posttests and of the amount of modified output for L1 English learners of Spanish. These results corroborated the findings from Mackey et al. (2002) in that the participants with higher WM capacity who had received feedback during a sentence completion task were better able to incorporate the feedback into their developing L2 system than those with lower WM capacity. Thus, studies of aural as well as visual processing of L2 morphology and morphosyntax have shown positive correlations between WM capacity and L2 learning.

1.6.1.3.4 Studies of WM and Processing of L2 Syntax

In addition to studies on WM and processing of L2 morphology, there have been various studies that have considered WM and its effects on processing of L2 syntax. For example, adopting the Just and Carpenter (1992) view of WM, Miyake and Friedman (1998) focused on WM for language and its limited supply of resources. In the study they reported on, Miyake and Friedman used the framework of the *Competition Model* (MacWhinney and Bates, 1989), as reviewed above. For example, in English, native speakers rely heavily on word order to determine the subject or agent of a sentence, whereas in Japanese, which has a more flexible word order, native speakers rely more on case markings to determine who did what to whom. Other studies have shown that Japanese learners of English have difficulties in adjusting their cue preferences as a population (Harrington, 1987; Kilborn and Ito, 1989), yet differ individually in their cue preferences in that some rely more on a word order cue (even though it is less valid in Japanese) than others. Suggesting that these differences may be due to WM, Miyake and Friedman conducted correlational analyses on the results from L1 Japanese L2 English learners' listening span and digit span tests both in Japanese and in English, from syntactic comprehension tests, and from cue preference distances (calculated by how far from a native English speaker's cue preferences the learners' cue preferences were). All but the L1 or L2 digit spans and cue preference distance variables correlated with each other. In particular, they found evidence that L1 WM and L2 WM may share the same resources. They also found a significant contribution of L2 WM on L2 syntactic comprehension and a strong contribution of L2 learner cue preference distance and how accurately the participant interpreted different sentence structures. The fact that syntactic comprehension and the ability to shift cue preferences to the non-native preferences in the L2 were positively correlated indicated that individual differences in cue preferences were related to WM capacity. Since word order is a global cue that places a high demand on WM, whereas other cues such as case-marking are locally processed, Miyake and Friedman maintained that those participants with lower WM capacity had more difficulty employing a word order strategy in the L2 to determine the agent or subject of the sentence than those with higher WM capacity. This seminal work in second language acquisition highlighted the important role of WM in L2 learning, and specifically in the processing of L2 syntax.

Another seminal work, Juffs (2004), found differing results. Juffs reported that based on his data, the central executive component of WM, as assessed by the reading span test (Daneman & Carpenter, 1980; Harrington & Sawyer, 1992), is not a source of individual variation in L2 online performance. Juffs tested and compared L1 Chinese, Japanese, and Spanish learners of English and L1 speakers of English on an adapted version of the reading span task both in their L1 and L2 (for the learners), on word span tasks both in their L1 and L2 (for the learners), and on a moving window grammaticality judgment task involving complex variations of garden-path sentences. Unlike some previous studies, he found correlations between the word span and reading span scores, which are often believed to test different aspects of WM. The strongest and most reliable correlation was between L1 and L2 reading span scores. More important, the only differences he found between the native and non-native speakers in online processing of the stimuli were that the non-natives were overall slower readers in their L2 and less accurate than the natives in their grammaticality judgments, as to be expected. It is interesting to note that the non-native speakers performed qualitatively the same as the native speakers in their processing of the various components of the garden path sentences, and that there was no correlation found between any of the WM measures and the scores at the point in the garden path sentences where the processing load was the greatest. Juffs concluded that his data were consistent with the Waters and Caplan view of a global WM that has independent memory capacity for online syntactic processing, thus explaining the lack of correlations found between WM capacity and L2 processing of syntactically ambiguous sentences.

In a subsequent study on adult learners of English from various language backgrounds (L1 Spanish, L1 Chinese, and L1 Japanese), Juffs (2005) examined whether WM measures predicted participants' performance on the processing of subject and object long distance extractions from finite and nonfinite clauses. He gave participants both a reading span test in their respective L1's and in their L2 English (Daneman & Carpenter, 1980), and two word span tests (one in the L1 and one in the L2) as measures of WM. Juffs predicted that participants may pattern differently according to the basic word orders of their L1's (Spanish and Chinese are mainly SVO, whereas Japanese is SOV). Participants were presented with stimuli based on Juffs and Harrington (1995), which included sentences with object and subject extractions with finite clauses (e.g. "Who did the woman suggest the manager liked at the office?" and, "Who did the woman suggest liked the manager at the office?" respectively), and in nonfinite clauses (e.g. "Who did the manager expect to hire for the job last week?" and, "Who did the manager expect to meet the customer last month?" respectively). Based on what they read, participants were asked to make a yes or no grammaticality judgment. Juffs found a reliable relationship between participants' scores on the word span and the reading span tasks in both their L1 and their L2, however overall there were no reliable relationships between these measures and proficiency measures or reading times on the stimuli. In addition, Juffs (2005) replicated Juffs and Harrington's (1995) findings that Chinese learners of English were sensitive to ungrammatical vs. grammatical *wh*-movement, and were more accurate on grammatical subject *wh*-movement from an embedded finite clause than on grammatical object extraction from a finite clause. Juffs suggested that the role of finiteness should be examined further, and that since he did not divide the participants into high and low span nor perform any further analyses on the effects of WM on parsing, that this also needs to be more carefully investigated.

Using similar materials, Dussias and Piñar (2010) did investigate the effects of WM on parsing more in depth. They found that cognitive capacity, as measured by the Waters and Caplan (1996b) reading span test, is an important factor that influences L1 and L2 syntactic processing differences. Dussias and Piñar manipulated the plausibility that the *wh*-filler could be an object filler for the main verb in the stimuli used in Juffs and Harrington (1995) in a self-paced reading grammaticality judgment task administered to Chinese learners of English. Like Juffs (2005) and Juffs and Harrington (1995), Dussias and Piñar found that the Chinese learners of English correctly rejected the ungrammatical *wh*-sentences at a level well above chance. Dividing the participants into two WM span groups using a median split, Dussias and Piñar found that the higher WM group patterned like the English monolinguals in their reading latencies, whereas the lower span participants did not display a subject-object extraction contrast across the two plausibility conditions. The authors concluded that the lower span group did not integrate syntactic and semantic information in the same way as their higher span counterparts or the native speakers, thus showing support for a relationship between WM span and L2 syntactic processing. Hence, the results from Dussias and Piñar

(2010) suggest that individual differences in processing resources play an important role during L2 syntactic processing.

In addition, looking at garden path sentences, Williams (2006) found evidence of individual differences based on WM capacity in L1 speakers and L2 learners of English. Those native-speaker participants with high WM span spent more time on the postverbal region of the sentences when it was plausible for either an argument or a direct object to follow the verb than when it was implausible. Overall, the high WM native-speakers spent the most time on this region. Then the low WM native-speakers spent the next longest amount of time, followed by the high WM non-native-speakers. Finally, the low WM non-native participants showed no effect. Williams concluded that due to the variability in plausibility effects found among both native and non-native speakers the incrementality of interpretation seemed to depend upon cognitive factors such as WM instead of upon whether a person was reading in the L1 or the L2. The division between high and low WM participants that Williams drew, however, was based on participants' performance on memory probe sentences presented after sets of two experimental sentences, which may not be the most reliable measure. If he had used one of the WM tasks described previously, his results would have been more comparable to other studies.

In sum, there have been numerous studies in the field of second language acquisition that have considered the role of WM and L2 processing of morphology, morphosyntax, and syntax. Although some studies have found no correlations between WM and L2 learning (e.g. Juffs, 2004, 2005) or have used non-standard measures of WM (e.g. Williams, 2006), there are numerous studies that have found correlations between L2 learning and WM abilities such as those that have analyzed reading comprehension ability, for example, Harrington and

Sawyer (1992), those that have analyzed processing of L2 morphology and morphosyntax, such as Sagarra (2007a, 2007b, 2008) and Sagarra and Herschensohn (2010), and those that have analyzed processing of L2 syntax, such as Miyake & Friedman (1998) and Dussias & Piñar (2010). As this is a currently expanding area of study, one goal of the present dissertation is to contribute to this line of research with a particular focus on WM and the L2 acquisition and processing of argument structure, in particular of differing word orders in the L2. In the following section, I review a group of studies that specifically have looked at processing of L2 word order, some of which have more systematically shown a strong relationship between WM span and L2 processing ability.

1.6.1.4.4 Studies of WM and L2 Processing of Argument Structure Word Order

Research done on WM and L1 processing of word order (e.g. King and Just, 1991; Vos, Gunter, Schriefers and Friederici, 2001) had found evidence that object-first structures are more cognitively taxing to process than subject-first structures for native speakers. Building on this research, Havik, Roberts, van Hout, Schreuder, and Haverkort (2009) tested native Dutch speakers and German learners of Dutch on their processing of subject-object ambiguities, and on their WM capacity in both their L1 and their L2, as measured by a Dutch version of the Daneman and Carpenter (1980) reading span test and a German reading span test (see Van den Noort, Bosch, and Hugdahl, 2006). For all groups the long object relative items caused longer reading times and lower accuracy scores, suggesting that this type of structure is more cognitively taxing to process both in the L1 and in the L2. Havik et al. also found an effect of WM span for all groups on reading times. Overall, Havik et al. found that all groups showed the predicted subject preference, based on their higher accuracy as compared to the object relative sentences. Only the high span learners demonstrated a processing advantage for subject relatives, however, which might be due to the complexity of the task itself, in that lower span learners may have also been of a lower proficiency level. Havik et al. suggested that a more highly proficient L2 group overall would possibly better show the online processing preference.

As reviewed previously, Grüter and Crago (2010) examined the processing of OVS main clauses. They found that L1 Spanish and L1 Chinese child learners of French performed equally well on a backward digit span task, while the L1 Chinese group performed significantly better than the L1 Spanish group on a non-word repetition task. Since there was no independent measure of proficiency, and the WM tasks were given in the L2, these results may have been confounded with proficiency level. The one provocative finding in regards to individual differences in WM ability and L2 processing of object-first constructions was the significant correlation for WM between the backward digit span task results and the frequency of omission of the object clitics in production for the L1 Chinese group. Grüter and Crago concluded that this finding provides new support to Prévost's (2006) hypothesis that object omission in L2 French production is linked to processing limitations. In sum, there appears to be a larger processing cost for object-first constructions over subject-first constructions as made clear by WM effects in the studies reviewed in this section.

Taken together, all of the studies reviewed in section 2.5.1 have shown correlations between WM in general and L2 visual and aural processing of multiple levels of grammar. The factor of proficiency appears to exert a strong influence on WM effects on L2 processing as well. In short, the importance of including WM capacity as an individual difference in L2 acquisition has become increasingly clear. There is one specific aspect of WM, however, that may play a particularly influential role in how L2 learners process argument structure. A whole other line of study has emerged regarding the relationship of the attentional component of WM and L2 processing. In the following section, I develop the case for including tasks that test this attentional aspect or executive attention, as defined by Kane, Conway, Hambrick, and Engle (2007), in studies considering WM as an individual difference factor in L2 processing.

1.6.2 Executive Attention

As stressed in the above literature, the human cognitive system is limited in capacity, and therefore, unable to process all of the information that is present in the environment at any one point in time. Hence, one's processing of information must be selective in that one must choose to process the information that is most relevant and advantageous to one's current goals. Many of the models of working memory that have been reviewed previously also incorporate a component of attentional control or executive attention. Researchers have debated whether this selective attention to task-related or goal-relevant material is achieved through a process of inhibition or suppression of unwanted information (e.g. Green, 1998), through a process of increased activation of desired information (e.g. Cowan, 1999), or by way of a more general mechanism of cognitive control (e.g. Kane, Conway, Hambrick, and Engle, 2007). However one chooses to look at the issue, it is clearly agreed upon in the literature that this ability to focus on particular information is a crucial component of working memory. Numerous studies have tested with various conflict-related tasks both monolinguals' and bilinguals' abilities to control what is being processed. In this dissertation, I will refer to the ability of the cognitive system to focus selectively on particular information to be processed and inhibit or suppress other less relevant to the task information as executive attention.

1.6.2.1 Models of Executive Attention

There have been various models of executive or attentional control employed in the literature; among the most prominent are those of inhibitory control (e.g. Green, 1998) or inhibitory mechanisms (e.g. Hasher, Lustig, and Zacks, 2007) and those of executive attention (e.g. Kane, Conway, Hambrick, and Engle, 2007). Elaborating upon each model and test is beyond the scope of this dissertation. In general, however, and what is of central relevance here, all of the models incorporate a central executive component that controls attention towards task-relevant information and inhibits distracting or interfering information. In the Green (1998) model, there are multiple levels of control. A higher level of control regulates a language task schema that in turn controls language selection through the inhibition of potential competitors in speech perception and production. In the Hasher, Lustig, and Zacks (2007) model, the authors include three separate functions of inhibition: access, deletion, and restraint, that all serve to eliminate irrelevant information from WM. They conclude that most likely individual and intra-individual variation derives from differences in inhibitory control processes. Finally, the Kane et al. (2007) model incorporates inhibition as well, however, it does not take for granted that inhibitory control determines WMC. Kane et al. (2007) treated the executive control aspect of WM, which they coined as "executive attention", as crucial in determining individual variation in learner outcomes. Kane et al. stated that their interpretation of WMC and variation involves "attentional" and "memorial" processes that work together in concert to maintain and recover access to task-relevant information while blocking access to task-irrelevant information (2007:22). Although many researchers have kept working memory capacity and inhibition separate, whenever Kane et al. (2007) refer to WMC they are referring to the attentional processes that link both WMC and inhibition. Their more general view of executive attention is more comprehensive than an inhibitory view since it accounts for a wide range of findings in respect to WMC and its relationship to other cognitive abilities and individual differences. For the present purposes, I adopt the Kane et al. (2007) version 1.2 of the measurement model of WM, reproduced in Figure 3.

1.6.2.2 Tests of Executive Attention and Their Use in Bilingualism Research

In order to identify individual differences in executive attention, a researcher must employ an appropriate task to tap into a participant's abilities. The most commonly used tests of executive attention in the field of psychology have been the Stroop Task (see MacLeod, 1991 for a review of studies employing this task) and the Simon Task (see Hilchey and Klein, 2011 for a review of studies using this task). The Stroop Task was first popularized as a measure of attention in Experiment 2 of Stroop (1935). The basic premise of a Stroop-like task is to measure the effects of interference by showing participants names of colors written in contrasting colors, requiring participants to suppress the urge to name the color of the word instead of reading the word itself. For example, one might see the word, "blue" written in the color red, and have to read, "blue" and block out the competing information of seeing the word written in the color red. Many tasks have been shown to elicit a Stroop effect; that is, when one type of stimulus interferes with another type of stimulus presented simultaneously, thus requiring the participant to focus his attention on the goal stimulus. Also requiring the participant to focus on the task at hand and ignore distracting stimuli, the Simon Task as first developed in Simon (1969) measured response times for participants who heard a tone in one ear and in one block had to orient a handle in the same direction as the tone and in a subsequent block had to orient the handle in the opposite direction. This task has undergone many iterations involving the use of auditory, visual, and auditory-visual stimulation with similar findings (e.g., Craft & Simon, 1970; Mewaldt, Connelly, & Simon, 1980; Simon & Acosta, 1982; Simon, Craft, & Small, 1971; Simon, Sly, & Vilapakkam, 1981). Most recently, Bialystok and colleagues' work on the bilingual advantage has featured the Simon Task. Bialystok and colleagues have shown with the use of the Simon Task that being fluent in two languages can protect against dementia in old age (e.g. Bialystok, Craik, and Freedman, 2007), that bilinguals are better at resolving conflict as compared to monolinguals (e.g. Bialystok, Craik, Klein, and Viswanathan, 2004), and that bilingual children are better at problem solving than their monolingual peers (e.g. Bialystok, Martin, and Viswanathan, 2005).

While both the Stroop and Simon tasks have shown their value in determining individual variation in executive attention, a third method has more recently been adopted to test executive attention in bilinguals. The Flanker Task (Eriksen and Ericksen, 1974) has been employed in the work of Costa and colleagues (e.g. Costa, Hernández, and Sebastián-Gallés, 2008), who have investigated how the fact that bilinguals continuously control two languages during speech production enhances their attentional capabilities, and Luk and colleagues (e.g. Luk, 2008; Luk, Anderson, Craik, Grady, and Bialystok, 2010), who have investigated how different sets of brain regions are associated with the suppression of interference in bilinguals when compared to monolinguals, pointing to a different network for cognitive control that is specific to bilinguals. Although the majority of the work done with the Flanker Task concerns bilingual advantages over monolinguals, the investigation of which is not the purpose of the present dissertation, the work by Luk and colleagues is most pertinent here since it is the source of the task used in the present dissertation. In the version of the Flanker Task adopted here, the participant views a red arrow that is flanked by black arrows on both sides. The participant must click the right mouse button if the red arrow points right, and the left mouse button if the red arrow points left. On one third of the trials (randomly presented), the red arrow points in the direction of the black arrows (congruent trials), and on another third of the trials, the red arrow points in the opposite direction (incongruent trials). On the final third of the trials, the red arrow is flanked by diamonds, which are neutral or "go" trials since they present no conflict with the arrow's direction. Reaction times (RTs) are measured and then both an inhibition effect (the difference in RTs between the incongruent trials and the go trials) and a facilitation effect (the difference in RTs between the go trials and the congruent trials) are calculated. The version that Luk and colleagues have employed also includes a no-go trial, which requires the participant to refrain from pressing any button.

The Flanker Task has been shown to be a test of more general cognitive control, and to the best of my knowledge, there have been no studies conducted that consider direct correlations between performance on the Flanker Task (or on the Stroop or Simon tasks) and L2 processing beyond the lexical level. However, since executive attention has been shown to be linked to individual differences in working memory capacity, it is now crucial to examine the relationship between executive attention abilities in L2 learners and processing of argument structure.

1.7 The Present Study

In the discussion above, I have reviewed the literature on universal strategies (e.g. VanPatten, 1984, 1996), on the transfer of L1 strategies to the L2 (e.g. MacWhinney, 1987, 1989, 1992; Ellis, 2006a, 2006b; Ellis and Sagarra, 2010), and on a mixed account of universal as well as transfer of strategies that involves proficiency as an important factor (e.g. Hopp, 2006), choosing to adopt Ellis' Associative-Cognitive theory as the frame of reference for the present dissertation because it offers the most complete account of why learners have difficulty processing certain argument structures in L2 Spanish. I also reviewed the research on working memory and its role in L2 processing of morphology, morphosyntax, and syntax. Finally, I considered the role of executive attention as a component of working memory and its influence on L2 learners' abilities to process argument structure in the L2. My goal in so doing is to contextualize the experiments in the present dissertation.

The research above raises important questions regarding our understanding of how adult second language learners process argument structure. For example, following Ellis' CREED, structures that are low salience, redundant, and provide cues that compete with other cues are less likely to be acquired successfully by L2 learners. The L1 offers additional hurdles, according to this theory, in regards to interference and overshadowing and blocking. Thus, it is crucial to test participants who hail from different L1 backgrounds when examining L2 processing. In addition, it has been suggested that processing difficulty could be related to the higher processing demands incurred by object-first constructions (e.g. Havik et al., 2009; Hopp, 2006). Although Harrington and Sawyer (1992), Sagarra and Herschensohn (2010), and Dussias and Piñar (2010) have found significant correlations between working memory ability and language processing ability, other studies (e.g. Juffs, 2004, 2005) have not found such correlations. Therefore, it is important to test learners' working memory ability to determine if L2 learners' processing patterns of first noun structures are mediated by this individual difference (as in Havik et al., 2009). As of yet, the link between executive attention and working memory capacity has not been extensively explored in the realm of second language learners' L2 processing of argument structure, and this possible relationship must be addressed. Moreover, in all of the literature reviewed, proficiency appears to be an important variable in how learners process L2 morphosyntax, yet some studies have used self-rated, teacher-rated, or no proficiency measures instead of providing an independent measure of proficiency (e.g. Isabelli, 2008), which could seriously confound their results. The inclusion of participants of multiple proficiency levels (or the treatment of proficiency as a continuous variable) within each language group is crucial so that these variables can be teased apart.

Taking all of the findings together from the literature reviewed previously, it becomes clear that many factors are involved in shaping individual differences in L2 learner outcomes in processing argument structure. Whether or not a learner transfers processing strategies or particular structures from the L1, his ability to process argument structure successfully in the L2 appears to also be modulated by his proficiency level and his working memory capacity, with his executive attention capabilities weighing in as the leading factor determining his WMC. In order to formulate the most representative picture of what is happening when a second language learner processes argument structure in the L2, it is imperative to include all of these factors. It is also crucial to create a study in which the structure that second language learners are asked to process is sufficiently challenging in order to have a wide range of results and to not have ceiling effects. With this in mind, the experiments here examine a structure in which there is a preverbal clitic pronoun and a postverbal subject, a construction which is both common and licit in Spanish but not in English, as in (16) and (17).

(16)	*Hoy*	***la***	*busca*	*el niño*	*en el parque*.
	[Today	3RD SINGFEMACC	look for-3RD SINGPRE	the boy	in the park]
	Today the boy looks for her in the park.				

(17) *Today **her** looks for the boy in the park.
Today the boy looks for her in the park.

In addition to L1 English learners of Spanish, which have been studied in large numbers (e.g. Liceras, 1985; VanPatten, 1984), it is important to consider learners from another L1

Romance language learning L2 Spanish in an attempt to tease apart the variable of first language. To the best of my knowledge, only Isabelli (2008) has done this (with L1 Italian speakers) with the present argument structure. However, her participants were not tested for proficiency level nor were they measured on WMC or executive attention capabilities.

Another Romance language that allows for preverbal clitic pronoun OVS structures like in Spanish is Romanian, for example, as in (18).

(18)	*Astăzi*	***o***	*caută*	*băiatul*	*în parc.*
	[Today	3^{RD}SINGFEMACC	look for-3^{RD}SINGPRE	boy-the	in the park]

Today the boy looks for her in the park.

Romanian has pronominal clitics that occupy preverbal positions (Dobrovie-Sorin, 1999), just as in Spanish. Thus, Romanians have significant experience with processing preverbal clitics. Given that one of the research questions of the present dissertation is to find out if transfer plays a role in L2 processing of argument structures at varying levels of proficiency, it is crucial to include participants with both the L1 of English and the L1 of Romanian in order to form a point of comparison based on first language. If one were to assume that there is transfer of L1 structures, we should predict that L1 Romanian-speakers, who already have an OVS clitic pronoun structure with a postverbal subject in their L1, would have fewer problems interpreting such a structure in L2 Spanish than L1 English-speakers, who do not possess such a structure in their L1. It can be predicted based on Ellis' theory that L1 Romanians would not have the same interference as L1 English learners of Spanish due to the presence of preverbal clitic structures in their L1. If there is no transfer of L1 structures, but some universal default for SVO word order or a first noun principle (see VanPatten, 1984), then there should be no differences in processing outcomes between learners of the same proficiency level who have an L1 that licenses such structures and learners who have a rigid SVO word order in their L1.

In addition, research has shown that individuals differ in working memory capacity and executive attention, and that this variation has surfaced in bilinguals as individual differences in L2 abilities and processing (e.g. Dussias and Piñar, 2010; Costa, Hernández, Costa-Faidella, and Sebastián-Gallés, 2009). Therefore, these variables should be taken into account as well when examining adult L2 acquisition of argument structure. Another research question asked in the present dissertation is whether and to what degree cognitive individual differences such as WMC and executive attention play a role specifically in how second language learners process a particular argument structure in the L2.

Thus, in the present study, I examine how L1 English and L1 Romanian learners of Spanish, of varying proficiency levels, process first nouns in OVS structures such as in (16) with the specific aims of (1) further clarifying the influence of the L1 in processing strategies, (2) assessing the relative impact of proficiency level, and (3) better establishing the roles of WMC and executive attention as cognitive individual difference measures on adult L2 learners' processing of argument structure. With this dissertation, I strive to enlighten the debate about why adult second language learners have difficulty attaining native-like processing abilities. To this end, I design and carry out two sentence processing experiments in which participants listen to and read sentences of varying word orders (SVO and OVS) and make a judgment on their meaning. Finally, I test the participants on both a working memory capacity task and an executive attention task, as well as on an independent measure of proficiency. I outline my method and procedure in greater detail in the following chapter.

CHAPTER 2: Methodology

2.0 Introduction

My aim in this chapter is to describe the methods and procedures used in the present study as well as provide some of the descriptive statistics for the L2 learner variables. First, in section 2.1, I present the experimental design of the study. Then, in section 2.2, I provide an overview of the target structure and some relevant background on clitic structures in both Spanish and in Romanian. In sections 2.3 and 2.4, I describe the participants and the materials. Next, in section 2.5, I outline the data collection procedure, and in section 2.6, I explain how I scored the data. In section 2.7, I include the descriptive statistics for the individual difference measures. Finally, in section 2.8, I discuss the statistical analyses used for the data.

2.1 Overview of Experimental Design

In order to gauge how successfully learners are processing structures with preverbal clitic pronouns and postverbal subjects in Spanish sentences, I tested 65 L1 English learners of Spanish of varying proficiency levels, 72 L1 Romanian learners of Spanish of varying proficiency levels, and 36 Spanish-speaking monolingual controls on seven to nine tasks in one session that lasted between one and two hours. An overview of the tasks employed in this study, each of which is addressed in further detail below, is provided in Table 1.

Table 1. *General Experimental Design*

Study target vocab.	Language history questionnaire	Proficiency test(s)	Sentence processing task	Flanker test	Sentence processing task	Letter-Number Sequencing test	Vocab. test
(5 min.)	(5 min.)	(10-20 min)	(20-40 min.)	(5 min.)	(20-40 min.)	(10 min.)	(10 min.)

In brief, participants began the experiment by studying the target vocabulary that would be employed in the sentence processing tasks. They subsequently completed a language history questionnaire (LHQ) (see Appendix A) used to measure homogeneity within and between the learner groups as well as to help assess whether the monolinguals were truly monolingual. After completing the LHQ, the learner participants also took an objective proficiency test that examined their linguistic knowledge, and determined their proficiency level in Spanish in order to later be analyzed as a separate variable. In addition, at the beginning of the first sentence processing task, participants were asked to rate their proficiency in reading, speaking, listening, and writing in Spanish based on a five-point scale[1].

The two sentence processing tasks (one testing reading and the other tapping into auditory processing of the target stimuli) employed in the study measured comprehension accuracy of both canonical SVO structures and structures containing preverbal clitics and postverbal subjects (OVS). A test of executive attention (the Flanker test) assessed the participants' ability to inhibit extraneous information while performing a task and was administered between each of the sentence processing tasks. This task also served as a break for participants between the two sentence processing tasks. A working memory test (a letter-number sequencing (LNS) test) gauged participants' level of working memory ability. In addition, the vocabulary test ensured that the participants had understood the words used in the tasks, and were thus, of a sufficient proficiency level to participate. Each of these tasks is discussed in further detail in sections 2.3 and 2.4.

The current study follows a repeated measures design with all of the participants performing all of the experimental tasks. Due to design and logistical factors, however, there were slight differences in the procedure for each language group. In the case of the

---

[1] The participants who were not L1 English speakers also took an English proficiency test. However, the results from this measure will not be part of the final analysis. I mention the results from the English proficiency test in the conclusion in chapter 5.

Romanian-Spanish learners and the Spanish monolinguals, the participants performed all of the tasks pertinent to their group during a single session. The Romanian-Spanish learners completed all of the aforementioned tasks, while the monolingual participants only completed the following tasks: (1) the Language History Questionnaire, (2) the English proficiency test, (3) either the reading or the listening sentence processing task, (4) the Flanker test, (5) the reading or the listening sentence processing task, and (6) the working memory test. In the case of the English-Spanish learners, the participants performed the following tasks on an online course management system before coming to the laboratory: (1) the Language History Questionnaire and (2) the Spanish Proficiency Test (which included multiple-choice grammar at three levels of difficulty, and listening and reading comprehension questions). Then, in the laboratory, the English-Spanish participants performed the following tasks: (1) study vocabulary for ten minutes, (2) the reading or listening sentence processing task (including self-ratings of proficiency in Spanish reading, writing, speaking, reading), (3) the Flanker test, (4) the reading or listening sentence processing task, (5) the working memory test, and (6) the vocabulary test (multiple choice).

2.2 The Critical Syntactic Configuration: Preverbal Clitics in OVS Sentences

A number of studies suggest that learners of Spanish show initial difficulty with learning clitic structures and their placement in Spanish, yet eventually they are successful (for L1 French learners, see Bruhn-Garavito & Montrul, 1996; for both L1 English and L1 French learners compared, see Duffield & White, 1999; Liceras, 1985). This could be attributable to important differences between English and Spanish pronoun structures and placement. Both English and Spanish have strong pronouns (e.g. *mí, ti,* me, you, him, etc.). Unlike English, Spanish has weak pronouns or clitics (e.g. *me, te, lo, la*, etc.). These clitics are marked for case, with direct object pronouns carrying accusative case and indirect object pronouns being marked for dative case. Because these clitic pronouns must attach to a host, they are always found adjacent to a verb in Spanish. This is also the case for pronominal clitics in Romanian. The present dissertation hypothesizes that Romanian learners of Spanish will have less difficulty acquiring and processing clitic structures due to the presence of similar structures in their L1, while the clitic processing data for the L1 English-speakers will replicate the results from the aforementioned studies. In this section, I present some background on clitic structures in Spanish and in Romanian, and highlight the similarities between these structures in the two languages. Then, I discuss the importance of examining how L2 learners of different language backgrounds process the particular clitic structure chosen for this dissertation.

2.2.1 Spanish Clitics

Spanish has a relatively robust system of clitics (Zagona, 2002). Spanish clitics are not uniformly pronominal, however, so I will not refer to them as such. Spanish clitics have other grammatical functions, and are not necessarily related to verbal arguments. The full inventory of Spanish clitics is illustrated in Table 2.

Table 2. *Spanish Clitics*

	Nominative	Accusative	Dative
1st person sg.		me	me
1st person pl.		nos	nos
2nd person sg.		te	te
2nd person pl.		os	os
3rd person sg. m.	se	lo	le
3rd person sg. f.	se	la	le
3rd person pl. m.		los	les
3rd person pl. f.		las	les
3rd person sg/pl. refl.		se	se

As shown in Table 2, Spanish accusative and dative clitics differ morphologically only in the third person in the standard dialect. There are certain dialects, *leísta* dialects, in which the appropriate dative form is used instead of any third person singular or plural accusative forms.

In addition, third person clitics in the dative case do not differentiate between gender. There is one nominative clitic in Spanish, *se*. It is used to indicate "one", and is referred to as "impersonal *se*".

2.2.1.1 Clitic Placement

The placement of clitics in Spanish is as follows, as adapted from Zagona (2002), in (19).

(19) a. For non-reflexive 3^{rd} person clitics, dative clitics precede accusative clitics
b. Non-3^{rd} person clitics precede 3^{rd} person clitics
c. Second person clitics precede 1^{st} person clitics
d. *Se* precedes other clitics
e. Phonetically identical sequences are excluded

Spanish clitics always occur adjacent to a verb, as in (20), and follow imperatives, as in (21), infinitives, as in (22), and gerunds, as in (23)[2].

(20) a. *Marta* *los* *lava.*
[Marta 3^{RD}PLMASCACC wash-3^{RD}SINGPRE]
Marta washes them.

b. *Los* *lava* *Marta.*
[3^{RD}PLMASCACC wash-3^{RD}SINGPRE Marta]
Marta washes them.

(21) a. *Lávalos* *ahora.*
[Wash-2^{ND}SINGIMP+3^{RD}PLMASCACC now]
Wash them now!

b.**Los lava ahora.*

(22) a. *Quise* *comprártelo.*
[Want-1^{ST}SINGPA buy-INF+2^{ND}SINGDAT+3^{RD}SINGMASCACC]
I wanted to buy it for you.

b. **Quise te lo comprar.*

[2] All Spanish examples are my own.

(23) a. *Estaba* *leyéndolo.*
[Be-1stSingPa read-prt+3rdSingMascAcc]
I was reading it.

b. **Estaba lo leyendo.*

When the imperative is negated, Spanish clitics precede them as well as other finite verbs, as illustrated in (24) and (25).

(24) *No* *lo* *comas.*
[not 3rdSingMascAcc eat-2ndSingImp]
Don't eat it!

(25) a. *Juana* *lo* *compró* *anoche.*
[Juana 3rdSingMascAcc buy-3rdSingPa last night]
Juana bought it last night.

b. **Juana compróIo anoche.*

In Spanish progressive constructions, clitics either precede the auxiliary or follow the participle, as in (26)[3]. However, it is ungrammatical for Spanish clitics to follow past or passive participles, as in (27b) and (28b).

