Reevaluating Learning Without Awareness: An extension of Williams (2005)

A Thesis

Presented to

The Division of Philosophy, Religion, Psychology, and Linguistics

Reed College

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Arts

Aaron Carreras

May 2013

Approved for the Division

(Psychology)

Enriqueta Canseco-Gonzalez

Acknowledgments

I would like to thank everyone who has offered their support and encouragement throughout this process. To Mami and Papi: I couldn’t have made it this far without you. Thank you for your constant love and support. To Kika, Eric, and Ernie: You guys have been the best support system that I could ask for. I would never have been able to run my experiment back home without you guys. To Enriqueta: Your support has been absolutely crucial. It has been a pleasure to work with you over the past four years. Thank you for putting up with me throughout this process. To Loretita and la Güerita: I would never have been able to get my experiment on its feet without you. Thanks for helping me throughout the process, especially the planning and coding phases. Your love, laughter, friendship, encouragement, and endless positive energy helped me maintain my sanity during those rough moments. To my Chittick and Spanish House families: thank you for the wonderful communities we created together. The meals, dances, laughter, and friendship that we shared have meant the world to me.

Table of Contents

List of Tables

List of Figures

Abstract

In prior studies Williams (2005) showed that people are able to learn form-meaning connections without awareness. Extending his findings, we investigate whether, while learning a new artificial grammar, bilingual speakers are capable of extracting a 'hidden' linguistic feature that exist only in their inactive language at the time of testing. With that purpose, we included three groups of speakers: Spanish-English Heritage speakers, second language learners of Spanish, and a control group of English monolinguals. Most participants remained unaware of the implicit rule during the training and testing phases. However, second language learners of Spanish chose the correct determiner-noun combination at significantly above-chance levels, presumably relying on their knowledge of Spanish grammar. Plausible explanations for differences observed between groups and the implication of our results are discussed.

Chapter 1

Introduction

Throughout this chapter I will present a brief review of prior studies investigating how monolingual and bilingual individuals respond to instances of implicit learning and set the foundation for my experiment. First, I define the different groups of participants I examined in my study. I then focus on studies that compared how individuals from various language backgrounds behaved in situations where they were expected to learn certain linguistic features of a given language without their awareness. Lastly, I thoroughly discuss a specific study and its reiterations, which have provided contradicting evidence as to whether individuals could implicitly learn particular aspects of an artificial grammar. Based on these studies I present my own research, designed to expand on and possibly reconcile these contradictory findings.

Different Kinds of Bilinguals

Given the diverse array of bilinguals, it becomes very difficult for researchers to classify these individuals into few, overarching, and meaningful categories. Over the years, researchers in psycholinguistics have defined and classified bilinguals in a variety of different ways (De Groot, 2011). Bilinguals have been divided by the differential competence between their two languages, with "balanced" bilinguals being equally proficient in both languages and "unbalanced" bilinguals being more dominant in one language over the other. Bilinguals can also be classified according to the level of proficiency they reach in each language, such that they are considered proficient bilinguals if their competence reaches near-native levels; whereas, non-proficient bilinguals fall short of native-like levels in one or both of their languages. Researchers have also categorized bilinguals according to the age at which they begin to acquire their second language. Those that learn their second language in childhood are considered "early" bilinguals and those that acquire their second language after childhood are deemed "late" bilinguals. Clearly, one can classify bilinguals in numerous ways. For the purpose of this study, I will limit myself to two groups of bilinguals differing in the age at which they acquired their second language: heritage language speakers, who would resemble the "early" bilinguals mentioned above, and second language (L2) learners, who could be considered "late" bilinguals.

Heritage Language Speakers

Over the past few years, there has been a growing body of literature investigating the differences between heritage language learners and second language learners across various linguistic abilities. Heritage speakers are typically either, the children of immigrants born in a given host country, or immigrant children who arrived in the host country some time during childhood (Montrul, 2012). The host country's language, spoken by the greater speech community, would be considered the country's majority language; whereas, the minority language, spoken by immigrants and their children, would be deemed the individual's heritage language.

Heritage speakers are a special case of bilinguals in that they are exposed to each language in different contexts and often acquire different levels of proficiency in their two languages. The minority family language is often spoken in the home environment, and the majority language is spoken throughout the host country. Some heritage speakers are exposed to both of these situations from birth. Others grow up in a monolingual setting initially, speaking only the minority language in early childhood, and then later, they become bilingual around age five or six when they begin attending school, taught in the majority language.