(26) a. *Jorge* *lo* *estaba* *comiendo.*
[Jorge 3rdSingMascAcc be-3rdSingPa eat-prt]
Jorge was eating it.

b. *Jorge estaba comiéndolo.*

(27) a. *Jorge* *ya* *lo* *había* *comprado.*
[Jorge already 3rdSingMascAcc have-3rdSingPa buy-pprt]
Jorge had already bought it.

b. **Jorge ya había comprádolo.*

[3] Although the topic of clitic climbing has been studied at great length in the literature, since the present dissertation did not include the environment in which this occurs (there were no infinitives or progressive constructions in the stimuli), I will not address this here. See Davies (1995) for a discussion on clitic climbing.

(28)	a. *La carta*	*ya*	*me*	*fue*	*escrita.*
	[The letter	already	1ST SINGDAT	be-3RD SINGPA	write-PPRT]

The letter was already written to me.

b. **La carta ya fue escrítale.*

2.2.1.2 Clitic Doubling

Although the present dissertation does not focus on a structure that involves clitic doubling, it is an important aspect of clitics that should be considered when comparing the processing of clitic structures across languages since the presence of two structures that provide redundant information may have an effect on L2 learners' processing of low salience structures, according to the Associative-Cognitive theory (Ellis 2006b). Clitic doubling, according to Torrens and Wexler (2000), is a construction in which both the full noun phrase and the clitic pronoun that receives the same theta role and is of the same case as the full noun phrase are present. An example of clitic doubling in Spanish would be as in (29).

(29)	*Juana*	*te*	*lo*	*dio*	*a*	*ti.*
	[Juana	2ND SINGDAT	3RD SINGMASCACC	give-1ST SINGPA	to	2ND SINGDAT]

Juana gave you it.

Spanish differs from other Romance languages like French and Italian in that it allows for clitic doubling or reduplication. Romanian is the only other widespread Romance language that allows for clitic doubling (Sánchez and Al-Kasey, 1999). Clitics may co-occur with phonologically independent pronouns, such as in (30a). The presence of the direct object clitic is obligatory with a pronominal direct object in Standard Spanish, as in (30a) and (30b), but impossible with a non-pronominal direct object, as shown in (31a) and (31b).

(30) a. *Los* *llamé* *a* *ellos*
[3RDPLACC call-1STSINGPA to them]
I called them.

b. **llamé* *a* *ellos*
[call-1STSINGPA to them]
I called them.

(31) a. *Comí* *arroz con pollo* *anoche*
[eat-1STSINGPA rice with chicken last night]
I ate rice with chicken last night.

b. **Lo* *comí* *arroz con pollo* *anoche*
[3RDSINGMASCACC eat-1STSINGPA rice with chicken last night]
I ate rice with chicken last night.

Indirect object or dative clitic doubling is obligatory with pronouns in all varieties, as illustrated in (32).

(32) a. *Le* *di* *la flor* *a* *ella.*
[3RDSINGDAT give-1STSINGPA the flower to her]
I gave the flower to her.

b. **Di* *la flor* *a* *ella.*
[give-1STSINGPA the flower to her]
I gave the flower to her.

In the case of nominative clitics, impersonal *se* does not double an overt subject, as shown in (33).

(33) **uno/él/el hombre,* *se* *habla* *demasiado* *allí.*
[One/he/the man, 3RDSINGNOM talk-3RDSINGPRE too much there]
One, one talks too much there.

2.2.2 Romanian clitics

Romanian pronominal clitics, which are of primary concern in the present dissertation, are full and non-full, accentuated and unaccentuated (Calude, 2001; Popescu, 2000). Table 3, adapted from Avram (1986), summarizes the forms for accusative and dative pronominal clitics.

Table 3. *Romanian Pronominal Clitics*

TYPE			CASE					
Person	Number	Gender	Accusative			Dative		
			Full		Non-Full	Full		Non-Full
			Accent.	Non-Acc.		Accent.	Non-Acc.	
1st	Sing.		mine	mă	m	mie	îmi	mi
	Pl.		noi	ne	ne	nouă	ne/ni	ne
2nd	Sing.		tine	te	te	ție	îți	ți
	Pl.		voi	vă	v	vouă	vă/vi	v
3rd	Sing.	Masc./Neu.	el	îl	l	lui	îi	i
	Pl.			îl	i	lor	le/li	li
	Sing.	Fem.	ea	o	o	ei	îi	i
	Pl.			le	le	lor	le/li	li

As shown in Table 3, there are no nominative pronominal clitics in Romanian, unlike the impersonal *se* in Spanish, mentioned previously. One fact worthy of mention is that Romanian clitics differ from the clitics of other Romance languages in that singular clitics show a morphological distinction between the dative and accusative case not only in the third person, but also in the first and second singular person. Additionally, many of Romanian pronominal clitics show alternations in form. For example, there are clitics that alternate between a full vowel (e and i) and a glide (e and j) in their pronunciation. Also, some show an alternation between the vowels î or ă and no vowel (e.g. the first person singular accusative *mă* and *m*). Finally, the first and second plural dative clitics have two possible forms, one ending in *e* and one ending in *i*. The reduced or non-full forms appear when other clitics are present or when the lexical verb that they precede begins with the vowels *a* or *o* (Săvescu Ciucivara, 2009).

2.2.2.1 Clitic Placement

In regards to the placement of clitics, although Romanian has free word order, it does not license the moving around of clitics in a sentence (Calude, 2001), which is similar to Spanish. Pronominal clitics must precede the verb (or auxiliary or modal if there is one present) even if the full clitic form is used[4], as in (34a, b, and c), adopted from Calude (2001). This is the same as in Spanish, as shown in (26)-(28).

(34) a. *Îi* *dau* *cartea.*
[3RD SINGMASCDATNON-FULLNON-ACCENTED give-1ST SINGPA book.DEFART]
I give him the book.

b. **Dau* *îi* *cartea.*
[give1ST SINGPA 3RD SINGMASCDATNON-FULLNON-ACCENTED book.DEFART]
I give him the book.

c. **Dau* *lui* *cartea.*
[give1ST SINGPA 3RD SINGMASCDATFULLNON-ACCENTED book.DEFART]
I give him the book.

The only exception to this rule is for past tense verbs with the auxiliary HAVE when a 3rd person singular feminine clitic is used. In this case, the clitic follows the lexical verb, as in (35).

(35) *Elefantul* *a* *stropit-* *o* *pe* *fată*
[Elephant-the has sprinkled her-ACC on girl]
The elephant sprinkled the girl.

Under certain conditions, preverbal clitics can also be phonologically attached to their left, as in (36), taken from Dobrovie-Sorin (1999), which is not attested in Spanish.

(36) *Maria-i* *scrie* *des*
[Maria-2ND SINGDAT write-3RD SINGPRE frequently]
Maria writes him/her frequently.

When more than one clitic is present in a structure, Romanian pronominal clitics can only occur in a certain order with respect to each other, with dative clitics first followed by accusative clitics, as in (37), adopted from Luís (2004). This mirrors the Spanish clitic order.

[4] The choice of the full form over the non-full form of a pronominal clitic in Romanian is simply a way to emphasize its meaning, indicating the “person” expressed by the clitic as the topic of the sentence, or to show contrast.

(37) *I* *-l* *dau*
[3RD SINGDAT 3RD SINGMASCACC give-1ST SINGPRE]
I give it to him.

Unlike other Romance languages, Romanian clitics can be separated from a lexical verb by an adverbial clitic (Săvescu Ciucivara, 2009).

2.2.2.2 Clitic Doubling

In the case of clitic doubling in Romanian, there are a few properties that are important to note. When direct objects occur as complements of the preposition *pe*, they must be doubled by a direct object clitic, as in the example (38), adopted from Babyonyshev and Marin (2006). When they surface as complements of a verb (no preposition), it is ungrammatical to double them with a clitic, as in (39), also adopted from Babyonyshev and Marin (2006).

(38) *Elefantul* *l-* *a* *atropit* *pe tigru*
[Elephant-the 3RD SINGMASCACC has sprinkled on tiger-M]
The elephant sprinkled the tiger.

(39) **Elefantul* *l-* *a* *stropit* *tigrul*
[Elephant-the 3RD SINGMASCACC has sprinkled tiger-the]

Clitic doubling in Romanian is not possible with inanimate objects. Furthermore, when a proper name acts as the direct object, only the clitic doubling construction is possible. Clitic doubling is also required for indirect object constructions, like in Spanish. I will not, however, elaborate on dative clitic doubling since these are not the focus of the present dissertation.

The presence of clitic doubling in Romanian has provided evidence in favor of the affixal status of Romanian pronominal clitics. Monachesi (1998) stated that since clitics can co-occur with full complements, then they are acting as agreement markers. I will briefly discuss the various arguments that have been posited based on data from Romanian in regards to the status of clitics in the following section.

2.2.3 The Status of Clitics as Words or Affixes

The apparent inflexibility of ordering of clitics in Romanian has been cause for researchers to claim that clitics are not words, but affixes (see Luís, 2004). Whether or not pronominal clitics are independent syntactic forms or affixes has been debated in the literature for many years with no decisive answer. Also, there are specific restrictions on certain combinations of clitics, such as first person dative clitics cannot be present with first person singular accusative clitics nor with first and second person plural accusative clitics. Luís (2004) made the case that these seemingly idiosyncratic limitations on combinations of clitics weaken the argument for the word status of clitics. She added that morphophonological alternations, such as the fact that the plural dative clitics *ne*, *vă*, *le* change to *ni*, *vi*, *li* when followed by accusative clitics, previously argued (Gerlach, 2001; Popescu, 2000) to follow from phonological rules are instead triggered by morphosyntactic features of the adjacent pronouns, which supports their classification as affixes. Luís proposed an inflectional approach to Romanian cliticization such that clitics are either stem-affixes or phrasal affixes, depending on the properties of the clause. Monachesi (1998) also posited an argument in favor of clitics as affixes. She included data on coordination. In Romanian, if two verbs that share the same clitic are coordinated, then the clitic has to be repeated, as in (40), adopted from Monachesi (1998).

(40)	*el*	*o*	*dorea*	*şi*	*o*	*căuta*
	[he	3RD SINGFEM	desire-3RD SINGPRE	and	3RD SINGFEM	look for-3RD SINGPRE]
	He desires her and looks for her.					

It is ungrammatical for clitics to have wide scope over coordination, as shown in (41), adopted from Monachesi (1998).

(41)	**el*	*o*	*dorea*	*şi*	*căuta*
	[he	3RD SINGFEM	desire-3RD SINGPRE	and	look for-3RD SINGPRE]
	He desires her and looks for her.				

This, Monachesi argued, supports the affixal status of Romanian pronominal clitics since syntactic words have wide scope over coordination.

In response to these and other accounts that clitics should be analyzed as affixes, Popescu (2000) argued that Romanian object clitics have more than one possible position in regards to the verb in that they appear before the verb in most moods and tenses but after it in the imperative and the gerund, just as in Spanish. This variation contradicts one of the criteria for affixes set by Zwicky and Pullum (1983): that affixes should have a fixed position. In addition, Popescu argued that since Romanian object clitics, as opposed to clitics in other Romance languages like Spanish, can optionally cliticize to a phonological word to which they do not belong, which contradicts the criterion for affixes that states that affixes always belong phonologically to their syntactic host. Also, the relative order of object clitics in Romanian is unchanged whether they occur before or after the verb, thus their syntactic behavior depends on other clitics and not on the verb stem like the behavior of affixes would (see Gerlach, 1998 for this criterion). Finally, splitting clitics with any other morpheme is not allowed in Romanian, which led Popescu to analyze clitics in her work as a morphological and syntactic unit instead of as affixes. Whether one believes clitics are affixes, affixal phrases, or full NPs, it is crucial to point out that for the purposes of the present dissertation it is the presence of clitics that behave in a similar way in both Spanish and in Romanian that leads me to hypothesize that there may be transfer of this structure.

2.2.4 Similarities with Spanish Clitics

Beyond the general similarities of form and placement, Romanian clitics are similar to Spanish clitics in that, regardless of type of clitic or position, neither direct nor indirect object clitics show agreement with participles. This is not the same in French or Italian, which is important to note. In fact, Babyonyshev and Marin (2006) found that in Romanian, children exhibit acquisition patterns of object clitics that are similar to those of children acquiring Spanish. Both language groups show a low rate of object clitic omission when obligatory, early appearance of object clitics in production data, and low use of definite direct objects. The relative ease at which children acquire clitics in Spanish and Romanian, Babyonyshev and Marin propose, is due to the lack of a need for participle agreement with clitics in these languages.

Calude (2001) proposed that Romanian clitics pattern more like Serbo-Croatian clitics than French or other Romance language clitics. However, the similarities that she found between Romanian and Serbo-Croatian clitics that were dissimilarities between French and Romanian clitics do not necessarily indicate that Romanian clitics are unlike Romance clitics as a whole. Some of the key aspects that Calude argues differ between Romanian and French clitics, such as the fact that pronominal clitics are placed before modal verbs in Romanian but after modal verbs in French[5] (as in (42) and (43)), are aspects that are, in fact, the same between Romanian and Spanish (as shown in (44)).

(42)	*Eu*	*vă*	*pot*	*sugera*	*Hotel Parc.*
	[1ST SINGNOM	2ND PLDATACCFULLNON-ACCENTED	can	suggest	Hotel Parc]

I can suggest Hotel Parc to you.

(43)	*Je*	*peux*	*vous*	*proposer*	*l'hôtel*	*du*	*Parc.*
	[1ST SINGNOM	can	2ND PLDAT	suggestINF	DEFART'hotel	of.the	Parc]

I can suggest Hotel Parc to you.

(44)	*Yo*	*os*	*puedo*	*sugerir*	*el*	*Hotel Parc.*
	[1ST SINGNOM	2ND PLDAT	can	suggestINF	DEFART	Hotel Parc]

I can suggest Hotel Parc to you.

Two other differences between Romanian clitics and French clitics that Calude points out are that in Romanian two consecutive vowels are maintained and not elided as in French and that Romanian does not use clitics after prepositions, but full pronominal forms, while

[5] All Romanian and French examples here are from Calude (2001). All Spanish examples are my own.

French does. Again, these aspects of clitics may differ between Romanian and French, but they do not differ between Romanian and Spanish, which supports my rationale to include Romanian learners of Spanish, as opposed to learners of Spanish from another Romance L1. Furthermore, since Romanian clitics have been shown to be similar in function to Spanish clitics, this strengthens my argument that experience with processing clitics in Romanian could aid in subsequent L2 processing of Spanish clitics.

2.2.5 The Present Structure

In particular, it is with respect to clitic placement and word order permutations that L1 English learners of Spanish have shown difficulty in previous research. In English, there is, generally speaking, a canonical subject-verb-object (SVO) word order, which in turn serves as a very important and reliable cue to the subject of the sentence. If we change the order of the constituents, for example, in the sentence "The girl kisses the boy" to "The boy kisses the girl", we have completely changed the meaning of the utterance. In Spanish, Italian, and Romanian, among other languages, however, this is not the case, given that subject noun phrases can appear pre- or postverbally in each of these languages. For example, in the sentence in (45) we can rearrange the constituents without changing the general meaning, to the sentence shown in (46), due to the presence of the object marker "a".

(45)	*Lo*	*besa*	*la niña.*
	[3^{RD}SINGMASCACC	kiss-3^{RD}SINGPRE	the girl]
	The girl kisses him.		

(46)	*O*	*caută*	*băiatul*
	[3^{RD}SINGMASCACC	kiss3^{RD}SINGPRE	girl-the]
	The girl kisses him.		

Such variable word orders have been shown to be confusing for learners, even in very simple structures (cf. VanPatten (1984) and VanPatten & Cadierno (1993)). The extent to which the word order variability in L1 may affect the acquisition of such structures in L2 remains a relatively open question. It is thus important to examine these structures further, incorporating learners from an L1 background that has clitics and licenses similar pronoun placement and usage (Romanian), and comparing them to learners from an L1 background that does not (English). The listening and reading experiments reported here examine this issue directly by presenting these groups of learners with sentences with preverbal clitics and postverbal subjects (as shown in (45) and in (46)), a structure that exists in both Spanish and Romanian, but is ungrammatical in English.

2.3 Participants

The participants for the study were 65 L1 English learners of Spanish, 72 L1 Romanian learners of Spanish, and 36 Spanish-speaking monolingual controls. The participants' ages ranged from 18-56, and there were 126 females. The L1 English learners hailed from two main constituencies: third- and fourth-year college-age learners of Spanish and professional, post-graduate and graduate learners of Spanish from two American universities. The L1 Romanian learners were recruited from individuals studying at the Instituto Cervantes' Spanish language school in Bucharest, Romania. They were college-age and professional, post-graduate first- through fourth-year learners of Spanish.

All of the learner participants scored above 88% correct on the post-experiment vocabulary test, which indicated that they all had sufficient knowledge of Spanish to participate in the study. To provide a descriptive baseline for the study, reading and listening data were also collected for 36 L1 Spanish-speakers from monolingual communities in northeastern Spain.

2.4 Materials

As mentioned above, participants completed between six and eight tasks during the course of the experiment, depending on their language group. Again, these were: vocabulary study, assessment measures (a language history questionnaire, proficiency tests, and a vocabulary test), reading and listening sentence processing tasks, and two cognitive tests (the Flanker test and the working memory LNS test). I describe each of the materials in detail in this section.

2.4.1 Assessment Measures

Prior to the language history questionnaire, the learner participants were given 10 minutes to study a list of the target vocabulary words from the sentence processing experiments. This pre-training was done to ensure (1) that the learners knew all of the words used in the experiment and would have no comprehension difficulties, which would confound the processing results, and (2) that there were no misunderstandings in regards to dialectal word choices (e.g. *piso*, which means "floor", can also refer to "apartment" in the standard dialect of Spain). All of the vocabulary were from the standard dialect of Spain, the dialect taught in the Instituto Cervantes and the dialect primarily taught in the universities in the United States from which the L1 English learner participants derived.

All participants completed a language history questionnaire in their L1 that asked them about the languages they use in their daily lives. This information was used to control for any discrepancies in language experience among participants. Specifically, the questionnaire determined the participants' L1, languages spoken at home and early in life, languages studied in school, languages that they could read, write, and speak, and which language was most comfortable for them. For the learner groups, the questionnaire also revealed a participant's age of first exposure and length and time of any study or residence in a Spanish-speaking country. Since the number of years spent studying the language is not necessarily comparable among participants, let alone among language groups, this variable was not included in the analysis. The length of study or residence in a Spanish-speaking country in months was included, however, since this measurement was more directly comparable across the participants. One participant whose first language was not that of her respective group (a Romanian participant who stated that her first language was not Romanian) was excluded from the study, and performed no further tasks.

In addition to the language history questionnaire, all of the participants except the monolingual controls performed a 68-item Spanish proficiency test, which consisted of a selection of sections from the Diploma de Español como Lengua Extranjera (DELE) exam, and included a basic, intermediate and advanced level grammar section, a reading comprehension section, and a listening comprehension section. The learner participants also gave a self-rating of their Spanish reading, writing, speaking, and listening abilities based on a scale from 1 (minimum ability) to 5 (native-like ability) before the first sentence processing task. Having both the objective measure (the proficiency test) and the subjective measure (the self-ratings) is important since it has been suggested, but not confirmed, that self-ratings correlate with L2 abilities and proficiency (Blanche and Merino, 1989; cf. Brantmeier, 2006).

I also administered an adapted version of the Test of English as a Foreign Language (TOEFL) exam as the English proficiency test to the participants in the L1 Romanian and Spanish monolingual groups. This was done to assess participants' knowledge of English. This test included a basic reading comprehension and basic grammar section. There were 20 items on this test. No participants were excluded from the study based on their scores from this test. I do not include this measure in the final analysis, but I do report the results from it in the final chapter. The full text for the two proficiency tests that I used can be found in Appendices B and C.

Finally, all of the learner participants completed a vocabulary post-test, which evaluated their knowledge of the Spanish vocabulary used in the experimental tasks. This was administered after the experimental tasks were completed to confirm that they knew the vocabulary used in the target sentences. There were 122 items on the vocabulary test: 40 verbs, eight adverbs, and 74 nouns. The participants matched the Spanish word to its equivalent in their L1. There was no time limit imposed on this test. The participants took between 15 and 45 minutes to complete this test, depending on their proficiency level, and all scored above 88% correct. This confirmed that they all had the necessary knowledge of the language to be able to complete the sentence processing tasks, and thus, were all included in the study. The vocabulary test is reproduced in Appendix D.

2.4.2 Sentence Processing Tasks

The data under consideration were presented in the context of a larger study employing eyetracking. Generally speaking, eyetracking is an online measure that records the location and time of eye fixations on a computer screen. Eyetracking, as opposed to other methods such as self-paced reading tasks, allows participants to read and process whole sentences in a more natural way, with the ability to regress and fixate as long as needed for comprehension of the stimuli, a process which more closely approximates the way that people read naturally. In addition, this methodology allows for the use of aural stimuli. While participants listen to a stimulus, their eye fixations on visual stimuli can be recorded. Due to the sensitivity of the measurement of eye movements, responses can be time-locked to the input without interrupting the natural flow of language comprehension (Tanenhaus, 2007). In addition, to make the processing even more naturalistic (Tomasello, 2003) and perhaps more cognitively taxing, the researcher can present the stimuli aurally in one task and then visually in another, and compare the results from the two modalities for further validity, which is what I chose to do here.

For the present dissertation, I focus on the accuracy data collected for both the reading and listening sentence processing tasks. While later research may analyze the data from eyetracking measures on the reading portion of the reading experiment, it is not the primary focus of this study. Here the primary goals are to investigate the possible contributions of L1 and proficiency to the successful processing of argument structure, and to do so by comparing as directly as possible performance in the visual and auditory modalities. For this, the dependent measure of accuracy measures provides the simplest and most straightforward basis for comparison across the two modalities.

The data were collected in a laboratory setting on the main campus of the Pennsylvania State University, on the campus of Temple University, at the Instituto Cervantes in Bucharest, Romania, and in the region of Aragón, Spain. The SR Research EyeLink 1000 eyetracker was the model used. This model required that the participant rest his chin on a chinrest at approximately 55 centimeters from the camera and 75 centimeters from the computer screen. Meanwhile, a desktop camera recorded the participant's eye movements. Before recording and during each break, the eyetracker was calibrated and validated.

In both the reading and the listening tasks, I tested participants on their processing of sentences with preverbal clitic object pronouns and postverbal subjects as well as on their processing of SVO sentences. For the reading task, the participants read 85 sentences in Spanish. First, they saw a fixation point on the screen, and then the entire sentence, which they read at their own pace. When they were done reading, they were instructed to look at a gray box located in the bottom right corner of the screen. This indicated that they were done reading and signaled to the computer to remove the sentence from the screen and present four pictures, which served as the comprehension question. Then, with the mouse, the participants chose the picture that they thought best described what they had just read. For the listening sentence processing task, the participants heard 85 sentences in Spanish, spoken by a native speaker, through headphones. The aural stimuli were pre-recorded using a Marantz professional portable solid-state recorder and Audio-Technica ath-m40fs studiophone microphone. First, the participants saw a fixation point on the screen. Then they saw a four-picture display, while they simultaneously heard the sentence.

For each experiment, the participants were presented with a total of five practice sentences, sixteen experimental sentences (eight per condition), and 64 filler sentences. The items were randomly presented. The experimental sentences for both tasks had two conditions: an SVO word order condition and an OVS word order condition, and were written in conjunction with a native speaker of Spanish. The SVO condition sentences were of the type: adverbial phrase, full noun phrase, verb, object marker, full noun phrase, prepositional phrase, for example, in (47).

(47)	*Por la tarde*	*el muchacho*	*llama*	*a*	*la muchacha*	*en la oficina.*
	[In the afternoon	the boy	call-3RDSINGPRE	to	the girl	in the office]

In the afternoon the boy calls the girl in the office.

The OVS condition sentences were of the type: adverbial phrase, direct object pronoun, verb, full noun phrase, prepositional phrase, for example, in (48).

(48)	*Por la tarde*	*lo*	*busca*	*la mujer*	*en la iglesia.*
	[In the afternoon	3RDSINGMASCACC	look for-3RDSINGPRE	the woman	in the church]

In the afternoon the woman looks for him in the church.

I included the extra material before and after the regions of interest (the clitic pronoun, the full noun phrase for the object and the full noun phrase for the subject) such as the prepositional phrases “por la tarde” and “en la iglesia” in (48) so as not to make the critical regions more salient than the rest of the sentence to the participant. A complete list of Spanish stimuli for both tasks is included in Appendices E and G.

All of the subjects and objects in the sentences were animate and singular so as to control for any animacy or subject-verb agreement cues for subjecthood, isolating word order and case as the only cues. The verbs were highly frequent and regular, and taken from the Spanish language textbook *Mosaicos*, used in the basic foreign language courses at the university from which the majority of the L1 English participants hailed. Both the listening and the reading task had a unique set of stimuli that included unique verbs. In addition, I controlled for the gender of the subject and object in order to not bias the participants’ processing toward a particular gender. The stimuli for each task included four sentences with the subject/object gender of female/female, four with female/male, four with male/female, and four with male/male. In addition to the experimental sentences, participants heard or read filler sentences that were similar in length and difficulty to the experimental ones, with one half of them created for a separate study on subject-verb agreement and the other half of them created for a separate study on tense morphology-adverb agreement, all of which were written in conjunction with a native speaker of Spanish.

For the listening task, while participants listened to the sentences, four pictures were presented simultaneously on a computer monitor. For the reading task, the participants saw the four pictures after they finished reading the sentence and had indicated that they were done reading by looking at the gray box in the bottom right hand corner of the screen. The same artist created all of the 128 (64 per task) pictures in order to maintain continuity in style. The pictures were black line drawings on a white background. Each picture was modified to a standard resolution and size, of the dimensions 280 x 280, and were of the format Tagged Image File Format (TIFF). For every sentence (both conditions) in both experiments, there were four picture conditions. Two of the conditions were the critical pair, in that they included the correct action and actors, and two were total distractors in that they had either incorrect actions or actors. The four conditions are shown in Table 4.

Table 4. *Experimental Conditions of Pictures*

CONDITION	**FOR THE SENTENCE:** *Por la noche la abuela busca al abuelo en la calle.* (At night the grandmother looks for the grandfather in the street)
1 (correct, critical pair)	Semantically Congruent and Grammatically Congruent: A picture of a grandmother looking for a grandfather
2 (incorrect, critical pair)	Semantically Congruent and Grammatically Incongruent: A picture of a grandfather looking for a grandmother
3 (incorrect, distractor)	Semantically Incongruent and Grammatically Congruent: A picture of a child looking for a grandfather
4 (incorrect, distractor)	Semantically Incongruent and Grammatically Incongruent: A picture of a grandfather looking for a child

For Condition 1, the correct picture, the pictures contained the same actors and action as the sentence stimulus and the actors had the same roles, patient and agent, as the sentence stimulus. For example, if the sentence were, *Por la noche la abuela busca al abuelo en la calle* (At night the grandmother looks for the grandfather in the street), then there would be a picture of a grandmother looking for a grandfather, as in Figure 4.

Figure 4. *Picture Condition 1: Correct Picture*

The pictures for Conditions 2, 3, and 4, were the incorrect pictures. For Condition 2, the actors switched roles, but the action was the same as the sentence stimulus. Using the example sentence above, this would be a picture of a grandfather looking for a grandmother, as shown in Figure 5.

Figure 5. *Picture Condition 2: Incorrect Picture, Critical Pair*

I considered the Condition 2 pictures part of the critical pair with the Condition 1 pictures since they were the exact inverse of each other. I hypothesized that participants would most likely choose this condition when incorrect. The pictures for Conditions 3 and 4 were the total distractors. For Condition 3, the roles were the same, but the actors or actions were different from the sentence stimulus. Using the example sentence above, the picture might be of a child looking for a grandfather, as shown in Figure 6.

Figure 6. *Picture Condition 3: Incorrect Picture, Distractor*

Finally, for Condition 4, neither the roles nor the actors or actions were the same as those in the sentence stimulus. Using the example sentence above, this might be a picture of a grandfather looking for a child, as shown in Figure 7.

Figure 7. *Picture Condition 4: Incorrect Picture, Distractor*

I hypothesized that participants would most likely not choose either one of these pictures unless there was a breakdown in their general comprehension of the sentence. The complete set of pictures from both experiments can be found in Appendices F and H.

In order to not draw attention specifically to the people involved in the action or to the action itself, I altered the semantic change in Conditions 3 and 4 throughout the stimuli in both tasks such that half of the sentences had changes of subject and half had changes in the verb. For example, in the case of the sentence, Conditions 3 and 4 could have a verb change (e.g. *baila* instead of *busca*) or a subject change (e.g. *la nieta* instead of *la abuela*), as shown here. Genders were kept constant. Correct pictures were counterbalanced so as not to create a bias for a particular picture location.

Furthermore, I conducted a separate norming study to ensure that the pictures were showing what I wanted them to be showing. In this study, 107 native English speakers wrote what they thought one person was doing to the other in each of the 64 listening pictures. The results revealed that participants ignorant of the sentence stimulus were able to correctly identify the actors and patients at or above 70% of the time for 48 out of 64 (75%) of the pictures. Due to the more flexible nature of the data analysis, which involved linear regression with mixed-effects models, I was able to build item in as a random effect, and therefore could see the effect of each item on the results. In addition, it is reasonable to keep all items in the final analysis because every participant saw the same items (i.e. the same pictures), which should in turn moderate the effect of any one picture or item.

2.4.3 Cognitive Individual Difference Measures

The two cognitive individual difference measures included in the present dissertation were the Letter-Number Sequencing Task (LNS), which has been shown to tap into the capacity component of working memory, and the Flanker Task, which has been shown to tap into a more executive attention/attentional control-type component. In section 2.4.1.1, I describe the LNS task and my rationale for using it in the present dissertation. In section 2.4.1.2, I describe the Flanker Task and my rationale for using it.

2.4.3.1 The Letter-Number Sequencing Task

All participants performed the Letter-Number Sequencing (LNS) test after the second of the two sentence processing tasks. This particular test was adapted from a subtest from the revised version of the Wechsler Adult Intelligence Scale (WAIS) (Wechsler, 1997). I chose this particular assessment of working memory since it allowed me to test each participant on the same exact task in the L1, thus minimizing potential confounds with L2 proficiency level. In this test, participants were asked to remember strings of numbers and letters shown on a computer screen, rearrange them so that the numbers were first and in numeric order and the letters were next in alphabetic order, and then type them into a computer. The participants were told that accuracy was more important than speed in this task. This task took approximately ten minutes to complete. Participants were given as much time as they needed to recall and type the sequences once they had been presented. The results from this task were then scored for overall accuracy (total number recalled correctly out of 21) and memory length (the length of the longest string that was recalled correctly out of 8). Although the two measures correlated highly with each other (Pearson correlation of .829, with $p>.01$), only the overall accuracy was used in the final analysis since it provided a more fine-grained distinction among participants, following previous research that has shown that the raw total score most highly correlated with intelligence testing (Crowe, 2000).

2.4.3.2 The Flanker Task

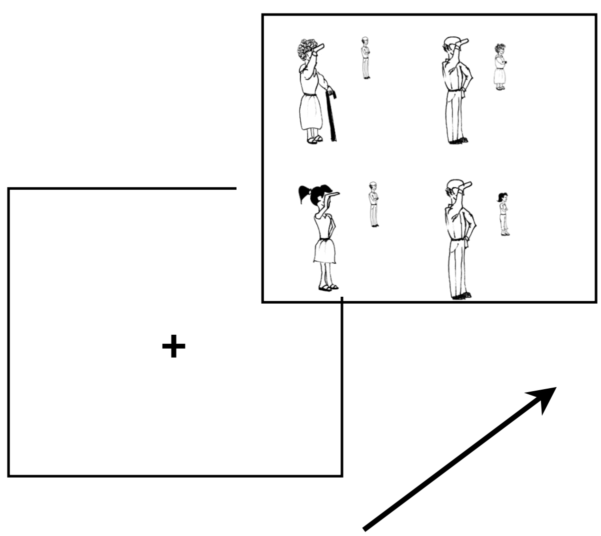
In addition to the LNS task, participants completed an executive attention test (Flanker Task) in between the two sentence processing tasks that assessed their ability to control their attention to task-relevant information and to block out extraneous information during processing. This task consists of a series of figures made up of arrows and diamonds in which one arrow is colored red. The participants saw each series on a computer screen, and were asked to click the mouse button that corresponded with the direction in which the red arrow was pointing, ignoring any cues from the other figures. They were told that both accuracy and speed were equally important in this task. This task took between five and eight minutes to complete. This task has been shown to draw out differences in executive attention among participants (e.g. Luk, 2008). The results from this task were then scored for facilitation and inhibition effects of reaction times on congruent and incongruent trials, which I will discuss in greater detail in the data scoring, descriptive statistics, and results sections.