While heritage speakers usually attain a strong command of the majority language, their proficiency and literacy level in the heritage language might suffer, as these speakers seldom have access to formal education in the minority language. As a result, the heritage language is often the weaker language for these individuals. These speakers are different from their monolingual, late-bilingual, and balanced-bilingual peers in various areas, including the age onset of bilingualism, their exposure to the heritage language, and their literacy skills in the heritage language (Bolger & Zapata, 2011). Additionally, the minority language is often stigmatized in the host country. All of these factors make it difficult for most heritage speakers to achieve native-like fluency in the minority language.

Research has found various areas in which heritage speakers show some linguistic deficiencies, including vocabulary, morphosyntax, word order, relative clauses, and conjunctions, among others (Montrul, 2010a). Spanish heritage speakers, for example, have been shown to produce gender errors in nouns with nontransparent word endings and have shown poor control of the subjunctive and conditional moods (Montrul, 2007; Silva-Corvalán, 1994; as cited by Montrul, 2010a). These heritage speakers also tend to make errors in verb choice (e.g., the use of 'estar' in contexts where 'ser' is appropriate as in 'ella está rubia' where it should be 'ella es rubia'), blur the distinction between the preterit and the imperfect tenses, and include optional and/or redundant pronouns (e.g., **Ella** estuvo en el tercer grado cuando **ella** vivió acá) (Silva-Corvalan 1986; Silvia Corvalan 1994; Zentella 1997; as cited by Suárez-Orozco & Paez, 2009).

Although heritage speakers may show some signs of language attrition, they may also display different degrees of proficiency in various aspects of the minority language. For example, research has found that heritage speakers behaved identically to advanced second language learners in regard to grammatical "subject pronoun expression, adjective-noun agreement…expression of aspect and hypotheticality, mood variability, and the occurrence of lexical Angliscisms" (Lipski, 1993 as cited by Roca & Colombi, 2003, p. 31) (e.g. suéter, líder, túnel etc.). One study that compared Spanish heritage speakers and Spanish L2 learners' morphosyntactic knowledge, found no differences in their knowledge of object clitic placement in a written grammaticality judgment task. However, heritage speakers outperformed the L2 learners in their acceptance of sentences with alternative word order (e.g. Olga lo puede comprar mañana, Olga puede comprarlo mañana), while the L2 learners were better at rejecting ungrammatical sentences without differential object marking (DOM) in Spanish (e.g. Les enseña una lección la profesora. [no DOM] Juan llama por teléfono **a** sus padres. [DOM]) (Montrul, 2010b).

In sum, past research has shown that heritage speakers' proficiency in the minority language does not totally resemble that of monolingual speakers, nor the proficiency of second language learners, which raises the question of where these individuals fall on the spectrum of bilingualism. Considering that heritage speakers and second language learners are treated similarly in terms of academic needs in most Spanish language classrooms, in spite of the differences in certain linguistic abilities mentioned above, research in this area may help introduce programs that would be better able to accommodate heritage speakers' attempts to learn their heritage language in an academic environment. For example, one study found that Spanish heritage speakers did not benefit as much as second language learners from different types of instruction of the imperfect subjunctive tense (K. Potowski, Jegerski, & Morgan-Short, 2009). If these different populations benefit from different kinds of instruction, then it is important to carry out further research to determine what kinds of instructional techniques could accommodate heritage speakers' needs. If, for example, Spanish heritage speakers would benefit more from a language approach that is not typically offered in a second-language classroom, it would be inefficient for these students to be placed in the same class as their L2 learner peers. Many heritage speakers that do not consider themselves "native" or "bilingual" Spanish speakers take Spanish courses designed for non-native speakers, where they lose valuable academic time that could have been used to help these students develop the skills appropriate for their needs (Roca & Colombi, 2003, p. 30).