2.5 Procedure

The total time for participants to complete the experiment was no more than two hours, depending on how many tasks they completed. For example, the L1 English participants did not have to complete the English proficiency test, and the L1 Spanish participants did not have to take the Spanish proficiency test; thus these two groups took less time overall to complete the experiment. The procedure was as follows: first the participants were asked to sign a consent-to-participate form that explained the study, their rights, and what would be expected of them. After this, they spent ten minutes familiarizing themselves with the vocabulary used in the experiments (in the case of the learners), and then completed a language history questionnaire (LHQ). Next, they took the appropriate proficiency tests at their leisure (a Spanish proficiency test for all of the participants except the monolingual control group, and an English proficiency test for all of the participants except the L1 English learners). Then, participants performed either the reading or the listening sentence processing task, which took 20-40 minutes and 10-20 minutes, respectively, to complete. In both sentence processing tasks, the participants responded to the sentences by choosing the picture that was best described by the sentence that they heard with the click of a mouse. This is similar to what Isabelli (2008) and VanPatten (1984) used as comprehension checks in their experiments, except that these researchers only provided participants with two pictures from which to choose. I included the additional two pictures so that it was clear whether the participants had understood the entire sentence, and that they were not just focusing on the order of the subject and object constituents to make their judgments. Illustrations of the sequence of screens for the listening and reading tasks are provided in Figures 8 and 9.

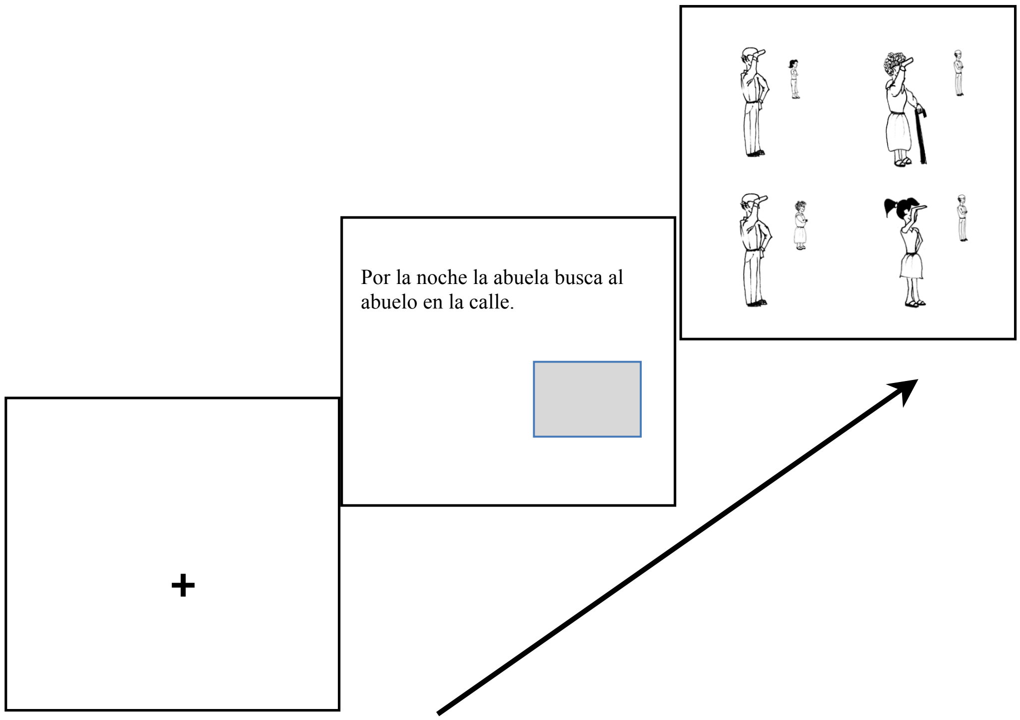
The participants who were not L1 English speakers also took an English proficiency test. However, the results from this measure will not be part of the final analysis. I mention the results from the English proficiency test in the conclusion in chapter 5.

Figure 8. *Sample Screens of Listening Sentence Processing Task*



After seeing the fixation point as shown above, participants heard, *Por la noche la abuela busca al abuelo en la calle*, as they saw the four line drawings on the screen. For the reading task, participants saw two screens after the fixation point, as shown in Figure 8. The first contained the stimulus sentence with a grey box in the bottom corner. The second contained a four-picture display.

Figure 9. *Sample Screens of Reading Sentence Processing Task*



After the first sentence processing task, participants took the Flanker test to assess their ability to inhibit extraneous information while performing a task. Illustrations of the three types of stimuli for this task are provided below in Figures 10, 11, and 12.

Figure 10. *Sample Screen of Incongruent Trial in the Flanker Task*

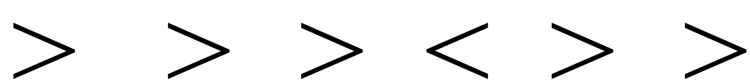


Figure 11. *Sample Screen of Congruent Trial in the Flanker Task*

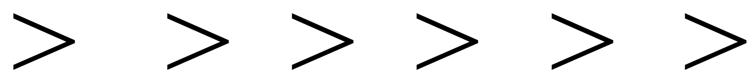
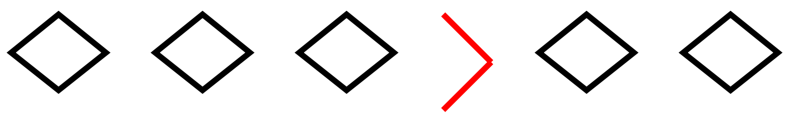
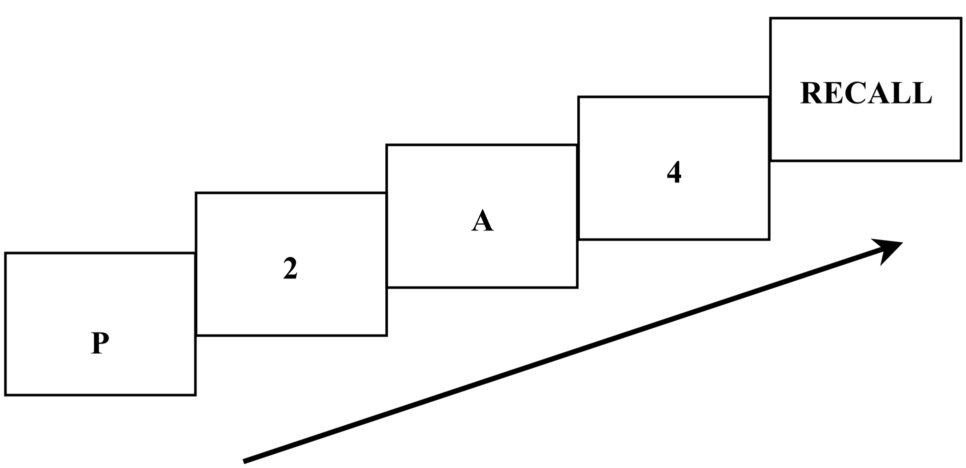


Figure 12. *Sample Screen of Go Trial in the Flanker Task*



The Flanker task also served as a break for participants between the two sentence processing tasks. After the executive attention task, participants completed the sentence processing task that they had not yet completed (either listening or reading). Subsequently, participants took a working memory test (a letter-number sequencing test) to gauge their level of working memory ability. This task was used since it employs the same stimuli for all participants regardless of language. An illustration of a sequence of screens from this task is provided in Figure 13.

Figure 13. *Sample Screens of Working Memory Test.*



Finally, all of the learners of Spanish took a vocabulary test at the end of the experiment to ensure that they had understood the words used in the tasks, and to control for any possible misunderstanding of materials due to a lack of vocabulary knowledge.

2.6 Data Scoring

The assessment measures provided information that was tallied and recorded. The language history questionnaire data provided information about the participants' age, sex, amount of time immersed in a Spanish-speaking country, and length of study of Spanish. I ran correlations on these factors in the analysis reported in the following section to see if any of them would share the variance with the variable of proficiency. Based on this, the amount of time immersed in a Spanish-speaking country was included as a variable in the analysis. Individual difference information was also tallied in the following categories: proficiency in Spanish for the learner participants and cognitive individual difference measures. Proficiency and working memory scores were significant predictors in the mixed effects model discussed in the following section.

The language proficiency tests objectively examined how familiar each learner participant was with her L2. For this test, participants received one point for each correct answer, and zero points for each incorrect answer. For the Spanish proficiency test, there were 68 items. These results were then scored as a percentage out of 100, and participants must have scored 88% or higher to be included in the study. For the English proficiency test, there were 20 items on this test. Participants received one point for each correct answer, and zero points for each incorrect answer. These results were then scored as a percentage out of 100. As mentioned before, the English proficiency test results will not be part of the final analysis.

For each of the sentence processing tasks, accuracy was analyzed for the 16 experimental items, and scores for overall accuracy and accuracy by condition were computed. Participants received 1 point for each correct picture selection. Wrong answers were also analyzed for patterns. Two participants were missing data for the listening sentence processing task and one participant was missing data for the reading sentence processing task. This was taken into account in the mixed-effects model described in the following section.

For the LNS test, a participant's working memory capacity was his raw score. Each participant was given a raw score, which was "all or nothing" for each item in the test. In other words, the participants received one point for each set of items that they typed correctly and zero points if they made one or more errors. A perfect score was 21. When scoring this test, due to the indistinguishable nature of the numeral "0" and the letter "o", I accepted these characters interchangeably as long as the participant wrote them in the correct order according to the rules of the task. Using the raw score has been shown to be the most reliable, and most highly correlated with intelligence testing (Crowe, 2000). Two participants, one L1 Romanian and one L1 English participant, were eliminated from the analysis due to missing or incomplete data.

For the Flanker test, accuracy means were only used as a way to eliminate outliers. I excluded one participant from the analysis because this participant's accuracy score was below 80%. Two additional participants were eliminated from the analysis due to missing or incomplete data. More important, the remaining participants were scored for their reaction times (RT) on each type of trial. First the outlier trials were removed by averaging all of the RT values for all of the participants, finding the Standard Deviation (SD), and removing all trials that were more than 2.5 SD (and over 1500 msec. as an absolute outlier value) or less than 2.5 SD (or 50 msec., an absolute outlier value) from the mean. The trials that followed incorrect trials, or recovery trials, were eliminated before the RT analysis as well because it has been shown that participants have a tendency to rush their response after an error (see Luk, 2008). The cleaning of the data was done with a script in the open source statistical programming environment R (Bates, 2007). The average RTs for each participant for each type of trial (congruent, go, incongruent) were then calculated in R. Then, I subtracted the mean of the RTs for the go trials from the mean of the RTs for the incongruent trials to get the RT Inhibition effect per participant. The RT Inhibition effect is the measure that reflects the degree to which a participant's ability to inhibit extraneous non-task-relevant information does not slow down her decision-making process when faced with a situation in which there is a conflict between the task goal and information given, such as when the flanker arrows and the target arrow point in the opposite direction, as opposed to a neutral situation, such as when there are diamonds surrounding the target arrow. Therefore, the smaller the number that a participant has for the RT Inhibition effect, the less the difference is between a participant's reaction times on incongruent trials and on neutral trials, which suggests a better ability to focus on task-relevant information and block out or inhibit non-task-relevant information. This measure has been known to show the most differences between bilinguals and monolinguals (see Luk, 2008).

For the Flanker task, I also subtracted the mean of the RTs of the congruent trials from the mean of the RTs for the go trials to get the RT Facilitation effect per participant. The RT Facilitation effect is the measure that reflects the degree to which a participant's ability to incorporate task-relevant information speeds up her decision-making process when faced with a situation in which there is not only no conflict between task goals and the information provided, but when there is also information present that reinforces the task goals, such as when the flanker arrows point in the same direction as the target arrow, as opposed to a neutral situation. Therefore, the RT Facilitation effect should be an even smaller number than the RT Inhibition effect. Finally, I found the Flanker Effect, which is the difference between each participant's mean RTs of the incongruent trials and his mean RTs of the congruent trials. It is assumed that participants will perform better on the congruent than the incongruent trials in the first place, regardless of executive attention abilities. If there is any RT facilitation provided by the presence of flanker arrows reinforcing the same information as the target arrow provides, then this would also predict overall smaller values for the RT Facilitation effect. Although participants with better executive attention would be expected to have the shortest RTs overall on correct trials, the RT Facilitation effect is less telling than the RT Inhibition effect since it is unclear whether or not those participants who have better executive attention would have the smallest values. If these participants are better able to include the additional information while they are processing the congruent trials, and thus process these trials more rapidly, then the smallest values would be predicted. However, if the participants with the best executive attention abilities are better at blocking out all irrelevant information, then they may not even process the flanker arrows' direction, whether it is congruent or not, and therefore, may show small, but not the smallest values for the RT Facilitation effect (see Luk, 2008). Neither the Flanker Facilitation scores nor the Flanker Effect scores were found to be significant predictors in the model in the following section, nor did they correlate strongly with the Flanker Inhibition scores.

2.7 Descriptive Statistics

2.7.1 Descriptive Statistics for the Assessment Measures

First I report the means and standard deviations for the learners' performance on the Spanish proficiency test (2.7.2) and the amount of time learners spent immersed in a Spanish-speaking country (2.7.3) since these two variables are important in the final analysis. Then, I report a series of correlations between the learner participants' background data such as age and time spent studying Spanish and proficiency scores (2.7.4). Finally, I report the descriptive statistics for the learners' performance on the cognitive individual difference measures (2.7.5).

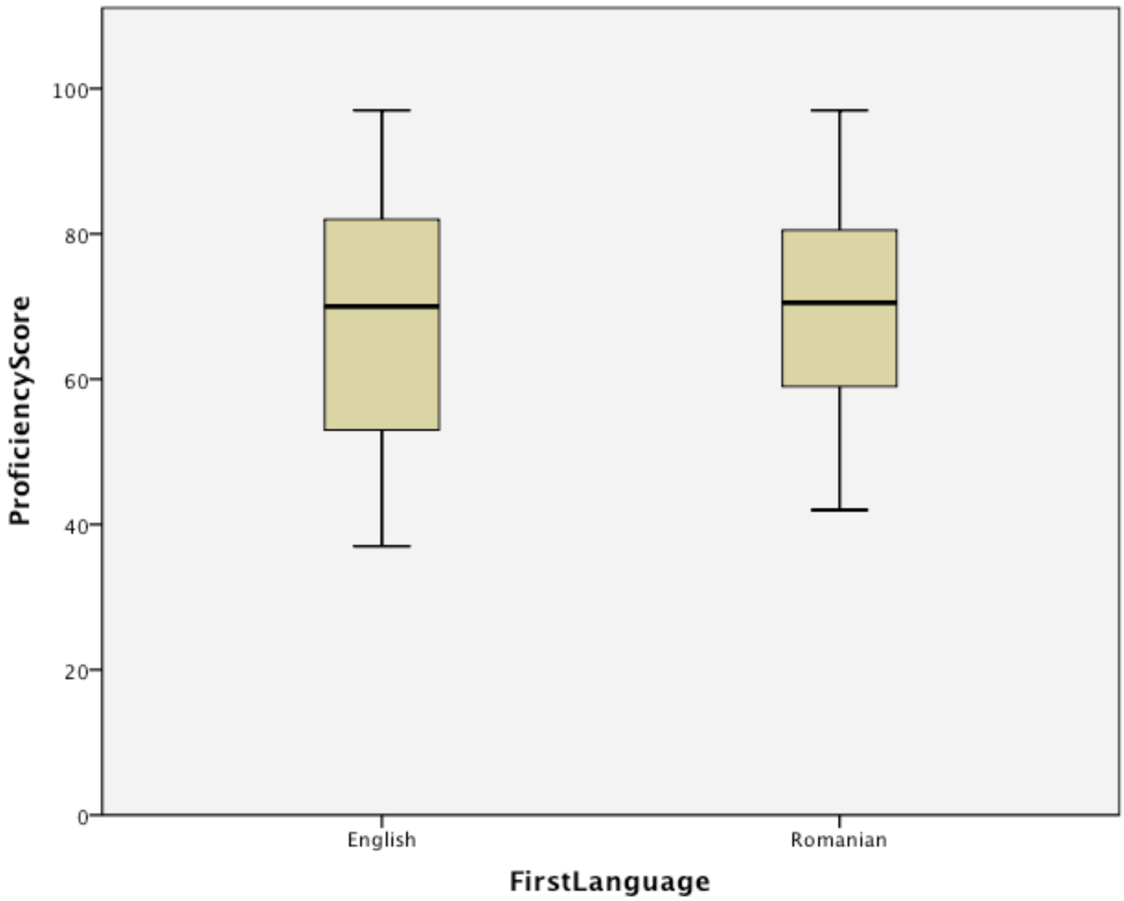
2.7.2 Spanish Proficiency Means

The overall mean for proficiency score on the adapted DELE test was 69.06 out of 100 with a SD of 15.17 for the 137 learner participants. The means and SDs by language group can be found in Table 5. The means and SDs are plotted in Figure 14.

Table 5. *Means for Proficiency by Language Group*

First Language Group	N	Mean	SD
English	65	68.11	16.44
Romanian	72	69.92	13.99

Figure 14. *Distribution and Variance of Proficiency Scores by First Language*



An independent-samples t-test revealed that the two learner groups were not significantly different in proficiency means. The results from the t-test are in Table 6.

Table 6. *T-Test Results for Proficiency Scores*

	T score	df	Sig. (2-tailed)
Proficiency Scores	-0.70	135	.49

2.7.3 Immersion Time Means

The overall mean for the learner participants' immersion time in a Spanish-speaking country as measured in months is 4.38 with a SD of 9.3 for the 137 participants. The means by first language group are shown in Table 7.

Table 7. *Immersion Time Means by First Language*

First Language Group	N	Mean	SD
English	65	6.92	11.90
Romanian	72	2.26	5.33

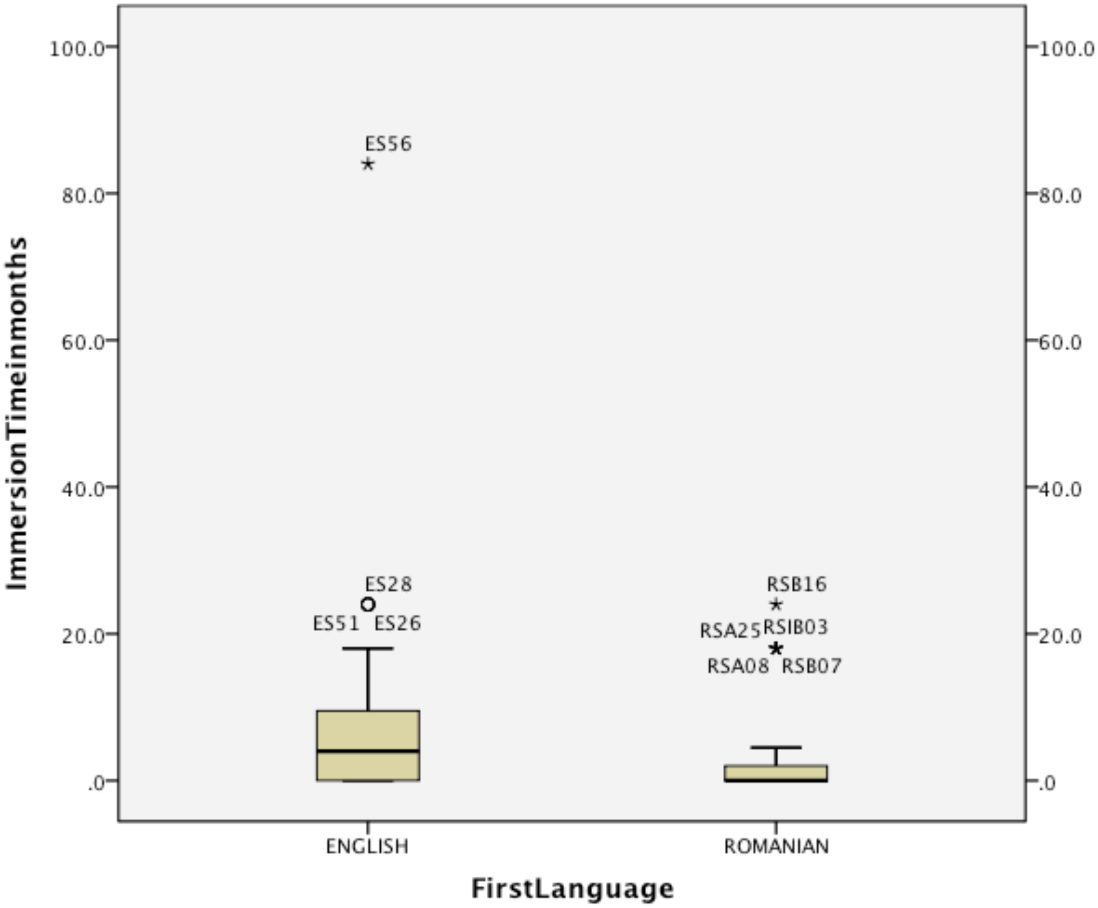
These means appear to be different from each other. The English-Spanish learner group on average had spent more than three times as many months in an immersed Spanish-speaking setting as the Romanian-Spanish participant group. I ran an independent-samples t-test on the means and found that the two groups are significantly different. The results from the t-test are in Table 8.

Table 8. *T-Test Results for Immersion Time*

	T score	df	Sig. (2-tailed)
Immersion Time	3.01	135	.003

In order to better understand the distribution of this variance, I created a box plot. The distribution can be seen in Figure 15.

Figure 15. *Distribution and Variance of Immersion Time Among Learner Participants*



The majority of the participants had less than 3 months of immersion time. There were 67 learners with no time abroad (18 of whom were from the English-Spanish group). One English-Spanish participant had spent 7 years abroad. I did not eliminate this participant from the study, however, because her proficiency and accuracy data matched that of a Romanian-Spanish participant. In order to double check, in the analysis, I ran the mixed-effects models both with and without these two participants and found that excluding them did not change the pattern of results, nor did it substantially change the predictor coefficient estimates.

2.7.4 Correlations between Proficiency Scores and Background Data

Here I present the correlations between the learners' Spanish proficiency scores (DELE exam scores and self-ratings) and the individual variables of age, immersion time, and years spent studying Spanish. Of the variables included here in the correlations, the most reliable predictors for accuracy were the DELE proficiency scores and immersion time. Therefore, these are the only two measures included in the final analysis.

The calculation for years spent studying Spanish was the difference between the age of first study of Spanish and the participant's age at the time of the experiment. This is not the most accurate nor reliable measure of length of study, even though it has been used in prior studies as a measure of proficiency (e.g. Grüter & Crago, 2010). It is included in the correlations with a caveat that although the participants reported the age at which they began studying Spanish, this does not necessarily mean that they had studied it continuously up to the time of the experiment or in the same way as the other participants (e.g. type and frequency of classes). Due to concerns regarding the unreliability of this variable as a predictor, as mentioned previously this measure is not included in the final analysis. The average of participants' self-ratings on four areas of proficiency (reading, writing speaking, and listening) was also calculated and included in the correlations. Many studies rely solely on self-ratings for proficiency measures (see Blanche and Merino, 1989), so it is important to see how these correlate in the present study with the independent measure of proficiency. It is crucial to note that these self-ratings were based only on a 5-point scale rather than a 10-point scale, which may limit variability among the scores. The results from the Pearson correlations run in SPSS are shown in Table 9.

Table 9. *Correlations between Proficiency Scores and Background Measures*

	Proficiency Score	Age	Avg. Self-Ratings	ImmersionTime in months	YearsStudied Spanish
Proficiency Score					
Correlation	1	0.17	0.43**	0.28**	0.19*
Sig.		0.06	0.00	0.00	0.04
N	137	131	136	137	122
Age					
Correlation	0.17	1	0.01	0.42**	0.31**
Sig.	0.06		0.87	0.00	0.00
N	131	162	130	131	122
Avg. Self-Ratings					
Correlation	0.43**	0.01	1	0.20*	0.57**
Sig.	0.00	0.87		0.02	0.00
N	136	130	136	136	121
ImmersionTime					
Correlation	0.28**	.42**	0.20*	1	0.42**
Sig.	0.00	0.00	0.02		0.00
N	137	131	136	137	122
YearsStudiedSpanish					
Correlation	0.19*	0.31**	0.57**	0.42**	1
Sig.	0.04	0.00	0.00	0.00	
N	122	122	121	122	122

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

As shown in Table 9, the proficiency scores from the adapted DELE exam correlate significantly and somewhat strongly with the average of the self-ratings. Assuming the validity of the independent measure of proficiency (the DELE), the correlation coefficient of 0.43 is not robust enough, however, to rule out using the DELE exam scores. If the self-ratings and DELE scores were more strongly correlated, then I could have used either one as the measure of proficiency. However, given the subjectivity and small scale of the self-ratings, I only include the DELE exam scores in the final analysis. The significant, but weak, correlations between the proficiency scores and immersion time and years spent studying Spanish also reinforce the importance of having an independent, objectively quantifiable measure of proficiency[6].

[6] Note that the question of how the different participants arrive at their respective levels of proficiency is beyond the scope of this study and will be left for future research.

It is reasonable that there is a moderately strong correlation between the measures of immersion time and years spent studying Spanish since most learners do not study or live abroad with little time first spent studying the language. It is important to remember that the variable for time spent studying Spanish was calculated by subtracting the participants' age of first exposure from their current age, which is not the most reliable measure due to discrepancies in how continuous and how regular one's study may have been over the course of the years included in this calculation. For example, a participant may have started studying Spanish in high school, but then not studied it for a few years, but then may have gone abroad for a job and been immersed again in Spanish. This participant could look the same, according to this calculation, as a participant who started studying Spanish in high school as well and continued to study Spanish in college and beyond. As evidenced in this case, the variable of years spent studying Spanish is not always an accurate depiction a participant's experience, and therefore, will not be included in the final analysis.

Other significant correlations include the correlations between the measures of immersion time and years spent studying Spanish and the measure of age, which is logical. More interesting, the strongest significant correlation is between years spent studying Spanish and average self-ratings. There may be some influence here of belief or even increased confidence in one's Spanish abilities that is gained with time dedicated to studying Spanish. Again, this is of no surprise.

2.7.5 Descriptive Statistics for Cognitive Individual Difference Measures

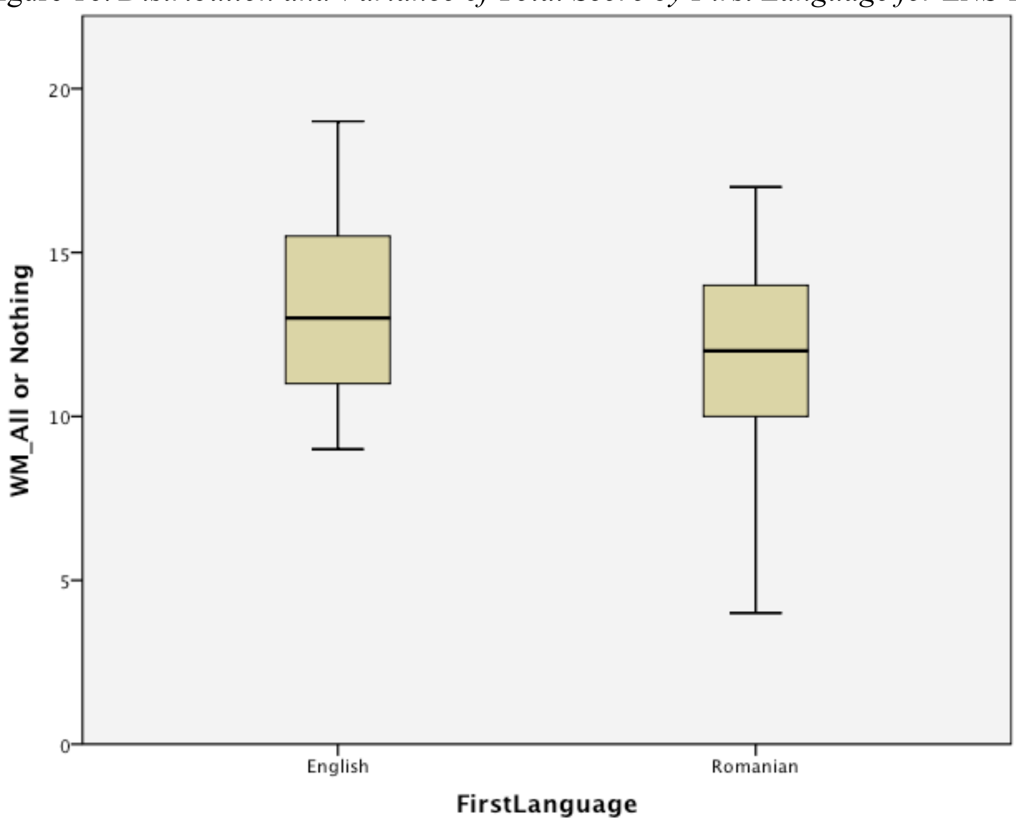
2.7.5.1 Letter-Number Sequencing Descriptive Statistics

The Letter-Number Sequencing (LNS) data for one participant in each language group were missing due to technological error. The total score was out of a possible 21 points, and the memory length score was out of a possible 8 points. I report the means and standard deviations in Table 10 for both measures in the LNS task by language group. The variance of the two groups can be seen in Figure 16.

Table 10. *LNS Means by First Language*

First Language	Measure	N	Mean	SD
English	Total Score	64	13.52	2.84
	Memory Length	64	6.61	1.09
Romanian	Total Score	71	11.85	2.69
	Memory Length	71	5.94	1.22

Figure 16. *Distribution and Variance of Total Score by First Language for LNS Task*



The average total score for the English-Spanish learner group was higher than the mean score of the Romanian-Spanish group. An independent-measures t-test on both the total scores and the memory length scores revealed that the two language groups were significantly different. The results from the t-test are in Table 11.

Table 11. *T-Test Results for LNS Measures*

	T score	df	Sig. (2-tailed)
Total Score	3.51	133	0.00
Memory Length	3.33	133	0.00

The fact that the English-Spanish group had higher LNS measure means than the Romanian-Spanish group lends me to believe that there may be a cultural difference in familiarity with this type of task. These two groups are most certainly matched on proficiency, but they may not be matched on other less tangible variables such as working memory capacity.

2.7.5.2 Flanker Inhibition Descriptive Statistics

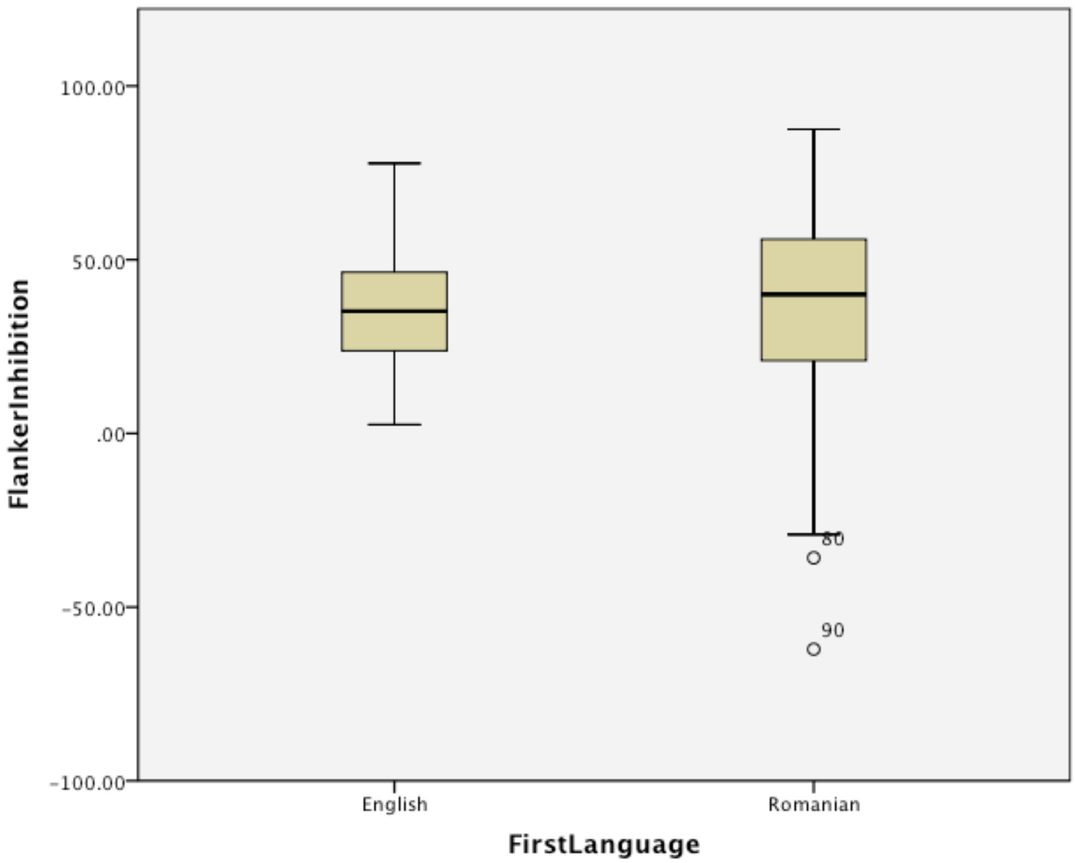
Here I report the means and standard deviations from the Flanker Inhibition measure. This measure was the difference in mean RTs between the neutral trials and the incongruent trials by participant, which has been shown to measure executive attention in bilinguals (see Luk, 2008). The means and SDs for the learner groups can be found in Table 12.