Despite the differences between heritage speakers and other types of speakers, they do share some linguistic characteristics not only with monolinguals of the minority language, as both groups experience an early onset of exposure to the language, but also with second language learners of that language, given their varying degrees of proficiency in the language (Bolger & Zapata, 2011). For example, one study found that L2 learners and heritage speakers *both* show grammatical deficits with regard to gender agreement (Montrul, Foote, & Perpiñán, 2008). These researchers stated that further investigation must continue investigating knowledge of gender agreement and other grammatical properties by utilizing tasks that minimize participants' direct use of *explicit* knowledge of gender. As research has shown that heritage speakers outperform L2 learners on tasks that rely on intuitive knowledge (Davidson, de la Fuente, & Montrul, 2011; Montrul & Perpiñán, 2011), it is crucial to follow the advice of Montrul et al. (2008) and investigate how these speakers perform on implicit learning tasks. In such tasks, individuals are expected to learn and apply concepts that are not explicitly taught.

Implicit Learning and Artificial Grammars

Implicit thought processes are often defined as fairly passive and automatic activities that occur outside the realm of individuals' conscious awareness; whereas, explicit processes are more conscious and attention demanding (Stadler & Frensch, 1998). Two criteria for implicit learning that are often highlighted in the literature are intentionality and awareness (Rebuschat & Williams, 2012). When one engages in implicit learning it occurs passively and without the participant's intention to learn. Studies involving the learning of artificial grammars often have researchers present participants with a set of rule-governed stimuli during a training phase and then test their new knowledge in a subsequent testing phase. Participants are typically not told about the nature or the existence of a rule that governs certain aspects of the language, but they are given an opportunity to apply those rules in the testing phase.

One early experiment on implicit learning by Reber (1967) had participants memorize a list of letter strings that followed a specific grammar rule. Participants were not given any information regarding the specific grammar rule or logic that the strings followed. They were simply asked to learn the strings in preparation for a memory test. After memorizing the strings, the participants were told about the existence of a grammar rule that governed the strings of letters. The dependent variable was the percentage of old and new strings that participants identified as following the grammatical rule that they were unaware of. Above average performance on this task was taken as an indicator of implicit learning ability (Reber, 1989 as cited by Stadler & Frensch, 1998).

Participants that perform significantly better than chance appear to have all the relevant information to complete the task; however, they are often unable to verbalize the knowledge that they gained. Of course, researchers in this area must be careful in selecting a task that gauges whether or not the participants actually learn the material implicitly, rather than explicitly. Without showing that participants have no explicit knowledge of the hidden rule, one could always claim that a particular experiment's results are due to the effects of explicit learning that are not considered by the researcher. For example, some experiments consider implicit language learning to have taken place when the participants perform statistically better than chance on testing items, while being unable to communicate how or why they were able to perform so well. It may be the case that participants did learn and apply an implicit rule or they may simply have trouble verbalizing their reasoning. Perhaps having more time to reflect on their decision-making process would allow them to eventually identify the rule correctly.

Implicit Learning in Heritage Speakers

A study by Davidson et al. (2011) compared the performance of Spanish native speakers, Spanish heritage speakers, and intermediate to advanced L2 learners of Spanish using tasks that tested both explicit and implicit knowledge of Spanish. Participants were given a series of tasks, including a Spanish word repetition task, in which participants listened to a series of three-word phrases (determiner, adjective, noun) and repeated the last word of each phrase as quickly and accurately as possible. Davidson et al. (2011) included the word repetition task to gauge a more implicit, automatic knowledge of the language. The dependent variable was the amount of time it took participants to repeat the final word in the three-word phrase. The notion was that participants would read sentences faster if they followed the phrase structure that subjects had been presented during training. The only significant finding was a difference between the Spanish L2 learners and the Spanish heritage speakers found in the reaction time data on the word repetition task. In fact, the performance of the heritage speakers resembled that of native Spanish monolinguals. The results from Davidson et al. (2011), along with those of other studies, illustrate that although Spanish heritage speakers often lack metalinguistic knowledge shared by L2 learners and monolinguals, they appear to resemble native Spanish monolinguals in tasks that rely on implicit knowledge (Montrul & Perpiñán, 2011).