Table 12. *Flanker Inhibition Means by First Language*

First Language	N	Mean	SD
English	64	35.26	17.86
Romanian	70	34.55	29.91

The data for two L1 Romanian participants and one L1 English participant were not included in the analysis due to technological error. Anticipating the analysis below, the measures of Flanker Facilitation and Flanker Effect were not found to be significant predictors in the models, nor have they been shown to indicate advantages in bilinguals (see Luk, 2008). Therefore, I do not report these measures here nor do I include them in the final analysis.

Figure 17. *Distribution and Variance of Flanker Inhibition Scores by First Language*



As shown in Figure 17, the learner groups appear to be quite similar in their ability to inhibit extraneous information, with slightly more variance in the L1 Romanian group. An independent-samples t-test showed that this was the case statistically as well. The results from this test are in Table 13.

Table 13. *T-Test Results for Flanker Inhibition*

	T score	df	Sig. (2-tailed)
Flanker Inhibition	0.17	132	0.87

As far as cognitive learner differences that may affect the ability to process the two structures that are central to this dissertation, it is important to note that the two first language groups are not significantly different in their executive attention abilities. The differences in group means for the LNS task are of interest. Therefore, I will address this issue and its implications in chapter 5 after presenting the analysis of the accuracy data. My predictions for the accuracy data are that I expect to see better accuracy for those participants with higher WMC and higher Flanker Inhibition scores, regardless of first language. I also expect to see better performance on the OVS (or more challenging) structures for those with higher WMC and better executive attention since they would be better able to maintain the crucial information in working memory, attend to the preverbal clitics while still processing the rest of the sentences, and therefore, be more accurate when finally choosing the picture that they thought was best depicted by what they had read or heard. These learner results in regards to the effect of WMC and executive attention on accuracy are found in the results sections in chapters 3 and 4, when I add these cognitive variables to the mixed-effects models.

2.8 Analyses Procedures

For this dissertation, I used logistic regression with mixed-effects models. As is becoming increasingly clear in the literature, logistic regression is better suited to analyze proportional data such as categorical outcomes (i.e. accuracy scores), which is what the present dissertation incorporates, than the more canonical ANOVA (see Jaeger, 2008). According to Jaeger's (2008) analysis, ANOVAs on proportional data may produce confidence intervals that are not interpretable, and can lead to spurious results. Using an actual psycholinguist data set, Jaeger (2008) showed that ANOVAs produced an interaction between conditions that was not there when logistic regression was employed. With binomially distributed conditions, ANOVAs assume that variances are homogeneous, which is not typically the case. Not only is logistic regression more accurate in that it provides information on the directionality and effect size of categorical data, it also allows researchers to keep every trial in the analysis, without having to eliminate participants from the entire analysis if they are missing trials as is necessary for ANOVAs.

There are additional benefits of using mixed-effects models with logistic regression. Mixed-effects models, which are a type of generalized linear mixed model, have become increasingly more prevalent in language processing research. First of all, mixed-effects models allow the researcher to model random subject and item effects because they illustrate outcomes as the linear combination of fixed and random effects. In addition, in mixed-effects models, all of the continuous variables are allowed to be continuous (such as the variables of proficiency and working memory here) and all of the discrete (or categorical) variables are allowed to be discrete (such as the variables of condition and first language here) in the analysis. So with mixed-effects models no gerrymandering of the variables is necessary to make the data fit the statistical model. Mixed-effects models allow for interactions between discrete and continuous variables as well. Not only do mixed-effects models offer a more complete picture of the data, but also, more importantly, they can be more accurate than ANOVA in that there is a reduced chance of the model being overfitted to the sample, and thus an increased chance of the model being more generalizable to the entire population (see Jaeger, 2008 for a detailed argument). In short, being able to combine subject and item analyses into one model is preferable and effective. This makes logistic regression with mixed-effects models the most appropriate method of analysis for the data of the present dissertation.

The data gathered for the present dissertation were analyzed using a linear regression model with mixed-effects built in the statistical environment R (R development core team, 2005). I built the independent variables (predictors) into the model with the dependent variable of participants' performance accuracy on the comprehension portion of the sentence processing tasks. The predictors that were fixed effects in the present analysis were first language, sentence condition, proficiency level, working memory total score, Flanker inhibition, and time spent immersed in a Spanish-speaking country. The random effects considered were the individual participants (subject) and the individual items (sentence). I elaborate on the results from the models and their implications in the following two chapters.

CHAPTER 3: Listening Task Results

3.0 Introduction

In the present chapter, I provide the results of the analysis of how first language and individual differences affect how accurately L2 learners of Spanish auditorily process sentences with SVO word order and OVS word order (preverbal clitics and postverbal subjects). I explain each of the mixed-effects models that were created for the listening task. First, I present the descriptive statistics for the aural sentence processing task and the predictor variables for the mixed-effects models (3.1). Then, I present the models (3.2 for monolinguals and 3.3 for learners). Finally, I provide a summary and a discussion of the listening task results (3.4).

3.1 Descriptive Statistics for the Listening Sentence Processing Task

Here I present the accuracy results from the listening sentence processing task. I include the Spanish monolingual data as well to serve as a baseline for comparison. Each task included 8 sentences in each condition (SVO and OVS). The accuracy was scored out of 100%. First, I report the means and SDs by language group and by condition for the listening task. Then, I report the means and SDs by language group and by condition for the reading task.

3.1.1 Accuracy on Listening Task

The accuracy means and SDs for the listening task by condition and language group is shown in Table 14.

Table 14. *Listening Accuracy Overall and by Condition for All Language Groups*

Language Group	Condition	N	Mean	SD
Eng-Span	SVO	65	0.88	0.11
	OVS	65	0.58	0.34
	Overall	65	0.73	0.20
Rom-Span	SVO	71	0.91	0.11
	OVS	71	0.68	0.31
	Overall	71	0.80	0.17
Span Mono	SVO	35	0.94	0.08
	OVS	35	0.90	0.10
	Overall	35	0.92	0.07

I have included the Spanish monolinguals in this table just as a baseline for comparison. The data for one Romanian and one Spanish monolingual participant were lost due to technological error. The distributions and variance for the overall accuracy and the accuracy on each condition by language group are shown in Figures 18-20.

Figure 18. *Listening Accuracy Overall Means for All Language Groups*

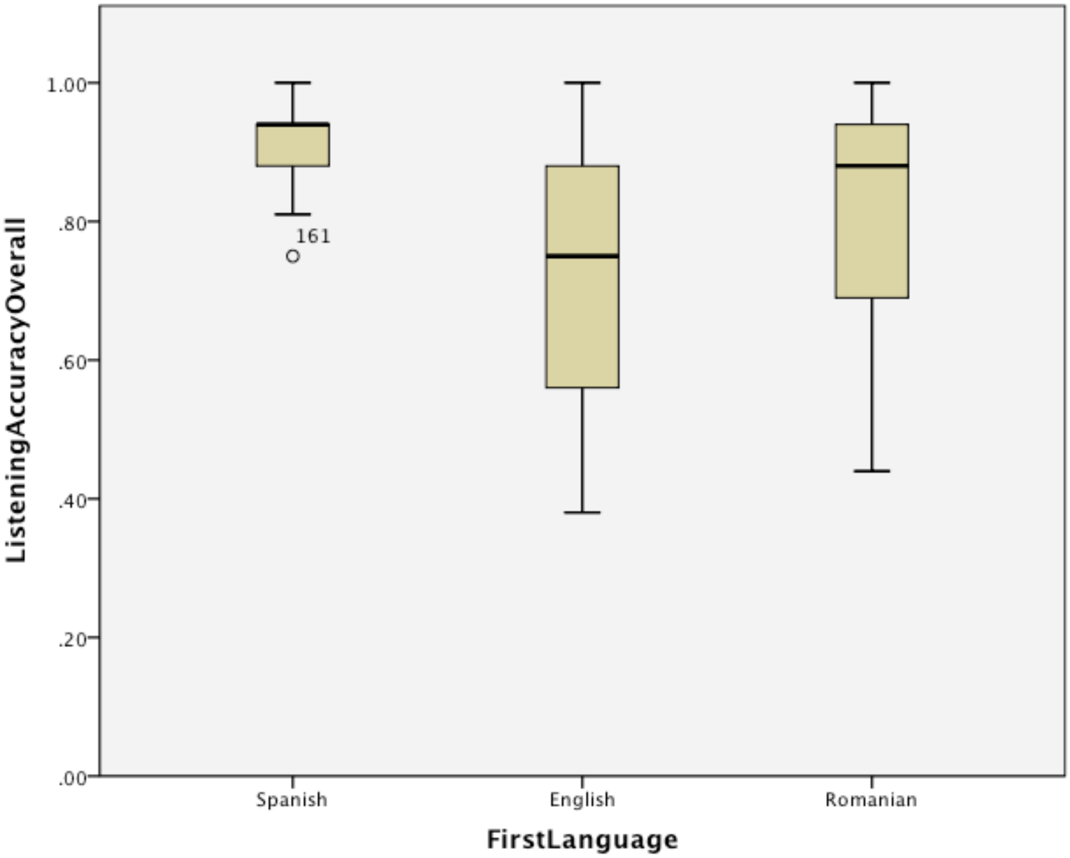


Figure 19. *Listening Accuracy SVO Means for All Language Groups*

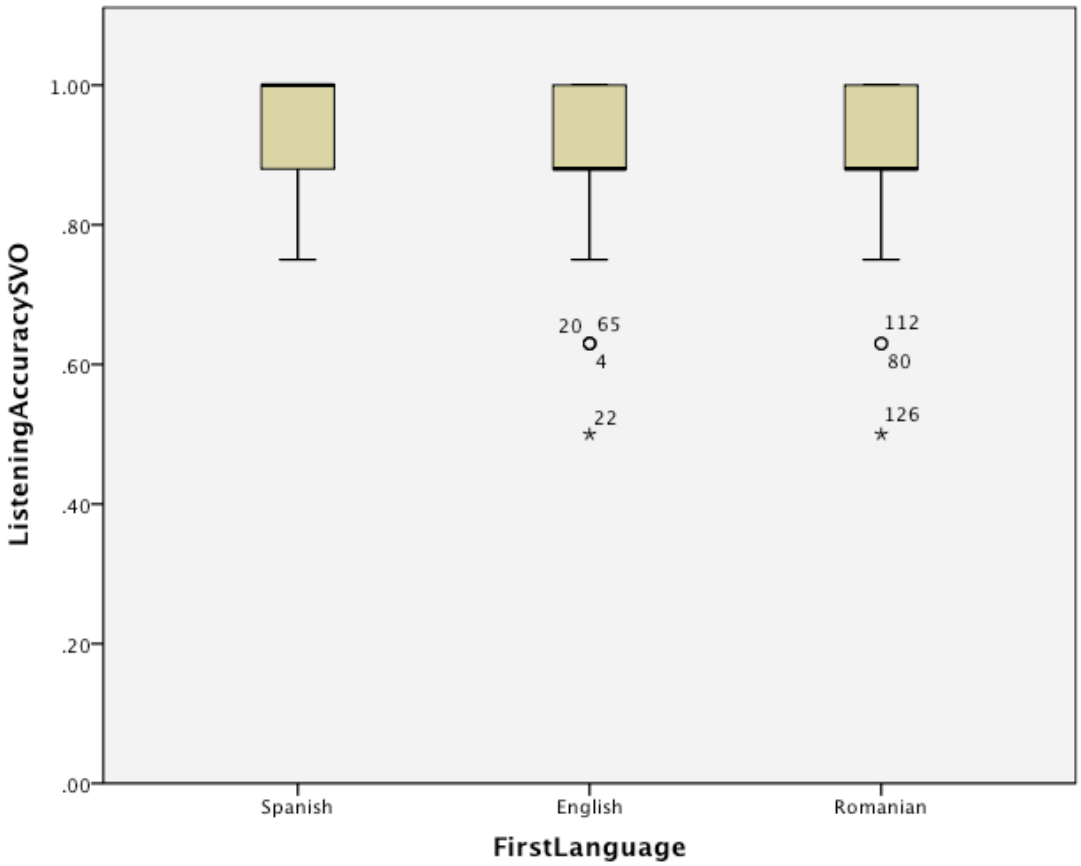
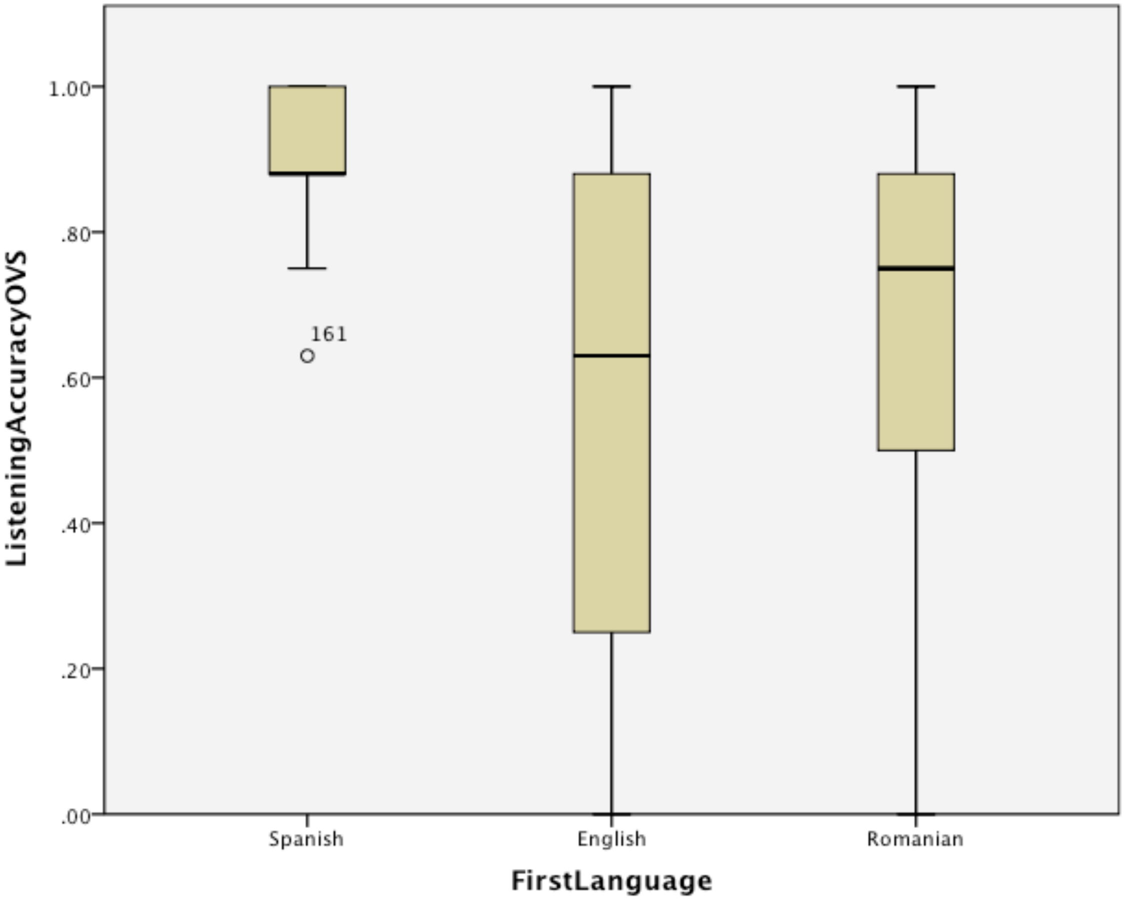


Figure 20. *Listening Accuracy OVS Means for All Language Groups*



As Figures 18-20 show for the listening task, the Spanish monolingual group was more accurate overall than the other two language groups, which is not surprising. Looking at the accuracy scores more in detail, it becomes clear that the source of the differences between the monolinguals and the learner groups is the OVS condition. The Spanish monolinguals outperformed both learner groups on the sentences with the preverbal clitics and postverbal subjects. It is interesting to note, however, that although the Romanian-Spanish learners did not differ much from either the Spanish monolinguals or from the English-Spanish learners on the SVO condition, there does appear to be a difference between the English-Spanish learners' mean performance and the Spanish monolinguals' mean performance on the SVO items. It is important to note that these are just the descriptive statistics, and a more detailed statistical analysis is provided below in the mixed models.

3.1.2 Predictors

I examined the participants' accuracy on the listening sentence processing task (a continuous variable based on the experimental sentences from these tasks) using logistic regression with a mixed effects model, fit by the Laplace approximation, which has been said to be extremely accurate (Harding and Hausman, 2007). The crossed random effects (variables specific to the present data) were for subjects (*SD* = 0.83 for the listening task for the learners) and for items (*SD* = 0.34 for the listening task for the learners). The predictor variables for this experiment can be found in Table 15.

Table 15. *Predictor Variables for Listening Task.*

Predictor Variables
Condition
Immersion Time
First Language
Working Memory
Flanker Inhibition
Proficiency

As shown in Table 15, I included in the model the fixed effects (independent variables) of proficiency in Spanish of the learner participants (a continuous variable based on scores from adapted DELE test), immersion time in a Spanish-speaking country (a continuous variable based on self-reports in the language history questionnaire), working memory total raw scores (a continuous variable based on LNS test), Flanker Inhibition scores (a continuous variable based on Flanker test), first language (a binary variable for the learners' data with English and Romanian as the two levels), and sentence condition (a binary variable for all the data with SVO and OVS as the two levels). As noted above, memory length scores from the

LNS test were not included in the model due to their high correlation with the raw total scores for working memory (Pearson correlation of .829, with $p > .01$). Nor were the average self-ratings of proficiency included due to their significant correlation with the scores from the independent and more objective measure of proficiency employed here (Pearson correlation of .433, with $p > .01$). The Flanker Effect and the Flanker Facilitation did not contribute significantly to the model, and were subsequently dropped from the analysis. In addition to the above fixed effects, I entered the random effects of participant (subject) and item (sentence) into the models.

Before running the models, I centered the proficiency and working memory raw scores. Since scores of zero in these two measures were not viable, this was the best way to prepare the data for this type of analysis. This was also done for the Flanker Inhibition measure because even though a score of zero is possible, it is highly unlikely that a participant with an average WM score and an average proficiency score would have a score of zero for Flanker Inhibition. With the centered Flanker Inhibition scores, the intercept estimates are clearer to interpret. In addition, for consistency, I set the reference levels for the variables of first language (as English) and condition (as SVO) so that the results would be clearer to interpret. Finally, I removed the incorrect responses in which participants selected pictures from the distractor conditions 3 and 4 (Semantically Incongruent/Grammatically Congruent and Semantically Incongruent/Grammatically Incongruent) from the main analysis in order to have a binary variable (correct and incorrect) that included only the critical pair of pictures, conditions 1 and 2. For the sake of completeness, I later analyzed the distractor choices in order to check the models. I present these results in the following sections after presenting the results from each of the models.

For purposes of clarity, for the learner group models, first I lay out a simple model with just the main effects. I then build in the interactions. The full output from each listening model, which also include the random slopes, are located in Appendices I-L.

3.2 Spanish Monolingual Listening Models

In order to provide a baseline for comparison for the learner participants, I first present the models for the Spanish monolinguals. Naturally, the predictors of proficiency in Spanish, immersion time in a Spanish-speaking country, and first language will not be part of these models. The predictors in the models for the Spanish monolinguals were the within-subject variable of condition and the between-subject variables of the centered working memory (cWM) scores and the Flanker Inhibition (cFlankInhib) scores. I also included the interaction between condition and working memory to test whether working memory scores affected accuracy on each sentence condition for this group. In this section, I report and discuss the results from the listening task model. Then, I show the effects of the distractor choices on the model.

3.2.1 Listening Task Model

The fixed effects estimate coefficients and statistical values for the data from the monolinguals' listening task with SVO as the reference level for Condition are shown in Table 16.

Table 16. *Fixed Effects for Listening Task by Monolinguals.*

Fixed Effects	**Estimate**	**Std. Error**	**Z value**	**P value**	**Sig.**
(Intercept)	5.32	0.93	5.72	1.09e-08	***
Condition	-0.33	1.11	-0.30	0.77	
cWM	0.10	0.26	0.39	0.70	
cFlankInhib	-0.01	0.01	-0.69	0.49	
Condition:cWM	0.27	0.32	0.84	0.40	

Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

The model shows that the intercept is highly significant and positive, which indicates that for a monolingual participant with average working memory capacity and average Flanker Inhibition scores accuracy on SVO sentences (reference level) was above chance. The model also shows that sentence condition (SVO vs. OVS) did not have a significant effect on the monolinguals' accuracy in this task. Although the accuracy means were slightly higher for the SVO than the OVS condition for the monolinguals as a group, as mentioned above, the monolinguals did not perform significantly better on the SVO condition than on the OVS condition. This is not surprising due to the high accuracy means for both conditions for this group. Arguably, these participants were at ceiling on this task, which is to be expected since they are adult native speakers of the language of the task.

In addition, the individuals' working memory scores and Flanker Inhibition scores did not play a significant role in their accuracy on any of the items in the listening task. This may be again due to their high accuracy overall on this task. Working memory and attentional control abilities do not modulate accuracy on this task when the participants are all native speakers and highly fluent in the language according to the model. The nature of the task may play a role in lessening the effect of working memory here as well. Although I had hypothesized that listening to the sentence would be cognitively taxing, and therefore accuracy on this task would be modulated by working memory ability, the participants saw the four pictures at the same time that they were listening to the sentences, and thus did not have to hold the relevant information in working memory long before making their picture selection.

Finally, the model shows that there was no interaction between condition and working memory. Thus, the working memory of the participants did not play a significant role in their accuracy on the items by condition. This, too, can be explained by their overall high accuracy on the task, allowing for little variation between participants, and possibly by the task not being taxing enough on working memory abilities. The full text of the output from the model including the data from the random effects can be seen in Appendix I.

3.2.2 Distractor Items Analysis

In order to check that the models above were accurate even though they did not take into consideration the distractor choices (conditions 3 and 4 of the pictures), I ran a model that included the data for these two wrong answer choices. For the listening data, the Spanish monolinguals, out of a total of 560 trials (data from one subject was missing due to technological error), chose distractor items only 32 times. The results from this model are in Table 17. The full text can be found in Appendix J.

Table 17. *Fixed Effects for Distractor Items in Listening Task by Monolinguals.*

Fixed Effects	Estimate	Std. Error	Z value	P value	Sig.
(Intercept)	5.42	0.89	6.10	1.06e-09	***
Condition	-0.43	0.43	-1.01	0.31	
cWM	0.10	0.09	1.07	0.29	
cFlankInhib	-0.01	0.01	-1.44	0.15	

Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

As shown above, the only significant effect was the intercept. Thus, none of the predictor variables determined the monolinguals' choice of a distractor item. This confirms that when the monolinguals chose the distractor items, it was not done in a systematic way, nor did the choice of these items have a significant impact on the model as a whole.

3.3 Learner Listening Models

Now that the baseline has been established by the Spanish monolingual listening data, I present the listening data for the L2 learners. The predictors in the models for the L2 learners were the within-subject variable of condition and the between-subject variables of first language (Romanian and English), the centered Spanish proficiency scores (cProf), the amount of immersion time spent in a Spanish-speaking country (ImmersionTime), the centered working memory (cWM) scores, and the centered Flanker Inhibition scores (cFlankInhib). First, I report and discuss the main effects results from the listening task model, and then build in the interactions. Then, as for the monolinguals, I show the fixed effects for the learners of the distractor choices on the models for the listening task. The data for the random effects as well as the full models for the learners can be found in Appendices K and L.

3.3.1 Listening Task Model

For the listening task for the L2 learners, I first built a model with just main effects. The fixed effects estimates with SVO (aSVO) as the reference level for Condition and English as the reference level for First Language are shown in Table 18.

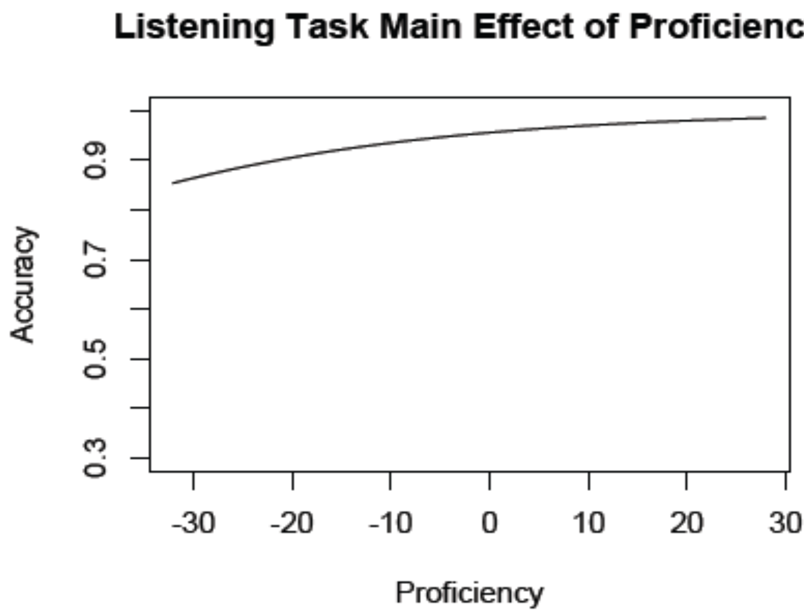
Table 18. *Fixed Effects for Listening Task by Learner Participants.*

Fixed Effects	**Estimate**	**Std. Error**	**Z value**	**P value**	**Sig.**
(Intercept)	3.64	0.25	14.66	<2e-16	***
Condition	-2.80	0.18	-15.47	<2e-16	***
ImmersionTime	-0.02	0.01	-1.46	0.14	
FirstLanguage	0.43	0.22	1.98	0.05	*
cWM	-0.07	0.04	-1.86	0.06	.
cFlankInhib	-0.01	0.01	-1.12	0.26	
cProf	0.08	0.01	10.29	<2e-16	***

Significance codes: ‘***’ 0.001, ‘**’ 0.01, ‘*’ 0.05, ‘.’ 0.1

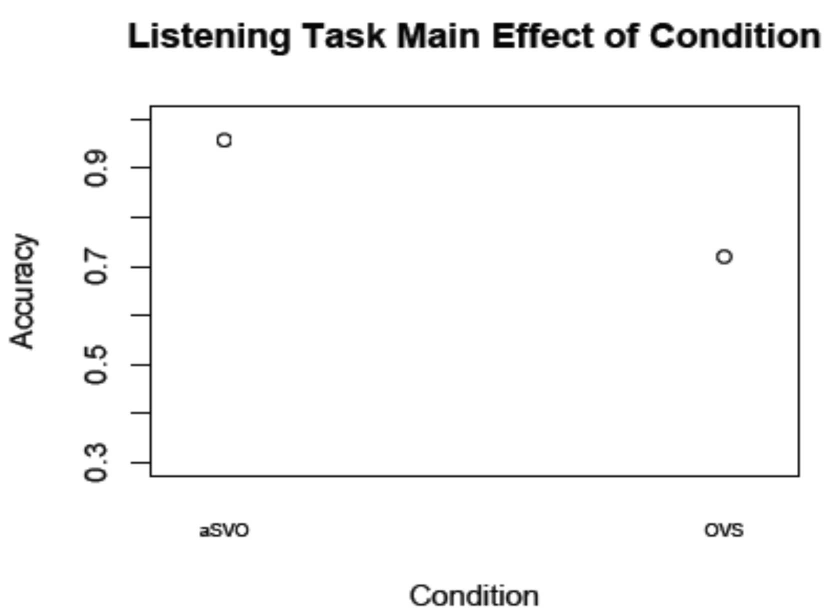
First, the intercept is highly significant. Second, there are highly significant main effects for both the predictors of condition and proficiency (cProf) in the model. This shows that the more proficient the learner, the more accurate he was on the listening task, as shown in Figure 21.

Figure 21. *Listening Task Main Effect of Proficiency*



Also, as predicted, there is an effect for condition such that the learner participants as a whole performed more accurately on the SVO condition than the OVS condition, as shown in Figure 22.

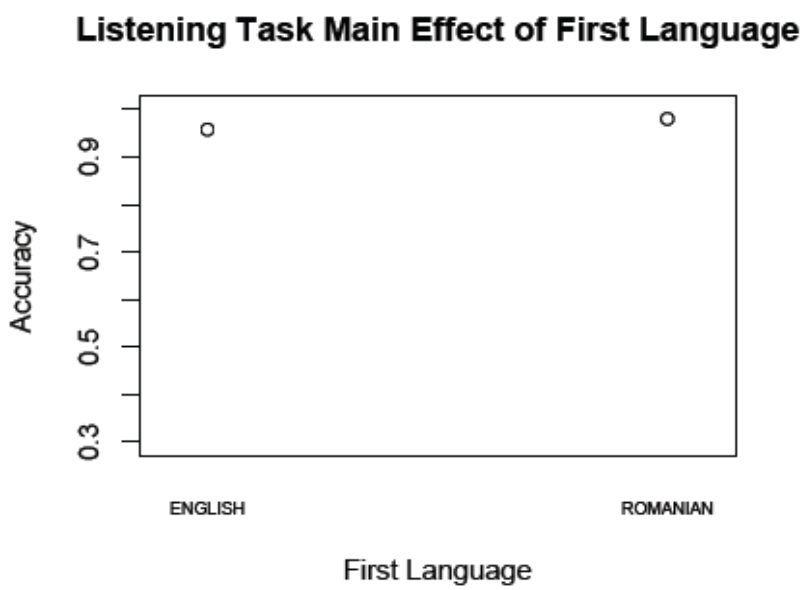
Figure 22. *Listening Task Main Effect of Condition*



This is not unexpected since the SVO word order is common in Spanish as well as in English (canonical, in fact) and in Romanian. Recall that although the Spanish monolingual participants did not perform significantly differently on the two conditions, descriptively they were more accurate on the SVO condition as well. In the next section, I break this main effect down into interactions to more clearly see how condition affects accuracy.

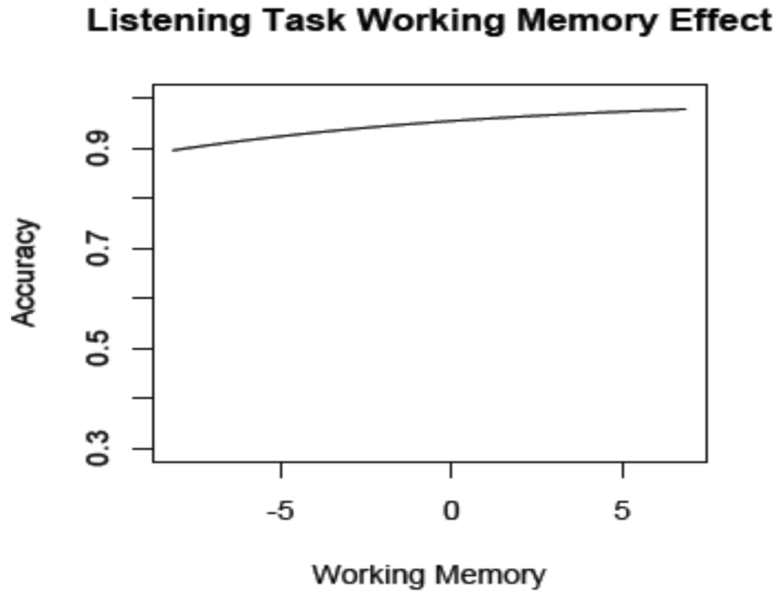
The effect of first language also reached significance in the model. This suggests that there were differences between how the Romanians and how the English-speakers performed on the accuracy of this task, with the Romanians scoring higher overall, as can be seen in Figure 23.

Figure 23. *Listening Task Main Effect of First Language*



There is also a main effect for working memory (cWM) that is marginally significant ($p = .06$), which indicates that working memory is affecting how accurate the learners are. This effect can be seen in Figure 24.