Implicit Learning in Monolinguals and Second Language Learners

A recent body of research has also shown that it is possible for participants to implicitly learn certain aspects of syntax in artificial languages. A study by Cleary and Langley (2007) trained college students with incoherent word sequences (e.g. MECHANICAL CONSUMERS SUBMIT COLDER SONGS) that (unbeknown to participants) followed a specific syntactic pattern (e.g. ENERGETIC TREES DEMAND SILLY FRAMEWORKS follows the same phrase structure as the aforementioned example). When individuals were presented with a recognition task (i.e. Rate the likelihood that each test string was previously studied on a scale of zero [definitely not studied] to 10 [definitely studied]), they were significantly more likely to rate new, non-studied word strings as old, if they followed the same syntactic pattern, suggesting that they had implicitly learned it.

A study by Francis, Schmidt, Carr, and Clegg (2009) illustrated that English native speakers were also able to implicitly learn basic word order patterns in an artificial grammar. Participants were presented with noun-verb-noun (N-V-N), noun-noun-verb (N-N-V) and verb-noun-noun (V-N-N) strings, the latter two constructions not occurring English. Participants were randomly assigned to receive one set of strings, either N-V-N or V-N-N strings, more often than the other. The set of strings that appeared most often was called the "More Practiced Rule". Participants were instructed to simply read each three-word string aloud. Participants' reading time was used as an indirect measure of learning. The second part of the study presented participants with old and novel N-V-N, N-N-V and V-N-N word strings. During this phase, researchers found that participants' reading speed decreased for strings that followed the "More Practiced Rule", even if the specific word strings had not been presented before. Participants appeared able to generalize information from one task to another with mere exposure to a construction that does not occur in the participants' native language.

Learning Without Awareness

Research has demonstrated conflicting results regarding people's ability to acquire grammatical knowledge without conscious awareness. In the studies I will present throughout this section, participants were explicitly given constructed rules governing specific aspects of an artificial language. Unbeknownst to these participants, researchers embedded an additional hidden rule in the stimuli that could help them decide between two possible stimulus choices in situations where the explicit rule did not provide an answer. Participants were taught four novel words that seemingly conveyed the same grammatical information and were asked to choose between them without being explicitly told when or how to choose one over the other. Nevertheless, these participants appeared to learn this hidden rule subconsciously and applied it correctly. Importantly, when asked to explain the reasoning behind their decisions, most participants were unable to explain the implicit rule they used. In short, these participants appeared to have learned a grammatical rule while being completely oblivious to it. In what follows, I briefly describe three of these studies that serve as a basis for our research.

A study by Williams (2005) examined the role of participants' prior linguistic knowledge on their ability to learn implicitly certain aspects of an artificial grammar. In the first experiment, 41 participants from various language backgrounds were given sentences and explicitly taught that the stimulus' article selection relied on a distinction between near and far (a choice between *gi* and *ro* for near and *ul* and *ne* for far). A vocabulary pre-training phase used flash cards to explicitly teach participants the distance information to which each determiner corresponded. In the training phase, participants were presented with sentences and were asked to decide whether the determiners meant near or far in each sentence (e.g., "I was terrified when I turned around and saw *gi* lion right behind me", in this case the lion is near the person). They were then asked to repeat the sentence orally and to form a mental image of the situation described. After the training, the testing phase began. Participants were presented with a sentence fragment providing contextual information (e.g., The woman was stung by __ ___), and participants were asked to choose the appropriate determiner-noun combination that would complete the sentence (e.g., *gi* bee / *ul* bee) according to the distance information.

Unbeknownst to participants, the animacy of the noun covertly played an important role in governing the determiner selection. The determiner *gi*, for example, indicated that the noun was not only close to the subject of the sentence, but also that it was a living entity (i.e., animate). Williams (2005) used four novel articles: *gi* referred to nouns that were near and animate, *ro* to nouns that were near and inanimate, *ul* to nouns that were far and animate, and *ne* to nouns that were far and inanimate. Importantly, participants were not informed about the role of animacy, but this feature was embedded into each sentence. After the testing phase, the participants were asked to report the criteria they had used to make their decisions. If participants mentioned animacy, they were asked *when* they had discovered the relevance of animacy. If participants failed to mention animacy, they completed a second testing phase where they were explicitly asked to search for any other rules governing the use of the novel determiners besides distance. The second testing phase was included to give individuals a second opportunity to identify the (implicit) rule. At the end of the second testing phase, participants were asked to provide a guess of the rule. Crucially, it was found that some participants who were unable to articulate the rule governing determiner selection, were still able to choose the correct determiner depending on each noun's animacy at a statistically significant level above chance. Only eight of the 41participants expressed awareness of the relevance of animacy; six of them became aware during the training phase and two during the testing phase.