Figure 24. *Listening Task Working Memory Effect for Learners*



The variable of immersion time is not a significant predictor of accuracy on the listening task in the model, although based on a previous model I built without this predictor variable its presence appears to sharpen up the model by slightly lowering the log likelihood. Along the same lines, the Flanker Inhibition scores do not appear to predict accuracy on the listening task. This may be due to the fact that the task involved does not require the participants to switch task goals at any point, and thus may not tap into the executive attention component of working memory. Although the predictor of Flanker Inhibition does not reach significance in any of the models analyzed here nor does it affect the model, and will therefore not be mentioned in the subsequent analyses, it remains in the model based on the theoretical reason to believe that it would have an impact on differentiating L2 learners' ability to process argument structure that I outlined in chapter 1. The next step is to look at the data in a more fine-grained way by analyzing possible interactions among the predictors and how their inclusion affects the model.

3.3.2 Listening Task Model with Interactions

The fixed effects estimates for the data from the L2 learners' listening task with the interactions between condition and first language, working memory, immersion time, and proficiency with SVO (aSVO) as the reference level for condition and English as the reference level for first language are shown in Table 19.

Table 19. *Fixed Effects for Listening Task by Learner Participants with Interactions.*

Fixed Effects	**Estimate**	**Std. Error**	**Z value**	**P value**	**Sig.**
(Intercept)	3.00	0.31	9.76	< 2e-16	***
Condition	-2.02	0.30	-6.67	2.53e-11	***
ImmersionTime	0.00	0.03	0.10	0.92	
FirstLanguage	0.79	0.38	2.07	0.04	*
cWM	0.11	0.07	1.58	0.11	
cFlankInhib	-0.01	0.01	-1.15	0.25	
cProf	0.04	0.01	3.16	0.00	**
Condition:ImmersionTime	-0.03	0.03	-1.03	0.30	
Condition:FirstLanguage	-0.47	0.38	-1.23	0.22	
Condition:cWM	-0.23	0.07	-3.32	0.00	***
Condition:cProf	0.05	0.01	3.66	0.00	***

Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

Before discussing the interactions found, it is important to note that the main effects for proficiency and condition are still present in the model, but that the impact of these predictors is more precisely shown with the inclusion of the interactions. The interaction between proficiency and condition is highly significant. This interaction by language group is illustrated in Figures 25 and 26.

Figure 25. *Proficiency and Condition Interaction for L1 English Participants*

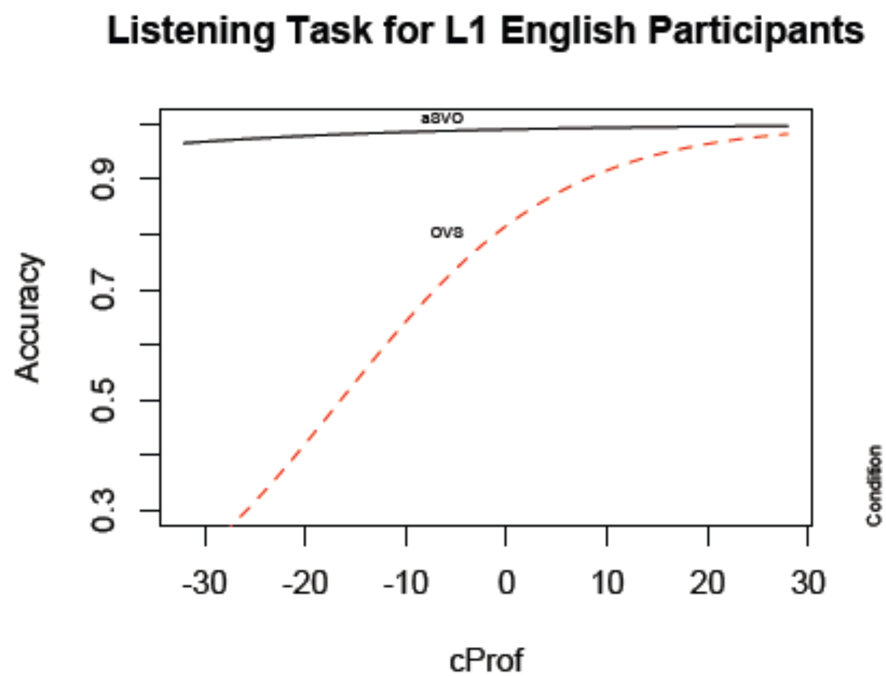
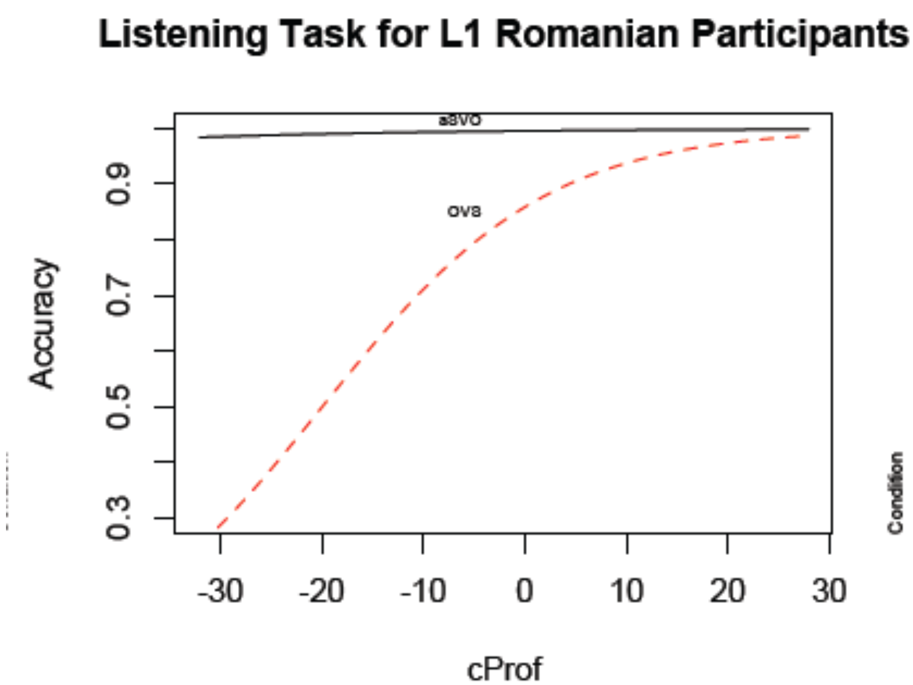


Figure 26. *Proficiency and Condition Interaction for L1 Romanian Participants*



Although the majority of the participants are highly accurate on the SVO condition items, the estimate coefficient for the interaction between proficiency and condition indicates that as proficiency increases, accuracy on the OVS items also increases. This is clear in the plots in Figures 25 and 26.

The interaction between working memory and condition is also highly significant, with the estimate coefficient indicating that participants with higher working memory performed slightly less accurately than participants with lower working memory. Although this is not predicted by the literature, the differences in accuracy are quite small, as illustrated in Figures 27 and 28, and could be due to the fact that the participants with lower proficiency levels happened to have higher working memory scores and vice versa.

Figure 27. *Working Memory and Condition Interaction for L1 English Participants*

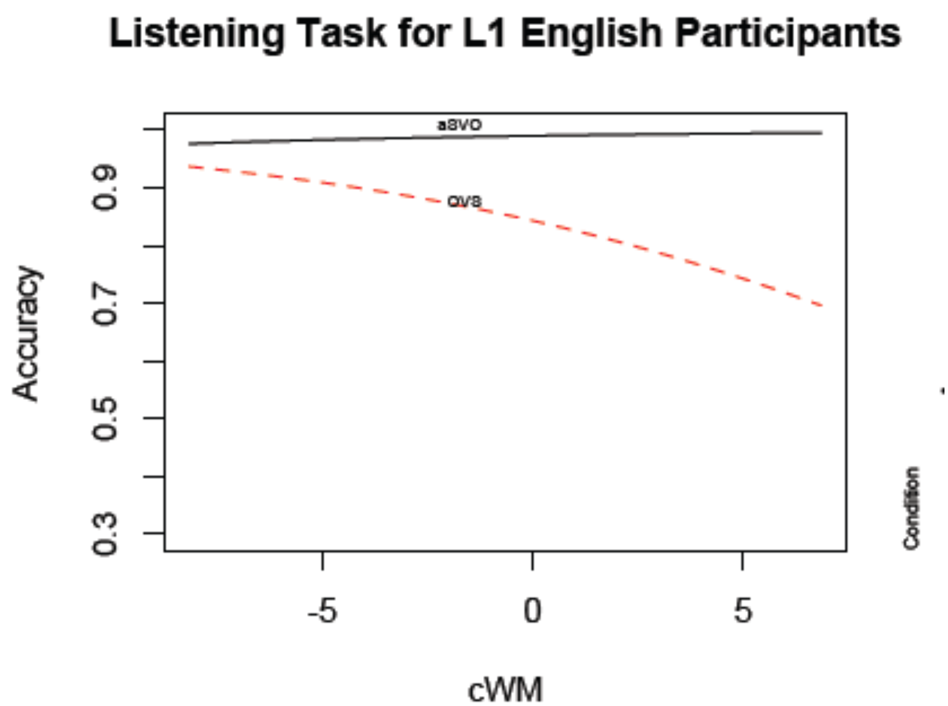
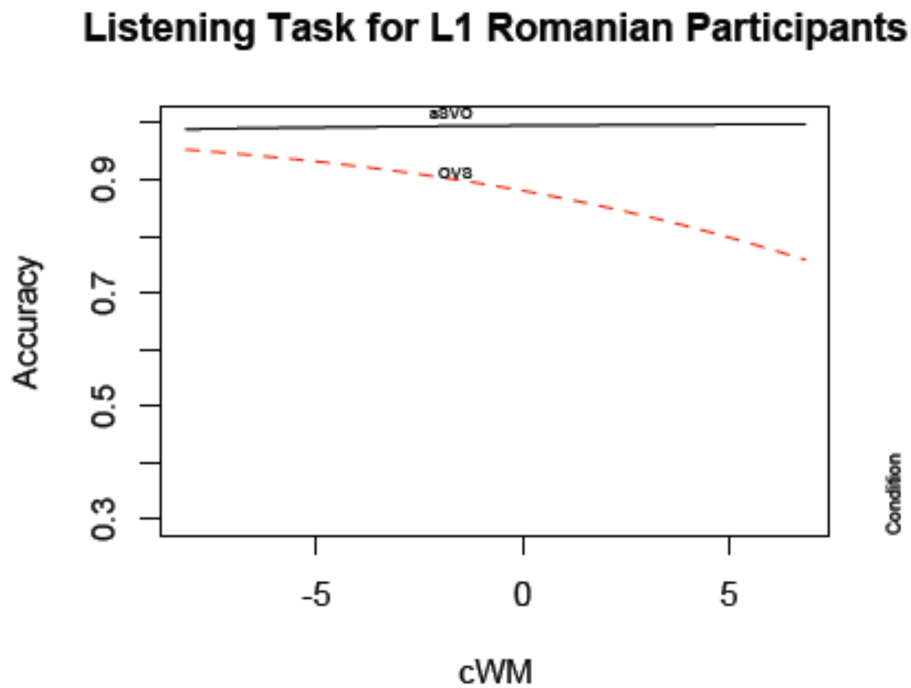


Figure 28. *Working Memory and Condition Interaction for L1 Romanian Participants*



Whether this was related to age or not is unclear. Adding age to the model did not return any significant results. This is an interesting venue to research further in future work.

It is interesting to note that there is no interaction between condition and first language. The Romanians appear to not be more accurate on the OVS items than the English-speakers even though these structures are present in their first language. Since first language does show a significant main effect, this finding indicates that the Romanians' success was not due to better accuracy on the OVS sentences per se, but to overall better performance on the listening sentence processing task. This suggests that their apparent advantage is not tied specifically to this structure, but may be due to more global similarities between Romanian and Spanish that do not exist between English and Spanish. The full text of the output from the model with the interactions can be seen in Appendix K. Now that I have presented the models built for the listening task for the learners, I present the model for the distractor items in the following section.

3.3.3 Distractor Items Analysis

Just as with the monolingual models, I removed the incorrect responses that included the two total distractor pictures and only analyzed the responses that included the pictures from the critical pair. Therefore, in order to check that the models above were accurate even though they did not take into consideration the distractor choices (conditions 3 and 4 of the pictures), I ran models that included the data for these two wrong answer choices. For the listening data, the learners chose the distractor choices a total of 124 times (80 for condition 3, 44 for condition 4), which is 5.7% of the total trials (2176) included in the analysis. The model for the distractor items can be found in Table 20.

Table 20. *Fixed Effects for Distractor Items in Learners' Listening Task Model.*

Fixed Effects	Estimate	Std. Error	Z value	P value	Sig.
(Intercept)	3.97	0.44	8.93	< 2e-16	***
Condition	0.02	0.21	0.09	0.93	
FirstLanguage	-0.25	0.23	-1.10	0.27	
cWM	-0.01	0.04	-0.35	0.73	
cProf	0.03	0.01	3.47	0.00	***
ImmersionTime	-0.01	0.01	-0.91	0.36	
cFlankInhib	-0.00	0.00	-0.97	0.33	

Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

As in the models for the Spanish monolinguals, condition, working memory, and Flanker Inhibition were not significant predictors for the choice of one of the distractor items. The only main effect was for proficiency, which is logical since I had included the distractor items as a secondary check for proficiency. Those participants with lower proficiency were more likely to choose the distractor items due to a more general proficiency breakdown and not because of the structure itself being challenging. The full text of the output from this model can be seen in Appendix L.

3.4 Summary of Results

For the monolingual participants, who served as a baseline for comparison, there were no main effects for condition, working memory, or Flanker Inhibition scores. Neither was there an interaction between condition and working memory nor a main effect for any predictor when the distractor items were tested. These results show that the monolinguals were at ceiling for accuracy on this task, and provide a point of departure for the L2 learner model results.

The L2 learner models for the listening task showed main effects for condition, first language, and proficiency, with an effect of working memory that was trending toward significance. In brief, the Romanians were more accurate overall, just as the more proficient learners were more accurate overall. For condition, participants were more accurate on the SVO than on the OVS structures. In addition to the main effects, there was a significant interaction between proficiency and condition such that as proficiency of the learners increased so did their accuracy on the OVS condition with the SVO condition accuracy relatively high across levels of proficiency. Somewhat surprisingly, there was a significant interaction between working memory and condition that appeared to be linked to worse performance on the OVS items as working memory abilities increased, an interesting result that warrants further study. No interactions or effects were found for immersion time or Flanker Inhibition scores in this task. Finally, there was only a significant main effect for proficiency when the distractor choices were examined, which showed that the lower a participant's proficiency, the more likely she would chose those items, and vice versa.

In sum, the important findings from the present analysis are twofold. Of the variables that I had predicted to have a significant impact on L2 learners' processing of preverbal clitics and postverbal subjects, proficiency appears to weigh the most. As proficiency improved, so did the accuracy results. The other important finding from the listening sentence processing task is that the first language of a participant also had an impact, but it had an impact that was not structure-specific. The L1 speakers of a language that is typologically similar to Spanish were more accurate overall. The other variables investigated here do not appear to be playing a strong role in determining processing accuracy on this task. In the next chapter, I reveal the results and analysis for the reading task, and discuss their implications.

CHAPTER 4: Reading Task Results

4.0 Introduction

In the present chapter, I provide the results of the analysis of the effects of first language and individual differences on how accurately L2 learners of Spanish visually process sentences with SVO word order and OVS word order (preverbal clitics and postverbal subjects). First, I present the descriptive statistics for the reading task (4.1) and the predictors that I use in the models (4.2). Then, I explain each of the mixed-effects models that were created for the reading task, presenting the models for the monolinguals in section 4.3 and the models for the learners in section 4.4. Finally, I provide a summary and a discussion of the reading task results (4.5).

4.1 Descriptive Statistics for the Reading Task

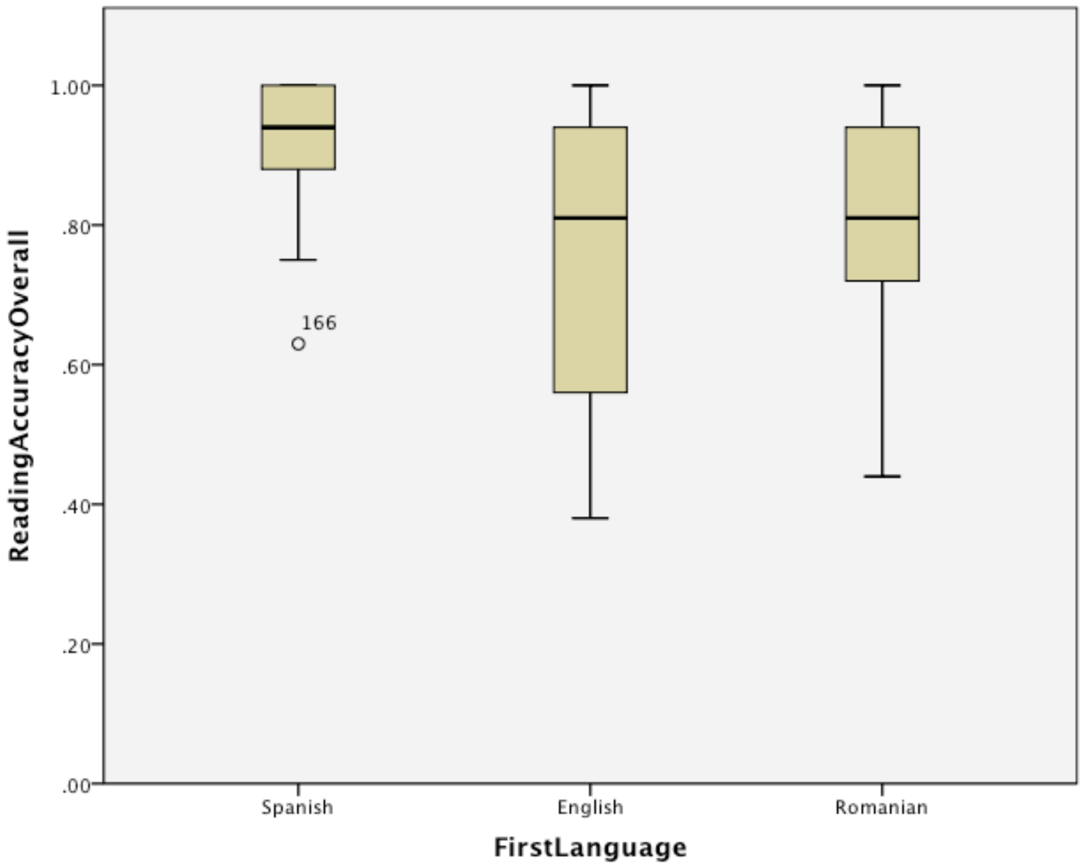
The accuracy means and SDs for each group and both conditions in the reading task are found in Table 21.

Table 21. *Reading Accuracy Overall and by Condition for All Language Groups*

Language Group	Condition	N	Mean	SD
Eng-Span	SVO	65	0.89	0.12
	OVS	65	0.63	0.33
	Overall	65	0.76	0.19
Rom-Span	SVO	71	0.89	0.11
	OVS	71	0.71	0.28
	Overall	71	0.80	0.15
Span Mono	SVO	36	0.92	0.11
	OVS	36	0.90	0.12
	Overall	36	0.90	0.09

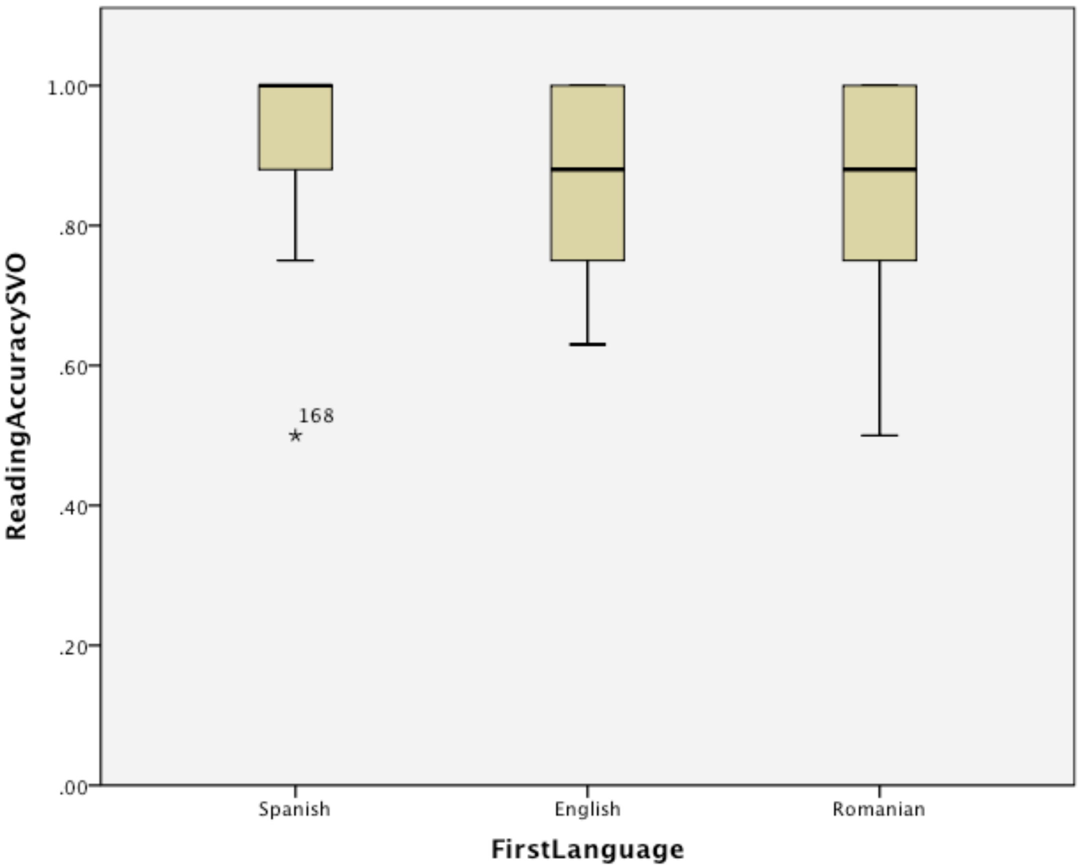
I have included the Spanish monolinguals here to provide a baseline for comparison for the accuracy means. The data from a different Romanian participant than for the listening task were missing due to technological error. As shown in Figure 29, the overall accuracy means for the reading task are similar to those for the listening task.

Figure 29. *Reading Accuracy Overall Means for All Language Groups*



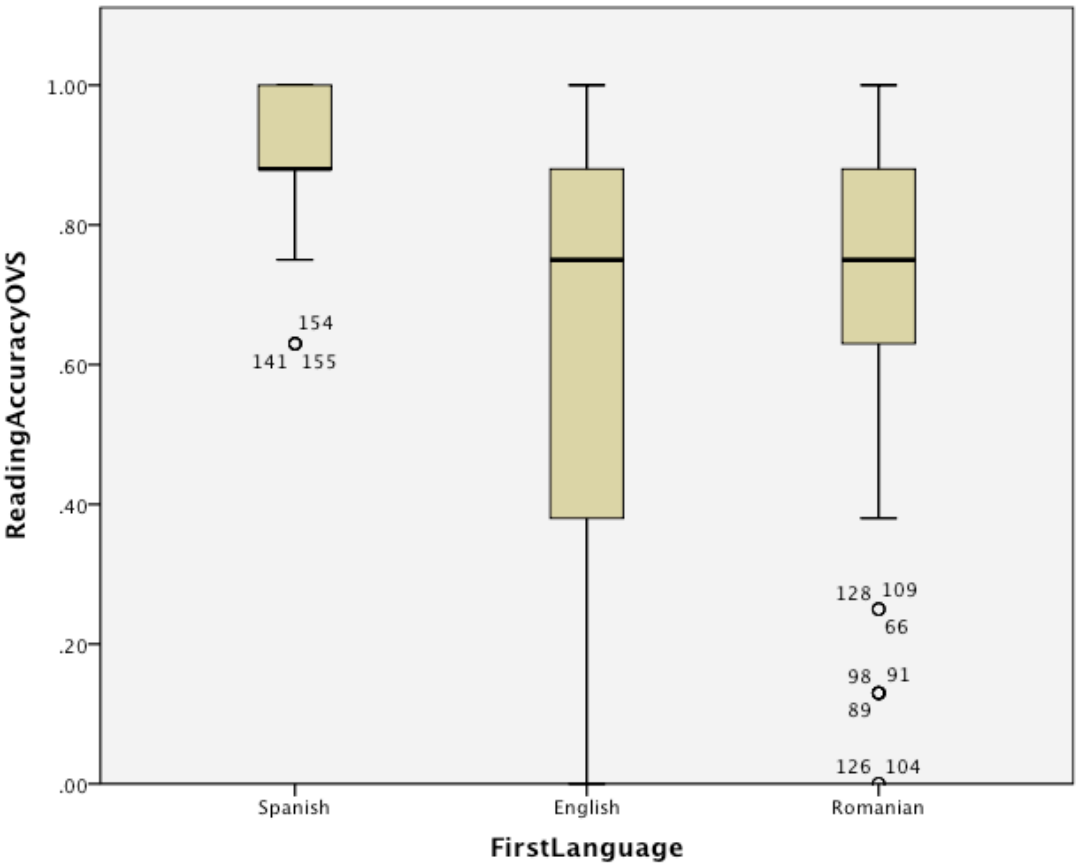
The Spanish monolinguals performed better than the two learner groups, whereas the two learner groups do not appear to be different from each other, even though the Romanian-Spanish participants have slightly higher scores for each condition than the English-Spanish participants. When the accuracy means for the two conditions are separated, however, we see that there are some differences in the results from the reading task as compared to the listening task in that for the SVO condition, unlike for the listening task, the learner groups in particular appear to be relatively similar in their means, as shown in Figure 30.

Figure 30. *Reading Accuracy SVO Means for All Language Groups*



For the OVS condition, however, the two learner groups performed much less accurately than the Spanish monolinguals, as shown in Figure 31, but not very differently from each other, which is also unlike the descriptive results from the listening task.

Figure 31. *Reading Accuracy OVS Means for All Language Groups*



It is important to note that these are merely the descriptive statistics, and that I provide a more detailed statistical analysis of the reading task data below in the mixed-effects models.

4.2 Predictors for Mixed-Effects Models

I examined the participants' accuracy on the reading sentence processing tasks (a continuous variable based on the experimental sentences from these tasks) using logistic regression with a mixed effects model, fit by the Laplace approximation, just as I did with the listening task data. The crossed random effects (variables specific to the present data) were for subjects (*SD* = 0.85 for the reading task for the learners) and for items (*SD* = 0.24 for the reading task for the learners). The fixed effects or predictor variables, which are the same as in the listening task, can be found in Table 22.

Table 22. *Predictor Variables for Reading Task*

Predictor Variables
Condition
Immersion Time
First Language
Working Memory
Flanker Inhibition
Proficiency

I included in the reading model the fixed effects (independent variables) of proficiency in Spanish of the learner participants (a continuous variable based on scores from adapted DELE test), immersion time in a Spanish-speaking country (a continuous variable based on self-reports in the language history questionnaire), working memory total raw scores (a continuous variable based on LNS test), Flanker Inhibition scores (a continuous variable based on Flanker test), first language (a binary variable for the learners' data with English and Romanian as the two levels), and sentence condition (a binary variable for all the data with SVO and OVS as the two levels). As noted above, memory length scores from the LNS test were not included in the model, nor were the average self-ratings of proficiency included. Again, the Flanker Effect and the Flanker Facilitation did not contribute significantly to the models, and were dropped from the analysis. In addition to the above fixed effects, I entered the random effects of participant (subject) and item (sentence) into the models.

As I did with the listening data, I used the centered proficiency, working memory, and Flanker Inhibition scores. In addition, for consistency, I set the reference levels for the variables of first language (as English) and condition (as SVO) so that the results would be clearer to interpret. Finally, I removed the incorrect responses in which participants selected pictures from the distractor conditions 3 and 4 (Semantically Incongruent/Grammatically

Congruent and Semantically Incongruent/Grammatically Incongruent) from the main analysis in order to have a binary variable (correct and incorrect) that included only the critical pair of pictures, conditions 1 and 2. For the sake of completeness, I later analyzed the distractor choices in order to check the models. I present these results in the following sections after presenting the results from each of the models.

For purposes of clarity of exposition, for the learner group models, first I lay out a simple model with just the main effects. I then build in the interactions. The full output from the reading models, which also include the random slopes, are located in Appendices M-P.

4.3 Spanish Monolingual Reading Models

In order to provide a baseline for comparison for the learner participants, I first present the models for the Spanish monolinguals. Again, the predictors of proficiency in Spanish, immersion time in a Spanish-speaking country, and first language will not be part of these models. The predictors in the models for the Spanish monolinguals were the within-subject variable of condition and the between-subject variables of the centered working memory (cWM) scores and the Flanker Inhibition (cFlankInhib) scores. I also included the interaction between condition and working memory to test whether working memory scores affected accuracy on each sentence condition for this group. In this section, I report and discuss the results from the reading task model. Then, I show the effects of the distractor choices on the model.

4.3.1 Reading Task Model for Monolingual Spanish Speakers

The fixed effects estimates for the data from the monolinguals' reading task with the reference level of SVO for Condition are shown in Table 23.

Table 23. *Fixed Effects for Reading Task by Monolinguals*

Fixed Effects	Estimate	Std. Error	Z value	P value	Sig.
(Intercept)	3.25	0.42	7.66	1.82e-14	***
Condition	0.33	0.56	0.59	0.56	
cWM	0.10	0.11	0.84	0.40	
cFlankInhib	0.00	0.01	0.25	0.80	
Condition:cWM	0.26	0.16	1.61	0.11	

Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

The intercept is highly significant and positive. The effects of sentence condition, working memory, and Flanker Inhibition scores on accuracy are not significant for the monolinguals in the model for this task. Just as in the listening task model, the reading task model shows no significant interaction between condition and working memory. This again could be due to the high overall accuracy and a lack of difference between the mean accuracy scores of these participants by condition. The full text of the output from the reading model can be seen in Appendix M.

4.3.2 Distractor Items Analysis

In order to check that the model above was accurate even though it did not take into consideration the distractor choices (conditions 3 and 4 of the pictures), I ran a model that included the data for these two wrong answer choices. The results from the reading task paint a similar picture as the results from the listening task distractor items analysis. The number of distractor choices is even lower for the reading task such that for a total of 576 trials, the monolinguals chose the distractor items only 16 times, which perhaps indicates that this task was less challenging for the monolinguals. The data from this model are included in Table 24. The full output for this model can be found in Appendix N.

Table 24. *Fixed Effects for Distractor Items in Reading Task by Monolinguals.*

Fixed Effects	Estimate	Std. Error	Z value	P value	Sig.
(Intercept)	5.63	0.81	6.98	3.05e-12	***
Condition	0.04	0.61	0.06	0.95	
cWM	0.16	0.20	0.80	0.42	
FlankerInhibition	-0.00	0.01	-0.21	0.83	

Significance codes: ‘***’ 0.001, ‘**’ 0.01, ‘*’ 0.05, ‘.’ 0.1

4.4 Learner Reading Task Models

4.4.1 Reading Task Model

The fixed effects for accuracy by the L2 learners of Spanish on the reading task with the reference levels of SVO (aSVO) for condition and English for first language are found in Table 25.

Table 25. *Fixed Effects for Reading Task by Learner Participants.*

Fixed Effects	Estimate	Std. Error	Z value	P value	Sig.
(Intercept)	3.11	0.22	14.01	< 2e-16	***
Condition	-2.09	0.16	-13.15	< 2e-16	***
FirstLanguage	0.25	0.22	1.15	0.25	
cWM	0.02	0.04	0.52	0.60	
cProf	0.06	0.01	8.62	< 2e-16	***
ImmersionTime	-0.00	0.01	-0.24	0.81	
cFlankInhib	0.00	0.00	0.23	0.82	

Significance codes: ‘***’ 0.001, ‘**’ 0.01, ‘*’ 0.05, ‘.’ 0.1

For the reading task, the intercept is highly significant and positive, which indicates that the learners are performing above chance. Like in the listening task, condition and proficiency are highly significant predictors of accuracy. The plots for these effects can be found in Figures 32 and 33.

Figure 32. *Reading Task Main Effect of Condition*

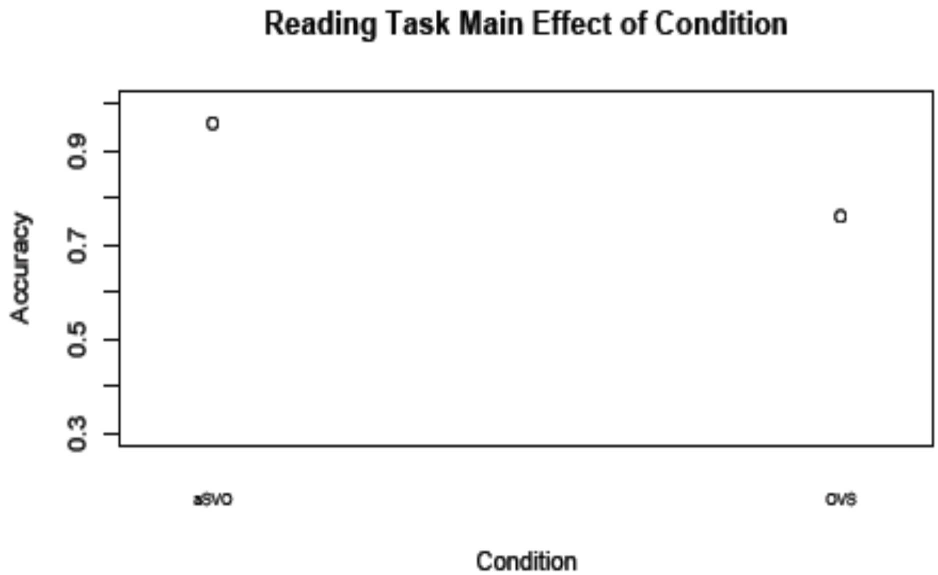
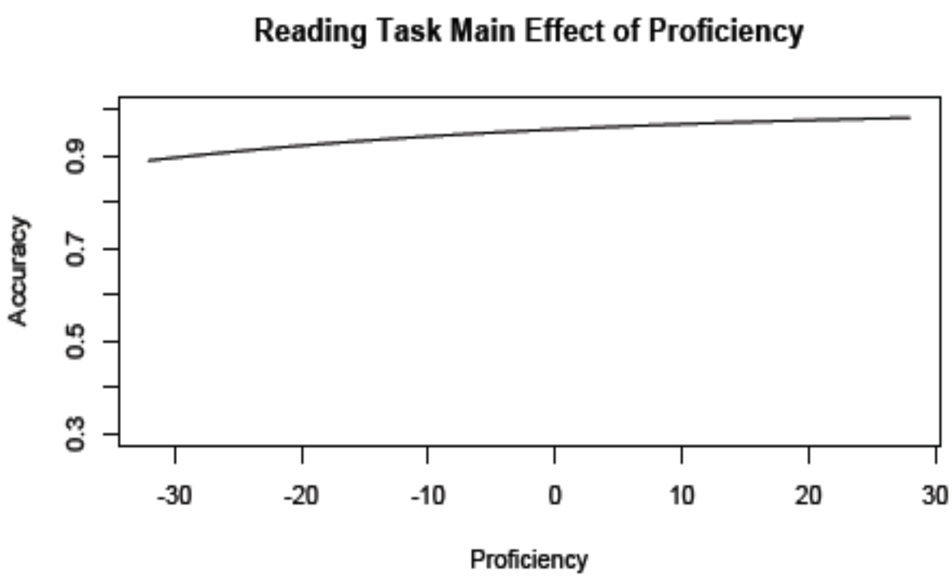


Figure 33. *Reading Task Main Effect of Proficiency*



First language and working memory, however, show no significant main effects in the reading task. Perhaps the main effects found in the listening task are mitigated here by the modality of the task. Just as the monolingual participants showed, the learner participants appeared to have performed on a more equal plane with respect to each other when faced with the task of reading the critical items instead of hearing them. Nevertheless, it is crucial to add the interactions to the model to see how to best interpret these results. The interactions for the reading task fixed effects can be found in the model in the following section.

4.4.2 Reading Task Model with Interactions

In Table 26, I present the reading task model for the L2 learners of Spanish that includes the interactions between the predictor of condition and the predictors of first language, working memory, proficiency, and immersion time with SVO (aSVO) as the reference level for condition and English as the reference level for first language.

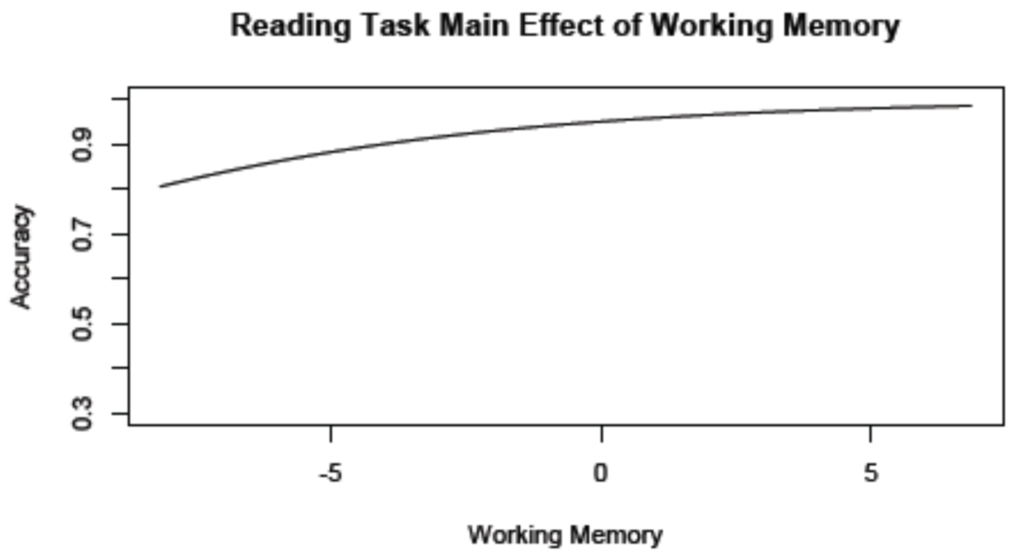
Table 26. *Fixed Effects for Reading Task by Learners with Interactions.*

Fixed Effects	**Estimate**	**Std. Error**	**Z value**	**P value**	**Sig.**
(Intercept)	2.92	0.29	10.18	< 2e-16	***
Condition	-1.80	0.28	-6.33	2.54e-10	***
FirstLanguage	0.02	0.34	0.04	0.97	
cWM	0.19	0.06	3.14	0.00	**
cProf	0.03	0.01	2.84	0.01	**
ImmersionTime	0.02	0.03	0.85	0.40	
cFlankInhib	0.00	0.00	0.19	0.85	
Condition:FirstLanguage	0.31	0.34	0.92	0.36	
Condition:cWM	-0.23	0.06	-3.75	0.00	***
Condition:cProf	0.04	0.01	0.83	0.00	***
Condition:ImmersionTime	-0.03	0.03	-1.27	0.21	

Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

The addition of the interactions to the model does not change the significance of either the predictor of condition or the predictor of proficiency, but it does draw out the effect of working memory and its significant interaction with condition. The main effect of working memory can be seen in Figure 34.

Figure 34. *Reading Task Main Effect of Working Memory*



This shows that overall as the participants' working memory capacity increases so does their overall accuracy. As for the interactions, the interaction between condition and proficiency is highly significant in this model, just as in the listening task model. In the present model, the slope of this interaction indicates that a learner of average proficiency performed better on the SVO condition than on the OVS condition. In other words, with increased proficiency there was increased accuracy on the OVS condition. These interactions are separated by first language and illustrated in Figures 35 and 36.

Figure 35. *Proficiency and Condition Interaction for L1 English Participants*

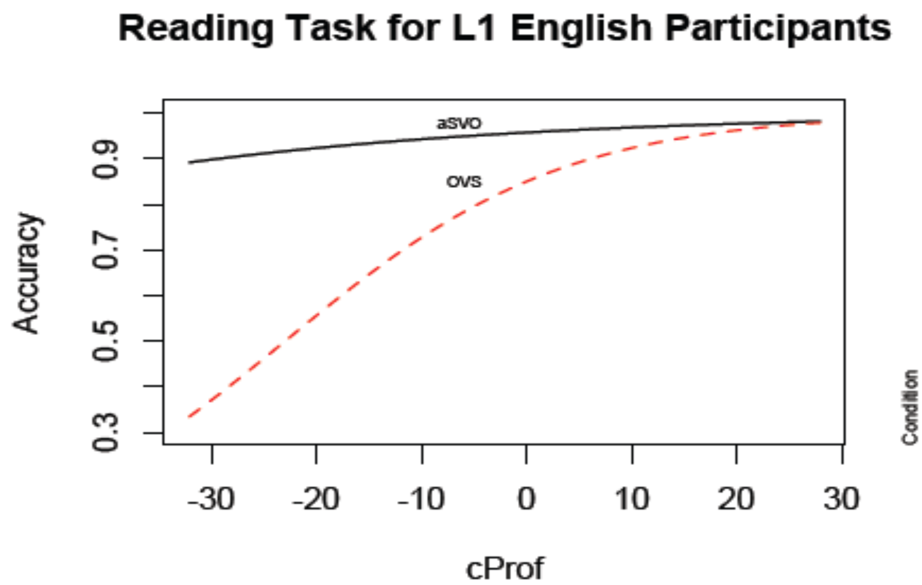
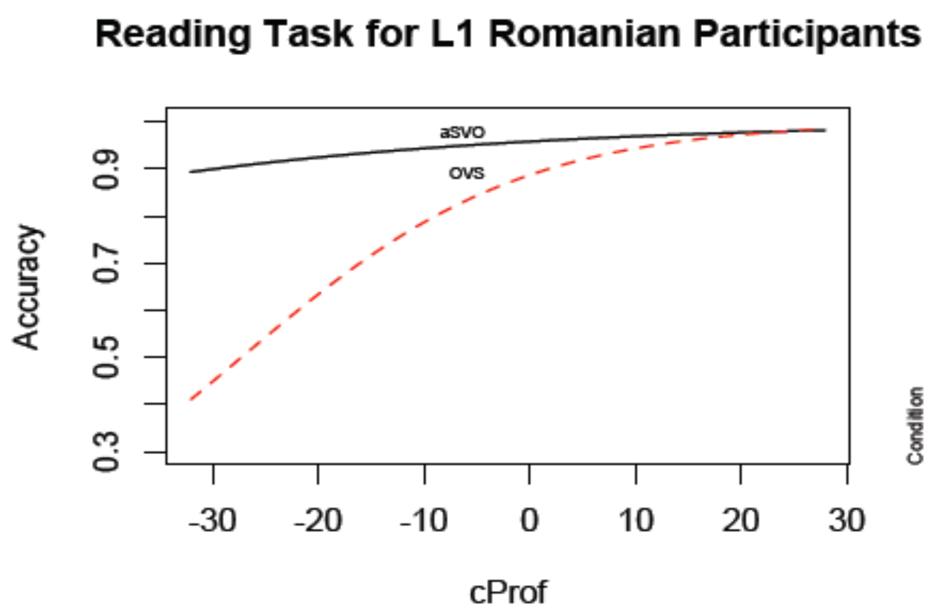


Figure 36. *Proficiency and Condition Interaction for L1 Romanian Participants*



As can be seen in Figures 35 and 36, there were no differences based on first language for the reading task. The interaction between working memory and condition was also significant in the reading model, just as in the listening model. Again, the slope shows that for increased working memory capacity, there is a decrease in accuracy on the OVS items. This result is again interesting in that there may be another factor involved that is not being considered in the present dissertation. This is an area that begs further research. These interactions are separated by first language and illustrated in Figures 37 and 38.

Figure 37. *Working Memory and Condition Interaction for L1 English Participants*

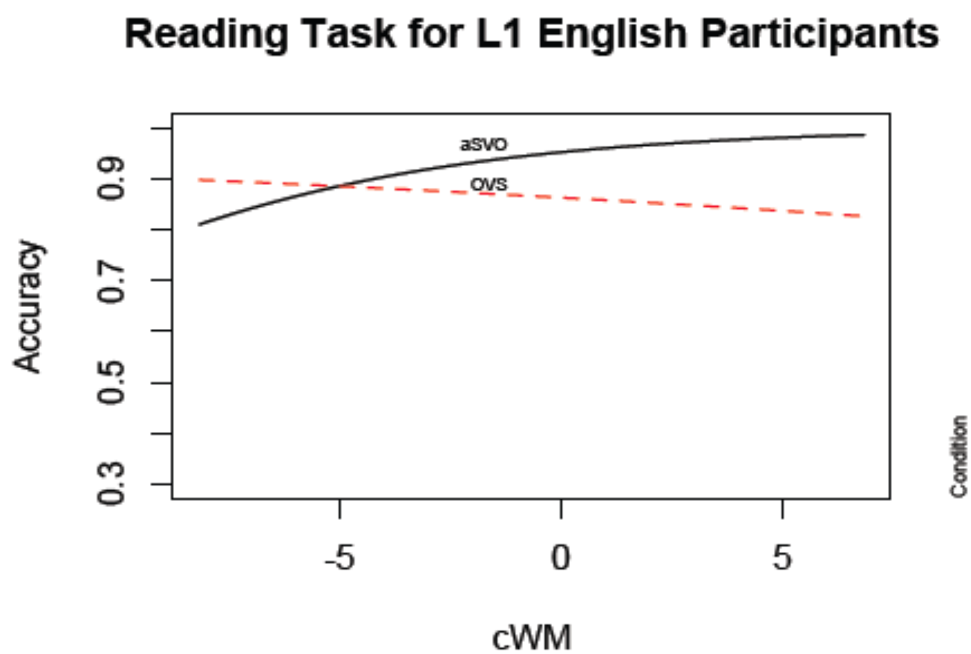
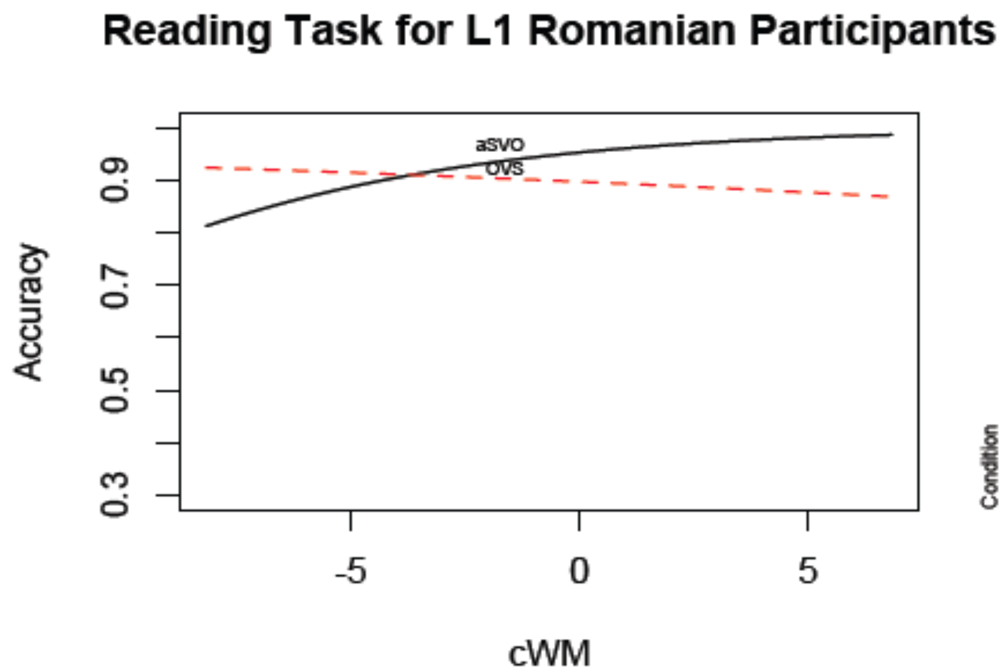


Figure 38. *Working Memory and Condition Interaction for L1 Romanian Participants*



The full text of the output from this reading model including the random effects can be seen in Appendix O.

4.4.3 Distractor Items Model

As with the monolingual models, for the models just presented, I removed the incorrect responses that included the two total distractor pictures and only analyzed the responses that included the pictures from the critical pair. In order to check that the models above were accurate even though they did not take into consideration the distractor choices (conditions 3 and 4 of the pictures), I ran a model that included the data for these two wrong answer choices. In the reading task, the learners chose the distractor choices a total of 106 times (57 for condition 3, 49 for condition 4), which is 4.9% of the total trials (2173) included in the analysis. The model for these data is shown in Table 27. The full text of the output from these models can be seen in Appendix P.

Table 27. *Fixed Effects for Distractor Items in Learners' Reading Task Model.*

Fixed Effects	Estimate	Std. Error	Z value	P value	Sig.
(Intercept)	4.15	0.49	8.43	< 2e-16	***
Condition	-0.31	0.22	-1.44	0.15	
FirstLanguage	0.15	0.24	0.60	0.55	
cWM	0.05	0.04	1.13	0.26	
cProf	0.02	0.01	2.62	0.01	**
ImmersionTime	-0.00	0.01	-0.13	0.90	
cFlankInhib	0.01	0.01	1.07	0.28	

Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

As in the models for the Spanish monolinguals and the listening task for the L2 learners, condition, working memory, and Flanker Inhibition were not significant predictors for the choice of one of the distractor items. The only main effect was for proficiency, which is not unexpected, given that I had included the distractor items as a secondary check for proficiency. Simply put, those participants with lower proficiency were more likely to choose the distractor items due to a more general proficiency breakdown and not because of the structure itself being challenging. Thus, these distractor items would not have added to the models and my decision to not include them in the main analysis is supported.

4.5 Summary of Results

Now that I have presented all of the models that fit the reading data, I provide a brief summary of the results. The Spanish monolinguals showed no main effects for condition, working memory, or Flanker Inhibition scores. There was no significant interaction between working memory and condition either. Their performance on the reading task was almost at ceiling, which is perhaps the reason for the lack of effects involving individual cognitive variables. This model served as a baseline for comparison with the L2 learner models.

The L2 learner reading task models showed main effects for condition and proficiency, and a main effect for working memory when the interactions were part of the model. Learners were more accurate on the SVO than on the OVS condition, more accurate if they were more proficient, and more accurate overall if they had a larger working memory. The interaction between condition and proficiency for the reading task indicated that proficiency modulated accuracy on the OVS items such that with increased proficiency there was increased accuracy. In addition, the interaction between condition and working memory for the reading task indicated that as working memory increased accuracy on the OVS items decreased, which may seem odd, but the decrease in accuracy is slight (<10%) and therefore requires more research to be completely understood. No interactions or effects were found for immersion time or Flanker Inhibition scores in this task, just as in the listening task. Finally, there was a significant main effect only for proficiency when the distractor choices were examined, which showed that only participants of lower proficiency chose those items, justifying their omission in the general models.

The main finding from the reading sentence processing task was that proficiency is the strongest factor in predicting the accuracy of learners' processing. The variable of first language was not found to be significant in the reading task. That is, there is no evidence in this task for either L1 transfer benefits in processing the OVS structure or for benefits of speaking a related language (Romanian, in this case) on accuracy. A comparison of the results from the listening task and the reading task as well as a general discussion and conclusions is provided in the following chapter.

CHAPTER 5: General Discussion and Conclusions

5.0 Introduction

In this chapter, I summarize and synthesize the results from the analysis of the data that was presented in chapters 3 and 4. First, in section 5.1, I discuss the results and relate them back to the literature. Then, I draw my conclusions based on the results and this comparison. In section 5.2, I describe some issues with the present research. Finally, I consider ways to further investigate these issues in future studies.

5.1 General Discussion and Conclusions

In general, when considering the original hypotheses advanced in Chapter 1, the results here are somewhat mixed. I hypothesized that first language and the individual differences of proficiency and cognitive abilities would have an effect on second language learners' processing of argument structure. The degree to which each of these components affected processing accuracy was in some cases unanticipated, and in others, just as predicted. Now, in more detail, I relate my findings to the literature reviewed. What follows is a discussion of the monolinguals' results, and then a discussion of the learners' results.

The monolingual models provided a baseline for comparison with the L2 learner models. There were no significant main effects for working memory or Flanker Inhibition on the accuracy of the sentence processing tasks, which was not surprising considering the high accuracy of the monolinguals. There was no significant main effect for condition either, which was of interest. Although prior literature had indicated that object-first structures are more challenging even for native speakers to process than subject-first structures (e.g. Havik et al., 2009) and the descriptive statistics showed that the monolinguals were slightly more accurate on the SVO condition overall, this was not borne out in the model. For these tasks, at least, the difference between the monolinguals' performance on the two conditions was not large enough to reach anything approximating statistical significance. This also could have been because the monolinguals were almost at ceiling in their accuracy on both tasks, and it may be the case that an earlier measure of the time course of processing, such as a measure found in an event related potentials design, might be able to tease out differences between these structures for native Spanish speakers. However, the accuracy measure here provides no evidence that the OVS structure is significantly harder for the native Spanish speakers to process.

The results from the L2 learner models, the focus of this dissertation, were more interesting. Both the models for the listening and the reading tasks for the L2 learners, unlike those for the monolinguals, resulted in highly significant main effects for condition. Specifically, the learners were significantly more accurate on the SVO condition than on the OVS condition. This finding was predicted by previous studies that have found that L2 learners have trouble understanding object clitic-first structures in Spanish (e.g. Isabelli, 2008; Liceras, 1985; VanPatten, 1984). Unlike the present dissertation, however, these previous studies did not incorporate proficiency as a continuous predictor variable. Here, the results showed a highly significant main effect for proficiency on both of the sentence processing tasks. This indicates that the more proficient a participant was, the more accurate she was on the task. Although this finding alone may not be surprising, it is nevertheless extremely important to assess the contribution of proficiency in order to rule out the potential of confounding proficiency and L1 (e.g. Isabelli, 2008).

Of primary interest here is the interaction between condition and proficiency. This interaction was also highly significant in both the reading and listening processing task models. As seen in the interaction plots in Chapters 3 and 4, it is the OVS structure in particular that is modulated by the proficiency levels of the learners. That is, while proficiency plays little if any role in predicting accuracy on SVO structures, its role is central to understanding the response accuracy rates for the OVS constructions.

In this context, it is striking to note that the L1 appears to play a very small role in the models. Based on prior literature that had primarily considered L1 English learners of Spanish and had assumed the negative influence of the L1, one hypothesis advanced at the beginning of this dissertation was that if learners of two different L1s were to be tested, then the two learner groups would perform differently on accuracy such that in the present dissertation the prediction was that the L1 Romanian participants would perform more accurately than the L1 English participants. This was borne out in the listening model with a significant main effect for first language, but not in the reading model. That is, the L1 Romanian learners of Spanish were more accurate overall on the listening task than the L1 English learners of Spanish. It is interesting, however, that there was no effect of L1 on the accuracy results for the reading task.

Of the most theoretical import here, despite the presence of a main effect for L1 in the listening task, there is no significant interaction between condition and L1 in both the reading and listening tasks. That is, whatever the benefits of L1 may be in the overall pattern of accuracy in the listening task, these benefits are not specific to a particular structure. I interpret this finding to be an indication that there is no direct transfer of the OVS structure. Simply put, there is no evidence here that the presence of preverbal clitics in Romanian facilitates the processing of analogous structures in Spanish. The results not only reinforce the general claim that the OVS condition is more challenging for learners than the SVO condition (e.g. LoCoco, 1982, VanPatten, 1984), but they further illuminate the nature of this difficulty, because the present study included learners from a typologically similar language to Spanish that licenses such structures as well as L1 English learners of Spanish. The difficulty with the preverbal clitic structures appears to apply to all second language learners. Following the phenomena outlined in Ellis' Associative-Cognitive framework (2006b), it appears logical that learners with various L1 backgrounds would have difficulty at low proficiency levels with such low salience structures as preverbal clitics.

One possible alternative interpretation of the results might reside in the possibility that the Romanian participants were processing the preverbal clitic and postverbal subject structures, such as *Lo besa el abuelo*, as clitic doubled structures due to the presence of these structures in their L1 (see Chapter 2 for a discussion on clitic doubling in Romanian). To rule out the possibility that the Romanian participants interpreted the sentences that included subjects and objects of the same gender as instances of clitic doubling, I later categorized each item from both experiments as either "same gender" (both constituents had the same gender) or "different gender" (both constituents had different genders), and ran a one-way ANOVA that included gender category as the independent variable and the accuracy on each item by each language group as the dependent variables. The results from this analysis are found in Table 28.

Table 28. *ANOVA Results for Gender Patterns of Constituents*

Group	**df**	**F**	**Sig.**
Accuracy L1 English	1	0.85	0.36
Accuracy L1 Romanian	1	0.32	0.57

As shown in Table 28, there were no significant differences between same gender and different gender items for the L2 learners. There appears to have been no disadvantage for the L1 Romanian learners based on a possible clitic doubling interpretation on the same gender items.

Turning to other individual differences across learners, there was a significant effect for working memory in the reading model, which did not reach significance in the listening model. Participants with higher working memory performed more accurately overall. In both models, however, working memory significantly interacted with condition. As noted in Chapters 3 and 4, both interactions revealed that the participants with higher working memory were actually less accurate on the OVS condition than those with lower working memory, a result that is not altogether congenial with previous work that has found results for WMC improving processing (e.g. Williams, 2006). This result may be somewhat compatible, however, with the results of Juffs (2004), who reported that based on his data, the central executive component of WM is not a source of individual variation in L2 online performance. Perhaps being able to hold more information in working memory did not aid in the processing of sentences that had less information (i.e. only direct object clitics, not full noun phrases). This decrease in accuracy is less than 10% for each task, but it remains provocative, requiring further study. Finally, I speculate that the presence of a main effect for working memory in the reading (but not in the listening) task may be an artifact of the experimental design. As I mention above in Chapter 4, in the reading task participants first read a sentence, which then disappears from the screen. Subsequently, the four-picture display appears, and participants must select the correct scene in the display. This method places a direct tax on working memory. By contrast, in the listening task, participants are able to continuously view the four-picture display while listening to the target sentences. Arguably, we might expect a working memory effect if participants were asked to listen to sentences without viewing the four-picture scene and were only able to view the pictures after having heard each target sentence. I leave this manipulation for future study, but it is worth noting that under a more taxing task, it is an empirical question whether we would find a WM and condition interaction.

It is important to note that the Flanker Inhibition scores did not show significance in any of the above models. Although the literature reviewed seemed to indicate that L2 learners with better executive attention abilities would be better able to attend to specific items and block out others (see Bialystok et al., 2009), an attentional benefit which might aid in the processing of clitics, no effect was found in the results of the present dissertation. The reason for this could be twofold: perhaps the sentence processing tasks failed to tap into the participants' executive attention abilities sufficiently enough to show variance among the individuals, or perhaps the Flanker task employed here was not a good measure of participants' executive attention abilities since it was a shortened version of the original from Luk (2008) (i.e. the present version did not include no-go trials, which are a challenging trial that may bring about more variance in the results). Also, the predictor of immersion time did not show significance in any of the learner models. This could be due to the uneven distribution of this variable across participants as well as between the two language groups on the whole. However, the predictor variable of immersion time was maintained in the models because it improved the fit of the models to the data for both sentence processing tasks by accounting for some of the extra variance among the participants.

Finally, the models for the distractor items in both tasks only showed proficiency as a significant predictor of accuracy. The participants with higher proficiency levels rarely chose the distractor items during the task. The participants with lower proficiency were more likely to do so than those with higher proficiency, but this was apparently not due to the condition of the sentence nor to a participant's individual differences since these predictors did not reach significance in the models presented here.

Taking this all together, the results from the sentence processing tasks show that proficiency and condition were by far the most significant predictors of accuracy. For listening, the model was different than for reading in that the predictor of first language was also significant as a main effect, and working memory was only trending toward significance instead of reaching significance, as it did in the reading model. Both the aural and the visual sentence processing tasks showed significant interactions between condition and proficiency and between condition and working memory, which allow me to conclude that at least some individual differences play a demonstrable role in L2 learners' processing accuracy of preverbal clitics and postverbal subjects.

In particular, the learner participants who scored the most accurately on the sentence processing tasks were (1) the most proficient, and (2) the L1 Romanian-speakers. Both of these characteristics point towards an experience-based theory of language acquisition (e.g. Ellis, 2006a, 2006b). With more L2 proficiency (and thus more experience with preverbal clitic structures), the participants were better able to navigate the two word order structures, regardless of whether the structures were present as part of the grammar of their L1. In short, there is little if any support here for a strong claim about structure-specific transfer playing a role in modulating learner performance on either the reading or listening tasks.

5.2 Current Issues and Future Research

The first step to take for future research will be to look closely at the eyetracking data that was also gathered in the reading sentence processing task. Fixation times and regressions to the target regions may reveal more granular information regarding how L2 learners are processing online the preverbal clitic and postverbal subject structures from the present dissertation. Differences between language groups or proficiency levels may become more obvious with these more sensitive measures.

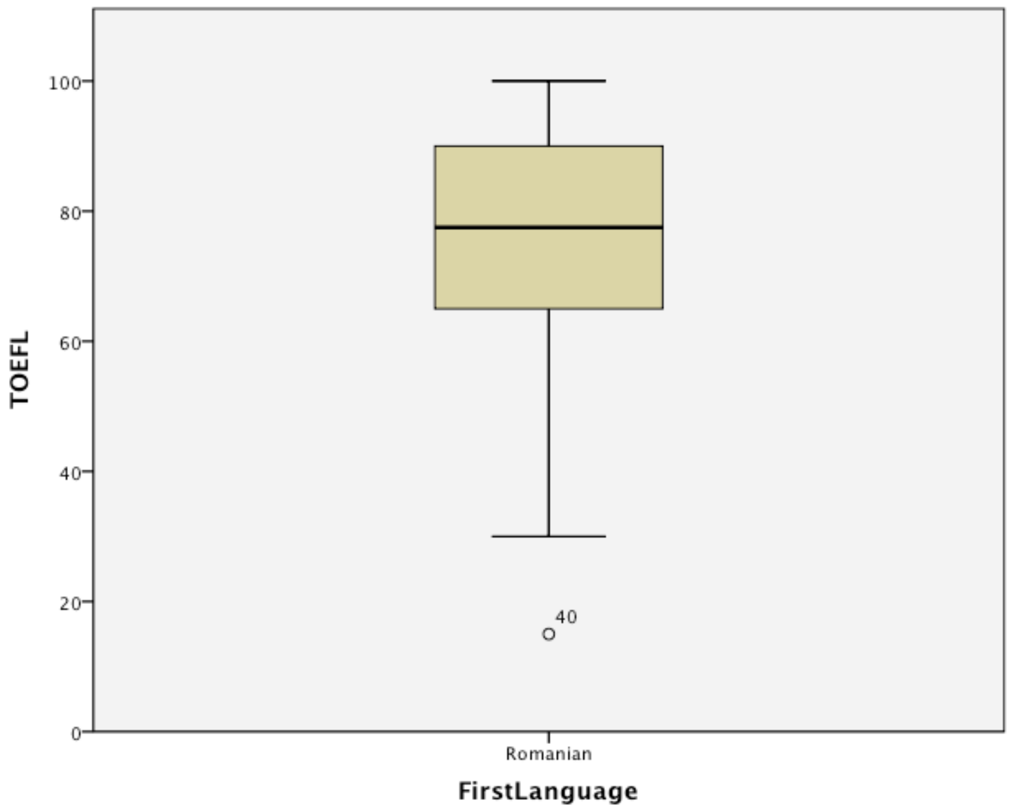
Now that this dissertation has contributed to the research by including the factor of proficiency, it will also be important in future research to assess how the participants arrive at their respective levels of proficiency. This is an open question here, but it is important to note that while the two groups in general were of equal proficiency, the Romanian group was characterized by less immersion and fewer overall years of study in Spanish. Many factors may come into play, such as self-selection, potential benefits of bilingualism (since many of the Romanians are also competent in English), or the relatedness of Romanian to Spanish as a Romance language with a fairly large shared vocabulary.

It could be proffered that the Romanians were more accurate overall on the sentence processing tasks not because of their experience with clitics in Romanian, but because they were already bilingual when they started learning Spanish, and tri- or multi-lingual when performing the experiments in the present dissertation. The majority of the Romanian participants spoke English well. The L1 Romanians also reported having significant knowledge of French, Russian, and other languages. I tested the participants who were not native speakers of English on an English proficiency test, which was adapted from the TOEFL exam. The means and SDs by language group are found in Table 29.

Table 29. *TOEFL Accuracy Means by Language Group*

Language Group	**Mean**	**SD**
Rom-Span	74.03	18.22
Span Mono	25.00	21.81

Figure 39. *L1 Romanians' Accuracy on TOEFL*



As seen in Table 28, the Spanish monolingual participants all scored below chance on the TOEFL exam, which was to be expected. The Romanians, however, as illustrated in Figure 39, were relatively proficient in English according to this measure. Since English does not have preverbal clitics or postverbal subjects, then, if anything, the Romanians' knowledge of English would bias them to being less familiar with such structures in Spanish. The time that the Romanians had spent using English instead of Romanian, would draw away from their experience with preverbal clitics, according to the Associate-Cognitive account (Ellis, 2006a). However, it could be argued that even though knowledge of English would not aid in their processing of preverbal clitics per se, being bilingual (or managing multiple languages) may have a more global cognitive benefit, according to the work done by Bialystok and colleagues. If this were the case, however, then the Romanians would have performed better on the Flanker Task by than their English-speaking cohort, which they did not.

Another line of future study would be to investigate the differences between the participants' performance on the two task modalities. The effect of first language was only significant in the listening task, not in the reading task. It had been hypothesized previously that the listening modality is more challenging for learners (and language processors in general) than the visual modality. Whether this discrepancy is due to particular characteristics of the L1 English group or whether it is due to the demands of the modality of the task is left to be determined in future research.

In respect to the lack of effects from the Flanker task, it will be important in future studies to consider using the full version of the Flanker task. Although the two language groups were not significantly different in their Flanker Inhibition scores, the version employed in the present dissertation, which lacked no-go trials, may not have been long enough nor have had distinctive enough trials to be able to thoroughly differentiate among the participants. Also, an additional measure could be used in future research (e.g. the Simon Task) for the sake of comparison. In addition, to better tap into executive attention abilities future research should ensure that the main experimental task is challenging enough along these lines. It is possible that the current sentence processing tasks were not a difficult enough of test of participants' executive attention abilities. It may also simply be the case that executive attention is not a crucial factor for this kind of processing task.

One final result that requires mention was that the interactions between condition and working memory in both the listening and reading tasks appeared odd in that they indicated that with increased working memory capacity there was decreased accuracy on the OVS items. This vexing result necessitates further research.

References

Ando, J., Fukunaga, N., Kurachachi, J., Suto, T., Nakano, T., & Kage, M. (1992). A comparative study on the two EFL teaching methods: The communicative and the grammatical approach. *Japanese Journal of Educational Psychology,* 40, 247-256.

Atkinson, R. C. and Shiffrin, R. M. (1968). Human Memory: A Proposed System and Its Control Processes. In K. W. Spence and J. T. Spence (Eds.), *The Psychology of learning and motivation: Advances in research and theory*. Vol. 2. (pp. 89-195). New York: Academic Press.

Avram, M. (1986). *Gramatica pentru toţi*. Editura Academiei Republicii Socialiste Romania. Bucureşti, Romania.

Babyonyshev, M. and Marin, S. (2006). Acquisition of Pronominal Clitics in Romanian. *Catalán Journal of Linguistics*, 5, 17-44.

Baddeley, A. D. and Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *The Psychology of learning and motivation: Advances in research and theory*. Vol. 8. (pp. 47-89). New York: Academic Press.

Baddeley, A. D. (1986). *Working Memory.* Oxford Psychology Series No. 11. Oxford: Clarendon Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, 4(11), 417-423.

Bader, M. and Meng, M. (1999). Subject-object ambiguities in German embedded clauses: an across-the-board comparison. *Journal of Psycholinguistic Research*, 28, 121-143.

Bates, D. M. (2007). *Linear mixed model implementation in lme4.* Manuscript, University of Wisconsin- Madison, January 2007.

Bates, E. and MacWhinney, B. (1979). A functionalist approach to the acquisition of grammar. In E. Ochs and B. Schieffelin (Eds.), *Developmental pragmatics*. New York: Academic.

Bates, E. and MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner and L. Gleitman (Eds.), *Language acquisition: the state of the art* (pp.173-218). New York: Cambridge University Press.

Bates, E. and MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney and E. Bates (Eds.), *The Crosslinguistic Study of Sentence Processing* (3-73). Cambridge: Cambridge University Press.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., and Venza, V. (1984).

A cross-linguistic study of the development of sentence interpretation strategies. *Child Development*, 55 (2), 341-354.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the Development of Language* (pp.279-362). New York: John Wiley & Sons, Inc.

Bialystok, E., Craik, F. I. M., Green, D. W., and Gollan, T. H. (2009). Bilingual Minds. *Psychological Science in the Public Interest*, 10(3), 89-129.

Bialystok, E., Craik, F. I. M., and Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, 45, 459–464.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, 19, 290–303.

Bialystok, E., Martin, M. M., and Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, 9, 103–119.

Blanche, P. and Merino, B. J. (1989). Self-assessment of foreign-language skills: implications for teacher and researchers. *Language Learning,* 39(3), 313-340.

Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, CBT, and subsequent performance. *System*, 34, 15-35.

Brown, R. (1973). *A First Language: The early stages.* Cambridge, MA: Harvard University Press.

Bruhn de Garavito, J. (2006). Knowledge of Clitic Doubling in Spanish: A Comparison of Early and Late Bilinguals. In S. Montrul, P. Prévost and R. Slabakova (Eds.) *Inquiries in Linguistic Development* (pp. 305-333).

Bruhn de Garavito, J. and Montrul, S. (1996). Verb movement and clitic placement in French and Spanish as a second language. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, and A. Zukowski, (Eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development* (pp. 123–134). Somerville, MA: Cascadilla Press.

Calude, A. (2001). Romanian Clitics: Siding with the French or the Serbo-Croatian? *Revue Roumaine de Linguistique,* 1-4, 91-105.

Caplan, D. and Waters, G. (1995). Aphasic disturbances of syntactic comprehension and WM capacity. *Cognitive Neuropsychology*, 12, 637–649.

Caplan, D. and Waters, G. (1996). Syntactic processing in sentence comprehension under dual-task conditions in aphasic patients. *Language and Cognitive Processes*, 11(5), 525–551.

Caplan, D. and Waters, G. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22, 77-126.

Carlson, M. and Gerfen, C. (2011). Spanish diphthongizing stems: Productivity, processing, and the shaping of the lexicon. *The Mental Lexicon*, 6:3, 351-373.

Carroll, J. B. and Sapon, S. (1959). *The modern language aptitude test.* New York: The Psychological Corporation.

Carroll, J. B. and Sapon, S. (2002). *Modern Language Aptitude Test: Manual 2002 Edition*. Bethesda, MD: Second Language Testing, Inc.

Chapman, G. B. and Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory and Cognition,* 18, 537-45.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786.

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, 106, 59–86.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, 113, 135-149.

Coughlan, A. and Hollows, S. (1985). *The Adult Memory and Information Processing Battery (AMIPB) Test Manual*. Leeds: A. K. Coughlan.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin,* 104(2), 163-191.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake and P. Shah (Eds.), *Models of WM: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge: Cambridge University Press.

Craik, F. I. M. and Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Behavior*, 12, 599-607.

Craft, J. L., & Simon, J. R. (1970). Processing symbolic information from a visual display: Interference from an irrelevant directional cue. *Journal of Experimental*

Psychology, 83, 415-420.

Crowe, S. F. (2000). Does the letter number sequencing task measure anything more than digit span? *Assessment*, 7(2), 113-117.

Cuervo, M. C. (2007). Double objects in Spanish as a second language: Acquisition of morphosyntax and semantics. *Studies in Second Language Acquisition*, 29, 583-615.

Daneman, M. and Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.

Daneman, M. and Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 25, 1-18.

Davies, M. (1995). Analyzing Syntactic Variation with Computer-Based Corpora: The Case of Modern Spanish Clitic Climbing. *Hispania*, 78(2), 370-380.

DeBot, K. and Van Montfort, R. (1988). "Cue-Validity" in Dutch as a First and Second Language. *Interdisciplinair Tijdschrift voor Taal- & Tekstwetenschap*, 8.2, 111-120.

De Vincenzi, M. (1991). *Syntactic Parsing Strategies in Italian*. Dordecht: Kluwer.

Dobrovie-Sorin, C. (1999). Clitics across categories: The case of Romanian. In H. von Riemsdijk (Ed.) *Clitics in the Languages of Europe* (pp. 515-540). Berlin: Mouton de Gruyter.

Duffield, N. and White, L. (1999). Assessing L2 knowledge of Spanish clitic placement: converging methodologies. *Second Language Research*,15(2), 133–160.

Dulay, H. C. and Burt, M. K. (1973). Should we teach children syntax? *Language Learning,* 23, 245-58.

Dussias, P.E. and Piñar, P. (2010). Effects of reading span plausibility in the reanalysis of *wh*-gaps by Chinese-English second language speakers. *Second Language Research*, 26(4), 443-472.

Dussias, P. E. and Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, 10 (1), 101-116.

Ellis, N. C. (2006a). Cognitive perspectives on SLA: The Associative-Cognitive CREED. *AILA Review*, 19, 100-121.

Ellis, N. C. (2006b). Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Salience, Interference, Overshadowing, Blocking, and Perceptual Learning. *Applied Linguistics*, 27(2), 164-194.

Ellis, N. C. and Sagarra, N. (2010). The Bounds of Adult Language Acquisition: Blocking and Learned Attention. *Studies in Second Language Acquisition*, 32, 553-580.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, 11(1), 19-23.

Engle, R. W., Cantor, J., and Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 18*,* 972–992.

Engle, R. W. and Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.) *The psychology of learning and motivation* (pp. 145-199). New York: Academic Press.

Engle, R. W., Kane, M. J. and Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake and P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge, England: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143–149.

Ervin-Tripp, S. M. (1974). Is second language learning like the first. *TESOL Quarterly*, 8 (2), 111-127.

Felser, C. and Roberts, L. (2007). Processing *wh*-dependencies in a second language: A cross-modal priming study. *Second Language Research*, 23, 9-36.

Fernández, E. M. (1995). Processing strategies in second language acquisition: Some preliminary results. In E. C. Klein and G. Martohardjono (Eds.), *The development of second language grammars* (pp. 217-240). Amsterdam: John Benjamins Publishers.

Fernández, E. (1998). Language dependency in parsing: Evidence from monolingual and bilingual processing. *Psychologica Belgica*, 38(3/4), 197-230.

Fernández, E. and Hirose, Y. (1997). *Bilingual sentence processing: Off-line results.* Paper presented at the Advanced Second Language Acquisition Seminar, May 15, 1997, CUNY Graduate School and University Center, New York, NY.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5, 519-559.

Frenck-Mestre, C. (1997). Syntactic ambiguity resolution while reading in second and native languages. *The Quarterly Journal of Experimental Psychology*, 50A, 119-148.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. R. Herredia and J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217-236). Amsterdam: Elsevier.

Gass, S. (1989). How Do Learners Resolve Linguistic Conflicts? In S. Gass and J. Schacter, (Eds.), *Linguistic Perspectives on Second-Language Acquisition* (pp. 183-199). Cambridge, CT: Cambridge University Press.

Gerlach, B. (1998). Optimale Klitiksequenzen. *Zeitschrift für Sprachwissenschaft*, 17(1), 36 -91.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, 1-76.

Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., and Weinberger, D. R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry*, 54, 159-165.

Goldschneider, J. M. and DeKeyser, R. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, 51, 1-50.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67-81.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131-149.

Grüter, T. and Crago, M. (2010). The roles of L1 transfer and processing limitations in the L2 acquisition of French object clitic constructions: Evidence from Chinese- and Spanish-speaking learners. In K. Franich, K. M. Iserman, and L. L. Keil (Eds.) *Proceedings of the 34th Annual Boston University Conference on Language Development* (pp. 150-161). Somerville, MA: Cascadilla Press.

Harding, M. C. and Hausman, J. (2007). Using a Laplace approximation to estimate the random coefficients logit model by non-linear least squares. *International Economic Review*, 48(4), 1311-1328.

Harrington, M. (1987). Processing transfer: Language-specific processing strategies as a source of interlanguage variation. *Applied Psycholinguistics*, 8, 351-377.

Harrington, M. and Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25-38.

Haut, M. W., Kuwabara, H., Leach, S., and Arias, R. G. (2000). Neural activation during performance of number-letter sequencing. *Applied Neuropsychology*, 7(4), 237-242.

Havik, E., Roberts, L., van Hout, R., Schreuder, R., and Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, 59(1), 73-112.

Hemforth, B. and Konieczny, L. (2000). *German sentence processing*. Kluwer.

Herschensohn, J. (2004). Functional categories and the acquisition of object clitics in L2 French. In P. Prévost and J. Paradis (Eds.), *The Acquisition of French in Different Contexts: Focus on functional categories* (pp. 207-242). Amsterdam: Benjamins.

Heilenman, L. K. and McDonald, J. L. (1993). Processing strategies in L2 learners of French: The role of transfer. *Language Learning*, 43(4), 507-557.

Hilchey, M. D. and Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomics Bulletin Review*, 18, 625-658.

Hopp, H. (2006). Subject-object ambiguities in L2 parsing. In D. Bamman, T. Magnitskaia, and C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (281-292). Somerville, MA: Cascadilla Press.

Isabelli, C. (2008). First Noun Principle or L1 Transfer Principle in SLA? *Hispania*, 91 (2), 465-478.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., and Kettermann, A. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, B47-B57.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434-446.

Juffs, A. (1998). Some effects of first language argument structure and morphosyntax on second language sentence processing. *Second Language Research*, 14(4), 406-424.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, 102, 199-255.

Juffs, A. (2005). The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research*, 21(2), 121-151.

Juffs, A. and Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in *wh*-extractions. *Studies in Second Language Acquisition*, 17, 483-516.

Juffs, A. and Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44(2), 137-166.

Just, M. A. and Carpenter, M. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329-354.

Just, M. A. and Carpenter, M. (1992). A capacity theory of comprehension: Individual differences in WM. *Psychological Review*, 99(1), 122-149.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., and Engle, R. W. (2007). Variation in Working Memory Capacity as Variation in Executive Attention and Control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (Eds.), *Variation in Working Memory* (pp. 21-48). New York: Oxford University Press.

Kane, M. J. and Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 333-358.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., and Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2), 189-217.

Kilborn, K. (1989). Sentence processing in a second language: The timing of transfer, *Language and Speech*, 32(1), 1-23.

Kilborn, K. and Ito, T. (1989). Sentence processing in Japanese-English and Dutch-English bilinguals. In B. MacWhinney and E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 257 – 291). New York: Cambridge University Press.

King, J. and Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language,* 30, 580-602.

Klein, W. (1998). The contribution of second language acquisition research. *Language Learning*, 48, 527-550.

Kormos, J. and Safar, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11(2), 261-271.

Kroll, J. F., Michael, E. B., Tokowicz, N., and Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, 18, 137-171.

Kruschke, J. K. and Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin and Review*, 7, 636-45.

Kyllonen, P. C. and Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14, 389-433.

Kyllonen, P. C. and Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning & Individual Differences*, 2, 129-160.

Lado, R. (1957). *Linguistics across Cultures: Applied Linguistics for Language Teachers.* Ann Arbor, MI: University of Michigan Press.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning*, 26, 125-34.

Li, P., Bates, E., Liu, H., and MacWhinney, B. (1992). Cues as Functional Constraints on Sentence Processing in Chinese. In H-C. Chen and O. J. L. Tzeng (Eds), *Language Processing in Chinese* (pp. 207-234). Amsterdam, The Netherlands: Elsevier Science Publishers B. V.

Li, P., Farkas, I., and MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, 17, 1345-62

Liceras, J. (1985). The value of clitics in non-native Spanish. *Second Language Research*, 1, 151-168.

Liu, H., Bates, E., and Li, P. (1992). Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics,* 13, 451-484.

LoCoco, V. (1982). The role of word order as a cue in L2 comprehension. Paper presented at the 11th annual Linguistics Symposium, University of Wisconsin-Milwaukee, March.

LoCoco, V. (1987). Learner Comprehension of Oral and Written Sentences in

German and Spanish: The Importance of Word Order. In B. VanPatten, T. Dvorak, and J. Lee (Eds.), *Foreign Language Learning: A Research Perspective* (pp. 119-129). Rowley, MA: Newbury House.

Luk, G. C. (2008). The anatomy of the bilingual influence on cognition: Levels of functional use and proficiency of language. Ph.D. dissertation (Unpublished), York University, Canada.

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., and Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition*, 74, 347-57.

Luís, A. R. (2004). Mixed clitic placement: evidence from Romanian. *Proceedings of the thirty-third Western Conference On Linguistics*, 16, 199-211. University of Southern California, LA.

Mackey, A., Philp, J., Egi, T., Fujii, A., Tatsumi, T. (2002). Individual differences in WM, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181-210). Amsterdam: John Benjamins.

MacDonald, M. C. and Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, 109(1), 35-54.

MacDonald, M. C., Pearlmutter, N., and Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.

MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin*, 109(2), 163-203.

MacWhinney, B. (1987). Applying the Competition Model to Bilingualism. *Applied Psycholinguistics*, 8, 315-327.

MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan, F. Eckman, and M. Noonan (Eds.), *Linguistic Categorization* (pp. 195-242). New York: John Benjamins.

MacWhinney, B. (1992). Transfer and competition in second language learning. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 371-390). Amsterdam: Elsevier.

MacWhinney, B. (1997). SLA and the Competition Model. In A. M. B. de Groot and J. F. Kroll (Eds.), *Tutorials in Bilingualism* (pp. 119-129). Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. and Bates, E. (1978). Sentential Devices for Conveying Giveness and Newness: A Cross-Cultural Developmental Study. *Journal of Verbal Learning and Verbal Behavior*, 17.5, 539-558.

MacWhinney, B. and Bates, E. (1989). *The cross-linguistic study of sentence processing*. New York: Cambridge University Press.

MacWhinney, B., Bates, E., and Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23, 127-150.

MacWhinney, B. and Pléh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition*, 29, 95-141.

MacWhinney, B., Pléh, C., and Bates, E. (1985). The development of sentence interpretation in Hungarian. *Cognitive Psychology*, 17(2), 178-209.

Marin, S. (2004). *Pronominal Clitic Positioning is Morpho-Phonological: Evidence from Romanian Adult and Acquisition Data.* Yale University, unpublished manuscript.

Marinis, T., Roberts, L., Felser, C., and Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27, 53-78.

McClelland, J. L. 2001. 'Failures to learn and their remediation: A Hebbian account' in J. L. McClelland and R. S. Siegler (eds): *Mechanisms of Cognitive Development: Behavioral and Neural Perspectives*. Carnegie Mellon symposia on cognition. Mahwah, NJ: Erlbaum, pp. 97–121.

Mewaldt, S. P., Connelly, C. L., & Simon, J. R. (1980). Response selection in choice reaction-time: Test of a buffer model. *Memory & Cognition, 8,* 606-611.

Miller, G. (1955). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 63(2), 343–355.

Miyake, A. and Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy and L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339-365). London: Lawrence Erlbaum Associates.

Monachesi, P. (1998). The Morphosyntax of Romanian cliticization. In P. Coppen, H. van Halteren, and L. Teunissen (Eds.), *Computational Linguistics in The Netherlands 1997* (pp. 99-118). Amsterdam-Atlanta: Rodopi.

Nevins, A. and Săvescu, O. (2010). An apparent 'number case constraint' in Romanian: The role of syncretism. In K. Arregi, Z. Fagyal, S. A. Montrul, and A. Trembley (Eds.) *Romance Linguistics 2008: Interactions in Romance. Selected Papers from the 38th Linguistic Symposium on Romance Languages (LSRL), Urbana-Champaign, April 2008* (pp.185-200). John Benjamins.

Olivella De Castells, M., Guzmán, E., Lapuerta, P., and Liskin-Gasparro, J. E. (2010). *Mosaícos*. NJ: Prentice Hall.

Osaka, M. and Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30(4), 287-289.

Osterhout, L. and Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Languages*, 31, 785-806.

Pléh, C. (1989). The Development of Sentence Interpretation in Hungarian. In B. MacWhinney and E. Bates (Eds.), *The Crosslinguistic Study of Sentence Processing* (158-184). Cambridge: Cambridge University Press.

Popescu, A. (2000). The morphophonology of the Romanian clitic sequence. *Lingua*, 110, 773-799.

Prévost, P. (2006). The phenomenon of object omission in child L2 French. *Bilingualism: Language and Cognition*, 9(3), 263-280.

R development core team (2005). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org

Rescorla, R. A., and Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black and W. F. Prokasy (Eds.), *Classical Conditioning II: Current Theory and Research*. New York: Appleton-Century-Crofts, 64-99.

Richardson, J. T. E. (1996). Evolving issues in working memory. In J. T. E. Richardson, R. W. Engle, L. Hasher, R. H. Logie, E. R. Stoltzfus, and R. T. Zacks (Eds.), *WM and human cognition* (pp.120-154). New York: Oxford University Press.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47(1), 45-99.

Sagarra, N. (2007a). Online processing of gender agreement in low proficient English-Spanish late bilinguals. In J. Camacho, N. Flores-Ferrán, L. Sánchez, V. Déprez, and M-J. Cabrera (Eds.), *Current issues in linguistic theory series* (pp. 240-253). Amsterdam: John Benjamins.

Sagarra, N. (2007b). From CALL to face-to-face interaction: The effect of computer-delivered recasts and working memory on L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 229–248). New York: Oxford University Press.

Sagarra, N. (2008). Working memory and L2 processing of redundant grammatical forms. In Z. Han (Ed.) in collaboration with E. S. Park, *Understanding second language process* (pp. 133-147). Clevedon, UK: Multilingual Matters.

Sagarra, N. and Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, 120, 2022-2039.

Sánchez, L. and Al-Kasey, T. (1999). L2 acquisition of Spanish direct objects. *Spanish Applied Linguistics*, 3(1), 1-32.

Săvescu Ciucivara, O. (2009). *A Syntactic Analysis of Pronominal Clitic Clusters in Romance: The View from Romanian.* Ph.D. dissertation (Unpublished). New York University.

Schneider, W. and Detweiler, M. (1987). A connectionist/control architecture for working memory. *The Psychology of learning and motivation*, 21, 53-119.

Schriefers, H., Friederici, A. D., and Kühn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, 34, 499-520.

Simon, J. R. (1969). Reactions towards the source of stimulation. *Journal of Experimental Psychology*, 81, 174–176.

Simon, J. R., and Acosta, E., Jr. (1982). Effect of irrelevant information on the processing of relevant information: Facilitation and/or interference? The influence of experimental design. *Perception & Psychophysics, 31,* 383-388.

Simon, J. R., Craft, J. L., and Small, A. M. (1971). Reactions toward the apparent source Of an auditory stimulus. *Journal of Experimental Psychology, 89,* 203-206.

Simon, J. R., Sly, P. E., and Vilapakkam, S. (1981). Effect of compatibility of S-R mapping on reaction toward the stimulus source. *Acta Psychologica, 47,* 63-81.

Skehan, P. (1989). *Individual differences in second-language learning.* London: Edward Arnold.

Slobin, D. I. and Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12, 229-265.

Sokolov, J. L. (1988). Cue validity in Hebrew sentence comprehension. *Journal of Child Language*, 15, 129-156.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Tanenhaus, M. (2007). Eye movements and spoken language processing. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, and R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 443-469). Amsterdam: Elsevier.

Tockowicz, N., Michael, E. B., and Kroll, J. F. (2004). The roles of study abroad experience and working memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, 7, 255-272.

Tomasello, Michael. (2003). Introduction: Some Surprises for Psychologists. In *Author* (Ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure, Vol. 2* (pp. 1-14). Mahwah, NJ: Lawrence Erlbaum Associates.

Torrens, V., Wexler, K. (2000) The acquisition of clitic doubling in Spanish. In C. Hamann and S. Powers (Eds.) *The Acquisition of Scrambling and Cliticization*. Dordrecht: Kluwer, 279-297.

Turner, M. L. and Engle, R. W. (1986) Working memory. *Proceedings of the Human Factors Society*, 30, 1273-1277.

Turner, M. L. and Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.

Van den Noort, M. W. M. L., Bosch, M. P. C., and Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist*, 11(4), 289-296.

VanPatten, B. (1984). Learners' Comprehension of Clitic Pronouns: More Evidence for a Word-Order Strategy. *Hispanic Linguistics*, 1, 57-67.

VanPatten, B. (1996). *Input Processing and Grammar Instruction in Second Language Acquisition.* Norwood, NJ: Ablex Publishing Corporation.

VanPatten, B. (2004). Input Processing in Second-Language Acquisition. In B. VanPatten (Ed.), *Processing Instruction: Theory, Research, and Commentary* (pp. 5-31). Mahwah, NJ: Lawrence Erlbaum Associates.

VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten and J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 115-135). Mahwah, NJ: Lawrence Erlbaum Associates.

VanPatten, B. and Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition,* 15, 225–243.

Vos, S. H., Gunter, T. C., Schriefers, H., and Friederici, A. D. (2001). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes*, 16(1), 65-103.

Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics*, 25(3), 315-339.

Waters, G. S. and Caplan, D. (1996a). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, 103(4), 761-772.

Waters, G. S. and Caplan, D. (1996b). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology*, 49A, 51-79.

Waters, G. S. and Caplan, D. (1996c). Processing resource capacity and the comprehension of garden path sentences. *Memory and Cognition*, 24, 342-355.

Wechsler, D. A. (1997). *WAIS–III, WMS–III technical manual*. San Antonio, TX: Psychological Corporation.

Werker, J. E. and Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, 24, 672-83.

Werker, J. E. and Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63.

Williams, J. N. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition*, 21, 1-48.

Williams, J. N. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition,* 9(1), 71-88.

Wright, C. (2010). Variation, asymmetry and working memory in the process of second language acquisition. *Proceedings of the 34th annual Boston University Conference on Language Development*, 34(2), 468-479.

Zagona, K. (2002). *The Syntax of Spanish.* Cambridge, UK: Cambridge University Press.

Zurif, E., Swinney, D., Prather, P., Wingfield, A. and Brownell, H. (1995). The allocation of memory resources during sentence comprehension: Evidence from the elderly. *Journal of Psycholinguistic Research,* 24(3), 165–182.

Zwicky, A. M. and Pullum, G. K. (1983). Cliticization vs. Inflection: English N'T. *Language*, 59(3), 502-513.

APPENDIX A. Language History Questionnaire

Language History Questionnaire

First name (print): ____________________ Last name (print): ______________________
Gender: M/F Age: _________________________
Your native language: _________________________
Language(s) spoken at home: _________________________
Language(s) spoken during first five years of your life: _________________________
Language(s) studied in:
School: _________________________ University/other: _______________________
Language(s) you use at home, at school, at work, or with friends/family___________________
Other language(s) you can:
Read: _______________ Write ________________ Speak _______________________
Which language do you feel most comfortable with at this time?_____________________
Contact with Spanish:
At what age were you first exposed to Spanish? _________________________
Approximate hours per week that you are in contact with Spanish: __________________
Have you ever lived in a Spanish-speaking country? Yes/No
If yes, Where? ______________ When? ____________ For how long? ______________

APPENDIX B. Spanish Proficiency Test.

<u>PART I.</u> Write the letter of the word or phrase that best completes the meaning of the sentence. If you do not know the meaning of a word in Spanish, <u>underline</u> it. "0" means that you do not think that anything goes in the blank.

GRAMÁTICA (NIVEL BÁSICO)

___ 1. ___________ edificio alto es la Torre 'Sears'.
A. Eso B. La C. Aquel D. 0

___ 2. Los autos que chocaron en el accidente iban ___________ el oeste.
A. dentro B. hacia C. fuera D. 0

___ 3. Los novios pasaron unas vacaciones fantásticas __________ fueron a Hawaii.
A. cuando B. que C. donde D. 0

___ 4. --¿Van a invitar al profesor y a su esposa a la reunión? –Sí, vamos a invitar ________.
A. ellos B. sus C. los D. 0

___ 5. Si no puedes usar tu bicicleta usa ___________.
A. nuestra B. de él C. la mía D. 0

___ 6. A Juana no ________ gustan las películas de ciencia ficción.
A. le B. se C. la D. 0

___7. En nuestro barrio hay muchas casas bonitas, pero _____ Juan es la más bonita.
A. su B. de la C. la de D. 0

___8. --¿Conoces _______ hombre de la camisa verde? --¿Es muy guapo verdad?
A. un B. al C. esto D. 0

___9. Oscar no va a graduarse este semestre, ni yo ________.
A. tampoco B. ningún C. además D. 0

__10. --¿Con quién saliste al bar anoche? –No salí con ______; fui sola.
A. tú B. alguien C. nadie D. 0

___11. Estamos comprando _______ pan francés para la cena de mañana.
A. la B. hay C. algo D. 0

___12. La palabra 'venir' viene _________ Latín.
A. por B. en C. del D. 0

GRAMÁTICA (NIVEL INTERMEDIO)

___ 1. Por favor, __________ llegues a Madrid, me llamas.
A. desde que B. antes de C. cuando D. después de

___ 2. – ¿Hasta qué hora estuvo Lorenzo en la consulta?
– Pues no sé, no lo vi. Cuando yo llegué, a las 12, ya se __________.
A. iba B. ha ido C. fue D. había ido

___ 3. Hoy invito yo __________ todos al café, que es mi cumpleaños.
A. para B. de C. a D. sobre

___ 4. ¿__________ has pedido ya a tus padres?
A. Se te B. Se lo C. Se les D. Se le

___ 5. Manuel, como no __________ más fruta, no tendremos suficiente.
A. compres B. compras C. compraras D. comprarás

___ 6. ¿Que te vas a París? ¡Quién __________ tú!
A. es B. sea C. sería D. fuera

___7. Sinceramente, yo que tú __________ un mapa antes de viajar.
A. compraré B. compro C. compraría D. comprara

___8. La música de los vecinos está muy alta. Estoy __________ llamar a la policía.
A. a B. por C. entre D. tras

___9. El médico me dijo que __________ que volver mañana.
A. había tenido B. tuve C. tenía D. he tenido

___10. Por favor, en cuanto __________ a Lucía, dile que me llame.
A. verás B. veas C. ves D. vieras

___11. El regalo que __________ he comprado a Andrés es muy bonito.
A. lo B. se C. la D. le

___12. El profesor me pidió que _________ a sus horas de oficina.
A. iré B. vaya C. iría D. iba

GRAMÁTICA (NIVEL AVANZADO)

___ 1. Ellos estaban dispuestos a que __________ nosotros en el coche y ellos andando.
A. íbamos B. fuimos C. iríamos D. fuéramos

___ 2. __________ como se enteraron de lo sucedido fueron a visitar a la familia.
A. Tan pronto B. No bien C. En cuanto D. Nada más

___ 3. Elisa llegó a la estación cuando el tren __________ de salir, ¡qué rabia!
A. acabó B. acaba C. acabaría D. acababa

___ 4. En cuanto deje la maleta en la habitación del hotel __________ meterme en la piscina, ¡qué calor!
A. creo B. debo C. pienso D. siento

___ 5. Carolina y Luis se casaron muy jóvenes, __________ cumplieron los 20 años.
A. al B. apenas C. de D. Pronto

___ 6. El perrito de María es muy gracioso, tan pronto salta __________ se tumba.
A. que B. de C. y D. como

___7. El jefe no se ha enfadado porque María _____ llegado tarde, sino porque no la había preparado bien.
A. ha B. haya C. había D. hubiera

___8. Al abuelo le encantaba que Juanito ___ a verle todos los días.
A. haya ido B. iba C.fuera D. iría

___9. Pedro va a hablar con el director, pero no quiere que ___ vaya con él.
A. algún B. alguien C. nadie D. todos

___10. Aunque ___ muy tarde, iré a verte al hospital, te lo prometo.
A. llegue B. llegara C. llegaría D. llegué

___11. Le dieron todo lo que pidió, ____ estuviera feliz y se quedara allí.
A. a saber B. por eso C. de ahí que D. por consiguiente

___12. Está __ nevar, así que abrígate bien.
A. para B. en C. si D. entre

PART II. Write the letter for the correct answer in the blank to the left of each item.

COMPRENSIÓN ESCRITA (NIVEL INTERMEDIO)
Las bicicletas también son para el otoño

El ciclismo está considerado por los especialistas como uno de los deportes más completos. Fortalece el cuerpo y también la mente, y a él puede __1__ cualquier persona porque no tiene __2__ de edad. La bicicleta es uno de los mejores deportes, sobre todo para la gente __3__ no puede hacer ejercicios de contacto con el suelo, como correr.

__4__ estemos ante un deporte muy beneficioso, ya que no sólo mejora nuestra condición física, a la vez que nos hace más resistentes: __5__ tiene unos efectos anímicos extraordinarios. Elimina el estrés y hace que __6__ más eufóricos y enérgicos, __7__ supone encontrarnos mejor. Un último elemento que añadir para lograr este óptimo estado es el contacto con la naturaleza.

Para practicar ese deporte, debemos __8__ en cuenta algunos aspectos. El tiempo es una de las dificultades con __9__ que se cuenta si se vive en la ciudad. Hay que intentar sacar tiempo de __10__ sea para poder practicar nuestro deporte preferido. En el caso de la bici, lo ideal es salir todos los días aunque sólo __11__ un cuarto de hora, si bien se recomienda pedalear __12__ 40 y 45 minutos. También se pueden realizar tres sesiones a la semana __13__ a los 60 minutos, y los fines de semana __14__ de entrenar un poco más porque tenemos más tiempo libre. La distancia a recorrer dependerá __15__ la velocidad y el ritmo que __16__, aunque no hay que obsesionarse con los kilómetros. Otro elemento __17__ importante es la elección que hagamos de la bicicleta: de carretera para los más deportivos, de montaña para los __18__ de la naturaleza, y las híbridas, que valen para todo.

Con la bici ya escogida, sólo __19__ resta equiparnos adecuadamente. En el atuendo no debe __20__ un buen *culotte*, un *maillot*, un chubasquero por si llueve, y un casco.

___ 1. a) acceder b) practicar c) ejecutar
___ 2. a) límite b) término c) frontera

___ 3.	a) quien	b) quienes	c) que
___ 4.	a) De modo que	b) De ahí que	c) Así que
___ 5.	a) pero	b) sino	c) también
___ 6.	a) estamos	b) estemos	c) estaremos
___ 7.	a) lo que	b) el cual	c) cuyo
___ 8.	a) tener	b) considerar	c) darnos
___ 9.	a) lo	b) las	c) la
___ 10.	a) donde	b) como	c) cuando
___ 11.	a) sería	b) es	c) sea
___ 12.	a) entre	b) hacia	c) de
___ 13.	a) alrededor	b) en torno	c) cerca
___ 14.	a) tratar	b) intentar	c) esforzarse
___ 15.	a) en	b) de	c) a
___ 16.	a) corramos	b) vayamos	c) llevemos
___ 17.	a) más	b) tan	c) muy
___ 18.	a) amantes	b) aficionados	c) interesados
___ 19.	a) se	b) nos	c) le
___ 20.	a) faltar	b) sobrar	c) quedar

COMPRENSIÓN AUDITIVA (NIVEL INTERMEDIO)

PART III. Escucha el texto 2 veces y responde a estas preguntas:
http://diplomas.cervantes.es/docs/ficheros/200906180001_7_23.mp3

___1. En la audición, el doctor Becerra afirma que los chicles…
A. blanquean los dientes
B. ayudan a mantener el color de los dientes
C. no son recomendables

___2. El doctor Becerra opina que…
A. las técnicas de blanqueamiento caseras no son las más adecuadas.
B. el proceso de blanqueamiento es conveniente en cualquier caso o situación.
C. la belleza de los dientes reside en su forma y armonía.

___3. En esta audición se dice que el blanqueamiento de dientes debe hacerse…
A. siempre que sea necesario.
B. sólo una vez en la vida.
C. cada cinco meses.

___4. El doctor Becerra dice que…
A. todos los chicles son malos
B. sólo los chicles con azúcar son malos
C. es mejor comer chicles sin ingredientes que blanqueen los dientes

APPENDIX C. English Proficiency Test

Write the letter for the correct answer in the blank to the left.

___ 1. Mr. Jones ...... the company since 1990
a runs b. is running c. has run d. ran

___ 2. The bookshop ..... next to the town hall
a. is put b. has the post c. position d. is located

___ 3. You will find more information in the .....
a. attached file b. attached to file c. file what is attached d. attachment file

___ 4. There has been an increase in the price of oil in ...... weeks.
a. recent b. just passed c. some d. Yesterday's

___ 5. Give me three minutes and I'll ....
a. return you call b. call you in back c. get back to you d. be phoned back

___ 6. You ...... to register for this course.
a. mustn't b. shouldn't c. don't need d. can have

___ 7. We have been working hard. Let's ...... a break.
a. do b. take c. make d. find

___ 8. ..... others ahead of yourself is easy to say, but harder to do.
a. Put b. Putting c. The putting d. For putting

___ 9. As you drive down the road, take the ......, and you will see his house ahead.
a. left second turn b. left turn c. second left turn d. left turning

___ 10. Can you tell me how to ...... the airport?
a. arrive in b. achieve c. get to d. attain

___ 11. The .... state of the US, Alaska, was purchased from Russia in the last century.
a. north b. northernmost c. northerly d. northeast

___ 12. ..... people know the town better than old Jake here.
a. The few b. Only the few c. Only few d. Few

___ 13. You must boil those vegetables before .... in the stew.
a. using them b. their used c. the use d. using

___ 14. The first congress .... in 1776.
a. was hold b. were held c. took place d. took over

___ 15. She got married .... while on holiday in Hawaii.
a. secretly b. together c. unexpected d. with Tom

___ 16. What did you want to do that ....?

a. reason b. for c. because d. thing

___ 17. Spring is .... It will be warmer soon.

a. on the way b. to be coming c. eventually d. prepared now

___ 18. The park is named ...... the town's first mayor.

a. in respect of b. owing to c. in honor of d. of the memory of

___ 19. ...... you paid me twice the salary, I wouldn't take that job.

a. Although b. Despite c. Though d. Even if

___ 20. It's raining again. It's ...... the weather improved.

a. in time b. for time c. the time d. about time

APPENDIX D. Vocabulary Test

Verbs: Match the Spanish verb on the left with the appropriate English verb on the right by writing the corresponding letter in the blank provided.

______	1.	preguntar	A. to yell	______	1.	necesitar	A. to recieve
______	2.	firmar	B. to climb	______	2.	apagar	B. to sweep
______	3.	compartir	C. to ask	______	3.	mejorar	C. to look at
______	4.	facturar	D. to guess	______	4.	gastar	D. to hug
______	5.	escalar	E. to keep	______	5.	recibir	E. to turn off/put out
______	6.	escuchar	F. to sign	______	6.	cepillar	F. to cut
______	7.	llevar	G. to listen to	______	7.	cancelar	G. to spend
______	8.	dibujar	H. to investigate	______	8.	cortar	H. to record
______	9.	cocinar	I. to send	______	9.	aprender	I. to call
______	10.	imprimir	J. to fill	______	10.	barrer	J. to need
______	11.	adivinar	K. to change	______	11.	grabar	K. to cancel
______	12.	gritar	L. to draw	______	12.	saltar	L. to learn
______	13.	guardar	M. to cook	______	13.	buscar	M. to understand
______	14.	limpiar	N. to collect/pick up	______	14.	abrazar	N. to wake up
______	15.	investigar	O. to check	______	15.	llamar	O. to improve
______	16.	cambiar	P. to clean	______	16.	mirar	P. to kiss
______	17.	recoger	Q. to carry	______	17.	comprender	Q. to brush
______	18.	esperar	R. to print	______	18.	despertar	R. to look for
______	19.	llenar	S. to share	______	19.	besar	S. to follow
______	20.	mandar	T. to wait for	______	20.	seguir	T. to jump

Adverbs: Now do the same for the following adverbs.

______	1.	ayer	A. now	______	1.	la semana pasada	A. today
______	2.	anoche	B. last night	______	2.	el mes pasado	B. last year
______	3.	ahora	C. yesterday	______	3.	hoy	C. last month
______	4.	anteayer	D. the day before yesterday	______	4.	el año pasado	D. last week

Nouns: Now do the same for the following nouns.

______	1.	tío	A. girlfriend	______	1.	compañía	A. aunt
______	2.	nieta	B. son	______	2.	hombre	B. lawyer
______	3.	hijo	C. uncle	______	3.	abuela	C. older gentleman
______	4.	niña	D. grandaughter	______	4.	hija	D. manager
______	5.	novia	E. boys	______	5.	tía	E. boy
______	6.	periodista	F. waitress	______	6.	esposa	F. worker
______	7.	alumno	G. spy	______	7.	gerente	G. man
______	8.	camarera	H. nephew	______	8.	empleado	H. daughter
______	9.	dueño	I. detective	______	9.	niño	I. cook
______	10.	esposo	J. girl	______	10.	anciano	J. company
______	11.	chicos	K. female student	______	11.	muchacha	K. employee
______	12.	espía	L. male student	______	12.	perro	L. wife

______	13. sobrino	M. journalist	______	13. abogado	M. grandmother
______	14. detective	N. owner	______	14. cocinero	N. sister
______	15. cirujano	O. husband	______	15. hermana	O. cousin
______	16. alumna	P. singer	______	16. trabajador	P. woman
______	17. amiga	Q. surgeon	______	17. mujer	Q. girl
______	18. cantante	R. grandfather	______	18. prima	R. witness
______	19. abuelo	S. female friend	______	19. testigo	S. dog

Nouns: Now do the same for the following nouns.

______	1. noticias	A. mobile phone	______	1. palos	A. answer
______	2. dinero	B. boots	______	2. suelo	B. teeth
______	3. codigo	C. money	______	3. vestido	C. apartment
______	4. opinion	D. French	______	4. bolígrafo	D. guitar
______	5. voz	E. car	______	5. robo	E. gift
______	6. coche	F. jumprope	______	6. jefe	F. floor
______	7. botas	G.news	______	7. coche	G. car
______	8. móvil	H. voice	______	8. piso	H. robbery
______	9. francés	I. gift	______	9. dientes	I. golf clubs
______	10. respuesta	J. code	______	10. guitarra	J. boss
______	11. paraguas	K. gas	______	11. nombre	K. dress
______	12. instrucciones	L. flight	______	12. corazón	L. lights
______	13. cuerda	M. contract	______	13. pollo	M. name
______	14. regalo	N. opinion	______	14. caja	N. pen
______	15. gasolina	O.umbrella	______	15. vaso	O. glass
______	16. fuego	P. answer	______	16. luces	P. heart
______	17. contrato	Q. fire	______	17. regalo	Q. box
______	18. vuelo	R. instructions	______	18. respuesta	R. chicken

APPENDIX E. Listening Task Stimuli

PRACTICE	Por la tarde la niña golpea al niño en la sala.
Item 1 Condition 1	Por la noche la abuela busca al abuelo en la calle.
Item 1 Condition 2	Por la noche lo busca la abuela en la calle.
Item 2 Condition 1	Por la mañana el esposo busca a la esposa en la iglesia.
Item 2 Condition 2	Por la mañana la busca el esposo en la iglesia.
Item 3 Condition 1	De repente la hija abraza a la tía en el comedor.
Item 3 Condition 2	De repente la abraza la hija en el comedor.
Item 4 Condition 1	En este momento el sobrino abraza al tío en la calle.
Item 4 Condition 2	En este momento lo abraza el sobrino en la calle.
Item 5 Condition 1	Por la noche la nieta llama a la abuela desde la habitación.
Item 5 Condition 2	Por la noche la llama la nieta desde la habitación.
Item 6 Condition 1	Normalmente el hombre llama a la mujer después del trabajo.
Item 6 Condition 2	Normalmente la llama el hombre después del trabajo.
Item 7 Condition 1	Hoy el abogado comprende al testigo en la oficina.
Item 7 Condition 2	Hoy lo comprende el abogado en la oficina.
Item 8 Condition 1	En este momento el perro sigue al hombre en el bosque.
Item 8 Condition 2	En este momento lo sigue el perro en el bosque.
Item 9 Condition 1	En este instante la abuela sigue a la tía en el barrio.
Item 9 Condition 2	En este instante la sigue la abuela en el barrio.
Item 10 Condition 1	Normalmente el trabajador comprende al jefe durante las reuniones.
Item 10 Condition 2	Normalmente lo comprende el trabajador durante las reuniones.
Item 11 Condition 1	Por la tarde el tío despierta a la tía de la siesta
Item 11 Condition 2	Por la tarde la despierta el tío de la siesta.
Item 12 Condition 1	Por la mañana la prima despierta al primo con la amiga.
Item 12 Condition 2	Por la mañana lo despierta la prima con la amiga.
Item 13 Condition 1	Por la tarde la amiga besa al amigo para la graduación.
Item 13 Condition 2	Por la tarde lo besa la amiga para la graduación.
Item 14 Condition 1	De repente la novia besa al novio en la boca.

Item 14 Condition 2	De repente lo besa la novia en la boca.
Item 15 Condition 1	En este instante la cantante mira a la periodista durante la entrevista.
Item 15 Condition 2	En este instante la mira la cantante durante la entrevista.
Item 16 Condition 1	Hoy el hermano mira a la hermana en el comedor.
Item 16 Condition 2	Hoy la mira el hermano en el comedor.

APPENDIX F. Listening Task Pictures

POL01a

POL01b

POL01c

POL01d

OL01a

OL01b

OL01c

OL01d

OL02a

OL02b

OL02c

OL02d

OL03a

OL03b

OL03c

OL03d

OL04a

OL04b

OL04c

OL04d

OL05a

OL05b

OL05c

OL05d

OL06a

OL06b

OL06c

OL06d

OL07a

OL07b

OL07c

OL07d

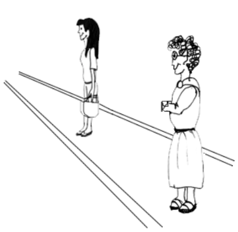
OL08a

OL08b

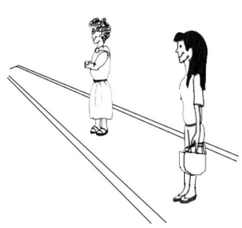
OL08c

OL08d

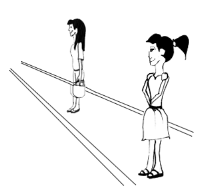
OL09a



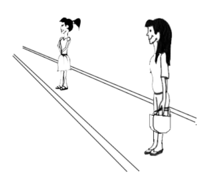
OL09b



OL09c



OL09d



OL10a

OL10b

OL10c

OL10d

OL11a

OL11b

OL11c

OL11d

OL12a

OL12b

OL12c

OL12d

OL13a

OL13b

OL13c

OL13d

OL14a

OL14b

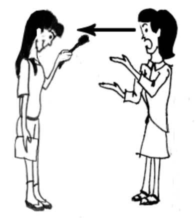
OL14c

OL14d

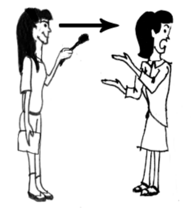
OL15a

OL15b

OL15c



OL15d



OL16a

OL16b

OL16c

OL16d

APPENDIX G. Reading Task Stimuli

PRACTICE	Hoy lo odia la niña por sus travesuras.
Item 1 Condition 1	Hoy el chico busca a la chica en\nel parque.
Item 1 Condition 2	Hoy la busca el chico en\nel parque.
Item 2 Condition 1	En este momento la niña busca al niño en\nla tienda.
Item 2 Condition 2	En este momento lo busca la niña en\nla tienda.
Item 3 Condition 1	Por la noche el hombre abraza al niño en la tienda.
Item 3 Condition 2	Por la noche lo abraza el hombre en la tienda.
Item 4 Condition 1	De repente el nieto abraza al abuelo en el comedor.
Item 4 Condition 2	De repente lo abraza el nieto en el comedor.
Item 5 Condition 1	Normalmente la abogada llama a la periodista desde\nla oficina.
Item 5 Condition 2	Normalmente la llama la abogada desde\nla oficina.
Item 6 Condition 1	Por la tarde el novio llama a la novia con\nsu teléfono.
Item 6 Condition 2	Por la tarde la llama el novio con\nsu teléfono.
Item 7 Condition 1	Normalmente la abuela comprende al abuelo durante\nlas peleas.
Item 7 Condition 2	Normalmente lo comprende la abuela durante\nlas peleas
Item 8 Condition 1	En este instante el cocinero sigue al hombre en la cocina.
Item 8 Condition 2	En este instante lo sigue el cocinero en la cocina.
Item 9 Condition 1	Normalmente la chica sigue al chico en\nel coche.
Item 9 Condition 2	Normalmente lo sigue la chica en\nel coche.
Item 10 Condition 1	Hoy la niña comprende a la tía durante\n el almuerzo.
Item 10 Condition 2	Hoy la comprende la niña durante\n el almuerzo.
Item 11 Condition 1	Por la mañana el esposo despierta a la esposa en\nla cama.
Item 11 Condition 2	Por la mañana la despierta el esposo en\nla cama.
Item 12 Condition 1	Por la tarde la abuela despierta a la nieta en\nla sala.
Item 12 Condition 2	Por la tarde la despierta la abuela en\nla sala.
Item 13 Condition 1	Por la noche la mujer besa al hombre en\n la cama.
Item 13 Condition 2	Por la noche lo besa la mujer en\n la cama.
Item 14 Condition 1	En este momento el abuelo besa al muchacho en\n la oficina.

Item 14 Condition 2	En este momento lo besa el abuelo en\n la oficina.
Item 15 Condition 1	De repente la alumna mira a la profesora en\nel aula.
Item 15 Condition 2	De repente la mira la alumna en\nel aula.
Item 16 Condition 1	En este instante el esposo mira a la esposa en\nla iglesia.
Item 16 Condition 2	En este instante la mira el esposo en\nla iglesia.

APPENDIX H. Reading Task Pictures

POR01a


POR01b


POR01c


POR01d

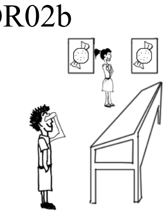

OR01a


OR01b

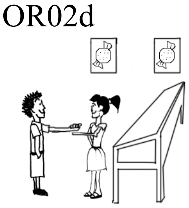

OR01c


OR01d


OR02a

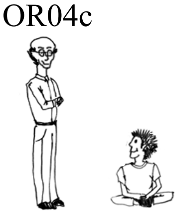

OR02b


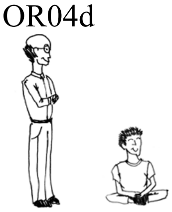
OR02c


OR02d


OR04a


OR04b


OR04c


OR04d


OR05a


OR05b


OR05c


OR05d


OR06a

OR06b

OR06c

OR06d

OR07a

OR07b

OR07c

OR07d

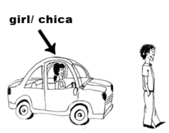
OR08a

OR08b

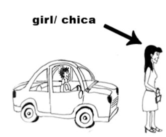
OR08c

OR08d

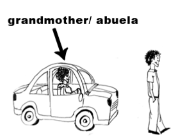
OR09a



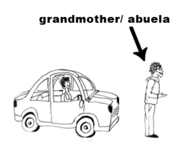
OR09b



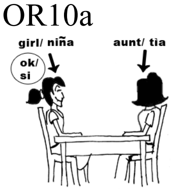
OR09c



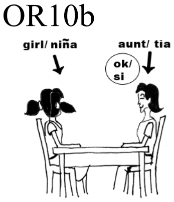
OR09d



OR10a



OR10b



OR10c

OR10d

OR11a

OR11b

OR11c

OR11d

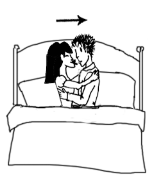
OR12a

OR12b

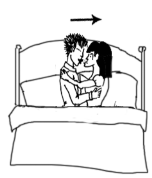
OR12c

OR12d

OR13a



OR13b



OR13c

OR13d

OR14a

OR14b

OR14c

OR14d

OR15a

OR15b

OR15c

OR15d

OR16a

OR16b

OR16c

OR16d

APPENDIX I. Monolingual Listening Task Model

```
Generalized linear mixed model fit by the Laplace approximation
Formula: Accuracy ~ (1 | SubjectID) + (1 | SentenceID) + Condition * cWM +
 cFlankInhib
   Data: list_mono
 Subset: PictureChosenifWrong < 3 & NativeSpeaker == "TRUE"
   AIC BIC logLik deviance
 130.1 160 -58.06    116.1
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 0.40307  0.63488
 SentenceID (Intercept) 1.67160  1.29290
Number of obs: 528, groups: SubjectID, 35; SentenceID, 16

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)       5.322946   0.931282   5.716 1.09e-08 ***
ConditionOVS     -0.331911   1.108667  -0.299    0.765
cWM               0.099248   0.254728   0.390    0.697
cFlankInhib      -0.006768   0.009839  -0.688    0.492
ConditionOVS:cWM  0.269041   0.321894   0.836    0.403
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) CndOVS cWM    cFlnkI
ConditinOVS -0.662
cWM          0.658 -0.539
cFlankInhib -0.125  0.023 -0.057
CndtnOVS:WM -0.488  0.785 -0.766 -0.023
```

APPENDIX J. Monolingual Distractor Items for Listening Task Model

```
Generalized linear mixed model fit by the Laplace approximation
Formula: distr_choices ~ (1 | SubjectID) + (1 | SentenceID) + Condition +      cWM +
 cFlankInhib
   Data: list_mono
   AIC   BIC logLik deviance
 199.1 225.1 -93.55    187.1
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 0.0000   0.0000
 SentenceID (Intercept) 6.8731   2.6217
Number of obs: 560, groups: SubjectID, 35; SentenceID, 16

Fixed effects:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   5.419216   0.888397   6.100 1.06e-09 ***
ConditionOVS -0.431225   0.427651  -1.008    0.313
cWM           0.097504   0.091367   1.067    0.286
cFlankInhib  -0.009149   0.006349  -1.441    0.150
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) CndOVS cWM
ConditinOVS -0.273
cWM          0.232  0.034
cFlankInhib -0.072 -0.054 -0.137
```

APPENDIX K. Learner Listening Task Model

```
Generalized linear mixed model fit by the Laplace approximation
Formula: Accuracy ~ (1 | SubjectID) + (1 | SentenceID) + Condition *
 ImmersionTimeinmonths +      Condition * FirstLanguage + Condition * cWM +
 cFlankInhib +      cProf * Condition
   Data: list_all
 Subset: PictureChosenifWrong < 3 & NativeSpeaker == "FALSE"
  AIC  BIC logLik deviance
 1397 1469 -685.3     1371
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 0.68485  0.82755
 SentenceID (Intercept) 0.11835  0.34402
Number of obs: 2012, groups: SubjectID, 133; SentenceID, 16

Fixed effects:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                  3.002912   0.307631   9.761  < 2e-16 ***
ConditionOVS                -2.014817   0.302011  -6.671 2.53e-11 ***
ImmersionTimeinmonths        0.002600   0.025541   0.102 0.918910
FirstLanguageROMANIAN        0.785236   0.378607   2.074 0.038078 *
cWM                          0.108538   0.068665   1.581 0.113946
cFlankInhib                 -0.005177   0.004506  -1.149 0.250630
cProf                        0.041047   0.012988   3.160 0.001576 **
ConditionOVS:ImmersionTimeinmonths -0.027092   0.026216  -1.033 0.301420
ConditionOVS:FirstLanguageROMANIAN -0.469723   0.383264  -1.226 0.220354
ConditionOVS:cWM            -0.232932   0.070168  -3.320 0.000901 ***
ConditionOVS:cProf           0.049439   0.013519   3.657 0.000255 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) CndOVS ImmrsT FLROMA cWM    cFlnkI cProf  COVS:I COVS:F COVS:W
ConditinOVS -0.752
ImmrsnTmnmn -0.533  0.468
FrLROMANIAN -0.601  0.468  0.216
cWM         -0.244  0.177  0.226  0.250
cFlankInhib -0.048 -0.001  0.014  0.010  0.052
cProf        0.472 -0.434 -0.368 -0.178 -0.109 -0.060
CndtnOVS:IT  0.446 -0.563 -0.859 -0.167 -0.192  0.003  0.314
COVS:FLROMA  0.456 -0.616 -0.171 -0.805 -0.192  0.019  0.148  0.212
CndtnOVS:WM  0.173 -0.270 -0.196 -0.189 -0.813  0.005  0.089  0.234  0.243
CndtnOVS:cP -0.407  0.514  0.312  0.140  0.082 -0.021 -0.796 -0.381 -0.175 -0.142
```

APPENDIX L. Learner Distractor Items for Listening Task Model

```
Generalized linear mixed model fit by the Laplace approximation
Formula: distr_choices ~ (1 | SubjectID) + (1 | SentenceID) + Condition +      FirstLanguage
 + cWM + cProf + ImmersionTimeinmonths + cFlankInhib
   Data: list_all
 Subset: NativeSpeaker == "FALSE"
   AIC   BIC logLik deviance
 771.7 822.7 -376.9    753.7
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 0.11369  0.33718
 SentenceID (Intercept) 2.11412  1.45400
Number of obs: 2128, groups: SubjectID, 133; SentenceID, 16

Fixed effects:
                       Estimate Std. Error z value Pr(>|z|)
(Intercept)            3.964659   0.444042   8.929  < 2e-16 ***
ConditionOVS           0.017597   0.204539   0.086 0.931442
FirstLanguageROMANIAN -0.251963   0.228939  -1.101 0.271084
cWM                   -0.014115   0.040643  -0.347 0.728377
cProf                  0.026033   0.007498   3.472 0.000516 ***
ImmersionTimeinmonths -0.011812   0.012937  -0.913 0.361219
cFlankInhib           -0.004250   0.004405  -0.965 0.334696
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) CndOVS FLROMA cWM    cProf  ImmrsT
ConditinOVS -0.226
FrLROMANIAN -0.333 -0.002
cWM         -0.158  0.009  0.290
cProf        0.163  0.001 -0.182 -0.129
ImmrsnTmnmn -0.214 -0.018  0.245  0.196 -0.313
cFlankInhib -0.045  0.012  0.028  0.115 -0.084  0.019
```

APPENDIX M. Monolingual Reading Task Model

```
Generalized linear mixed model fit by the Laplace approximation
Formula: Accuracy ~ (1 | SubjectID) + (1 | SentenceID) + Condition * cWM +
cFlankInhib
   Data: rdg_all
 Subset: PictureChosenifWrong < 3 & NativeSpeaker == "TRUE"
   AIC BIC logLik deviance
 281.7 312 -133.8    267.7
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 0.12639  0.35552
 SentenceID (Intercept) 0.61944  0.78705
Number of obs: 560, groups: SubjectID, 36; SentenceID, 16

Fixed effects:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)      3.245668   0.423556   7.663 1.82e-14 ***
ConditionOVS     0.329801   0.563387   0.585    0.558
cWM              0.094942   0.113405   0.837    0.402
cFlankInhib      0.001445   0.005739   0.252    0.801
ConditionOVS:cWM 0.255458   0.158986   1.607    0.108
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) CndOVS cWM    cFlnkI
ConditinOVS -0.544
cWM          0.592 -0.423
cFlankInhib  0.034 -0.026 -0.084
CndtnOVS:WM -0.399  0.772 -0.668 -0.026
```

APPENDIX N. Monolingual Distractor Items for Reading Task

```
Generalized linear mixed model fit by the Laplace approximation
Formula: distr_choices ~ (1 | SubjectID) + (1 | SentenceID) + Condition +  cWM +
cFlankInhib
   Data: rdg_all
 Subset: NativeSpeaker == "TRUE"
   AIC   BIC logLik deviance
 140.6 166.7  -64.3    128.6
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 2.5150   1.5859
 SentenceID (Intercept) 1.7363   1.3177
Number of obs: 576, groups: SubjectID, 36; SentenceID, 16

Fixed effects:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   5.624367   0.806255   6.976 3.04e-12 ***
ConditionOVS  0.037502   0.612734   0.061    0.951
cWM           0.159036   0.198615   0.801    0.423
cFlankInhib  -0.002855   0.013501  -0.211    0.833
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) CndOVS cWM
ConditinOVS -0.366
cWM          0.557 -0.013
cFlankInhib -0.077  0.026 -0.214
```

APPENDIX O. Learner Reading Task Model

```
Generalized linear mixed model fit by the Laplace approximation
Formula: Accuracy ~ (1 | SubjectID) + (1 | SentenceID) + Condition * FirstLanguage +
 Condition * cWM + Condition * cProf + Condition * ImmersionTimeinmonths +
 cFlankInhib
   Data: rdg_all
 Subset: PictureChosenifWrong < 3 & NativeSpeaker == "FALSE"
  AIC  BIC logLik deviance
 1486 1559 -729.9     1460
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 0.716453 0.84644
 SentenceID (Intercept) 0.055088 0.23471
Number of obs: 2023, groups: SubjectID, 133; SentenceID, 16

Fixed effects:
                                   Estimate Std. Error z value Pr(>|z|)
(Intercept)                       2.9189256  0.2867884  10.178  < 2e-16 ***
ConditionOVS                     -1.7974736  0.2842071  -6.325 2.54e-10 ***
FirstLanguageROMANIAN             0.0148414  0.3367317   0.044 0.964845
cWM                               0.1885437  0.0599942   3.143 0.001674 **
cProf                             0.0314365  0.0110822   2.837 0.004559 **
ImmersionTimeinmonths             0.0225919  0.0267450   0.845 0.398270
cFlankInhib                       0.0008082  0.0041791   0.193 0.846650
ConditionOVS:FirstLanguageROMANIAN  0.3110670  0.3369957   0.923 0.355976
ConditionOVS:cWM                 -0.2293123  0.0612361  -3.745 0.000181 ***
ConditionOVS:cProf                0.0441451  0.0115357   3.827 0.000130 ***
ConditionOVS:ImmersionTimeinmonths -0.0344274  0.0271938  -1.266 0.205512
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) CndOVS FLROMA cWM    cProf  ImmrsT cFlnkI COVS:F COVS:W COVS:P
ConditinOVS -0.735
FrLROMANIAN -0.726  0.545
cWM         -0.158  0.079  0.261
cProf        0.364 -0.309 -0.204 -0.005
ImmrsnTmnmn -0.459  0.368  0.215  0.165 -0.203
cFlankInhib -0.016 -0.013  0.020  0.064 -0.015 -0.023
COVS:FLROMA  0.538 -0.718 -0.746 -0.186  0.158 -0.154  0.013
CndtnOVS:WM  0.080 -0.183 -0.181 -0.758 -0.001 -0.140  0.000  0.253
CndtnOVS:cP -0.293  0.411  0.152 -0.015 -0.737  0.141 -0.022 -0.183 -0.041
CndtnOVS:IT  0.359 -0.471 -0.154 -0.137  0.145 -0.842  0.010  0.191  0.185 -0.214
```

APPENDIX P. Learner Distractor Items for Reading Task Model

```
Generalized linear mixed model fit by the Laplace approximation
Formula: distr_choices ~ (1 | SubjectID) + (1 | SentenceID) + Condition +      FirstLanguage
 + cWM + cProf + ImmersionTimeinmonths + cFlankInhib
   Data: rdg_all
 Subset: NativeSpeaker == "FALSE"
   AIC   BIC logLik deviance
 700.3 751.3 -341.2    682.3
Random effects:
 Groups     Name        Variance Std.Dev.
 SubjectID  (Intercept) 0.10006  0.31633
 SentenceID (Intercept) 2.64132  1.62522
Number of obs: 2125, groups: SubjectID, 133; SentenceID, 16

Fixed effects:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)          4.149658   0.492510   8.426   <2e-16 ***
ConditionOVS         -0.312278   0.217210  -1.438   0.1505
FirstLanguageROMANIAN  0.145007   0.240584   0.603   0.5467
cWM                  0.047434   0.041967   1.130   0.2584
cProf               0.020586   0.007846   2.624   0.0087 **
ImmersionTimeinmonths -0.001759   0.013590  -0.129   0.8970
cFlankInhib          0.004837   0.004502   1.074   0.2826
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
          (Intr) CndOVS FLROMA cWM    cProf  ImmrsT
ConditinOVS -0.247
FrLROMANIAN -0.288  0.001
cWM        -0.117  0.000  0.289
cProf       0.149 -0.003 -0.196 -0.097
ImmrsnTmnmn -0.193 -0.021  0.267  0.186 -0.331
                  cFlankInhib -0.012  0.011  0.068  0.102 -0.062  0.013
```

VITA

Aroline E. Seibert Hanson
arolinhanson@gmail.com

EDUCATION

PhD in Spanish and Language Science, Dual Title August 2012
The Pennsylvania State University, University Park, PA

MA in Spanish Linguistics May 2009
The Pennsylvania State University, University Park, PA

MA in Spanish August 2005
Middlebury College, Middlebury, VT

BA in Spanish May 2001
Tufts University, Medford, MA

RESEARCH INTERESTS

Second Language Acquisition and Processing
Cognitive Individual Differences
Teaching Methodologies

TEACHING EXPERIENCE

Teaching Assistant, Intensive Language Institute, Penn State	2011
Teaching Assistant, Introduction to Hispanic Linguistics, Penn State	2009-2010
Supervisor for Spanish 2, Penn State	2008-2009
Instructor and Tutor for GRE and TOEFL, Kaplan, Inc.	2007-2008
Spanish Teacher, Portsmouth Abbey School	2001-2007

SELECTED CONFERENCE PRESENTATIONS AND PUBLICATIONS

Seibert Hanson, A. (2012). *If my memory serves me correctly: L2 processing of Spanish clitics.* Paper presented at GURT, Georgetown University.

Seibert Hanson, A. (2011). *L2 word order processing: Proficiency matters.* Paper presented at Second Language Research Forum, Iowa State University.

Seibert Hanson, A. & Sagarra, N. (2011). *L2 comprehension and production of first nouns: Universal processing, but not universal understanding.* Paper presented at International Symposium on Bilingualism 8, University of Oslo, Norway.

Sagarra, N. & Seibert Hanson, A. (2011). Eyetracking methodology: A user's guide to linguistic research. *Studies in Hispanic and Lusophone Linguistics*, 4:2.

FELLOWSHIPS AND AWARDS

Graduate Exhibition, 3rd Place Poster, Arts and Humanities	2012
Graduate University Fellowship, Penn State	2007-2012
Teaching with Technology Certificate, Penn State	2008