Since five of the six participants who became aware during the training phase were native speakers of English, Williams ran a second experiment in order to address the possibility of different processing demands for native and non-native speakers of English. Because the experiment was conducted mostly in English (plus the artificial grammatical rules), it was possible that the native English speakers were given an unfair advantage over participants who had learned English as a second language. For example, the native English speakers could have had more experience working with determiners compared to the other participants. To address this possible confounding effect, this second experiment included only non-native speakers of English from various language backgrounds, including Albanian, Arabic, Bosnian, Mandarin Chinese, Croatian, German, Korean, Macedonian, Serbian, Singhalese, Spanish, Taiwanese, and Ukrainian.

In this iteration, participants were again unable to explicitly identify the rule governing the determiner selection, and were also able to choose the determiner appropriate for each noun's animacy at levels significantly above chance. Importantly, this was the case even when participants were faced with determiner-noun combinations that did not appear during training. Williams (2005) took these findings to suggest that the participants had implicitly learned and applied the hidden animacy rule. Seven of the 24 participants expressed awareness of the relevance of animacy; two of them became aware during the training phase and the other five during the first test phase. The remaining seventeen unaware participants showed slightly higher scores than unaware participants from the first experiment. These results suggest again that connections between form and meaning can be learned implicitly.

Interestingly, there was also a correlation between test performance on the novel determiner-noun combinations and knowledge of languages that encode grammatical gender. Williams stated that the only thing that the participants' languages shared with the language system used in the experiment "was a formal structure in which nouns belong to different classes and determiner use is controlled in part by agreement processes with respect to those classes" (2005: 297). In short, this language was similar to other languages in that some nouns require certain determiners in specific contexts. However, none of the participants' spoken languages distinguished word classes based on animacy. Another concern that was raised was the possibility that participants with knowledge of a language that encoded grammatical gender may have showed a greater sensitivity to noun-determiner classes with sophisticated agreement processes. For example, languages like Spanish place a restriction on which articles a certain noun can take, as the article must agree with the grammatical gender of the particular noun. Williams took this to suggest that implicit learning of form and meaning may also depend on the individuals' prior linguistic knowledge.

Hama and Leow (2010) failed to replicate Williams' (2005) experiment after making a few methodological changes. First, they excluded participants with backgrounds in fields related to language or linguistics, as individuals in these fields may have been more sensitive to identifying the hidden grammatical rule. They used a training task that was similar to that of Williams's (2005) study; however, participants were asked to "think aloud" throughout the training and testing procedures in order for the experimenters to gain a better idea of how participants processed the input during these two phases. The participants' reported thoughts were coded in order to place participants in different groups with regard to their level awareness to the role of animacy. Participants were categorized as "noticing" if something related to animacy was mentioned, "understanding" if they were able to correctly identify the role of animacy, and "no report" if they did not fall in either of the previous categories. Unlike Williams (2005), who offered participants a choice between two of the determiner-noun combinations, these participants had to choose between all four possible determiner-noun combinations during the testing phase. This was done in order to obtain a deeper analysis of participants' awareness of the targeted rule and make it more difficult for participants to choose the correct choice by chance. Furthermore, in order to ensure that the methodology of the two phases remained consistent, the sentences in both the testing and training phases were presented orally. Hama and Leow (2010) also added a production assessment task during the testing phase, in which participants listened to a sentence where the object was missing a determiner and a short beep indicated the position of the article. They were asked to orally provide the missing determiner, making sure to think aloud during the task. This extra task was added to help experimenters gain a better measure of when and how the participants would discover the role of animacy. Hama and Leow (2010) found that participants in the "no report" group failed to consider the role of animacy either when asked to choose between the articles in the testing phase, or when asked to verbally produce the missing determiner. In fact, participants in this study showed difficulty in merely learning the determiner-noun combinations that occurred during the training phase.

A similar study by Faretta-Stutenberg and Morgan-Short (2011) also failed to find evidence of implicit learning despite nearly mirroring Williams' (2005) study. This study had some key differences concerning the participant pool and how the data were analyzed. First off, the experiment only included thirty native English speakers who were not studying language-related disciplines. Of the native English speakers, fifteen of them reported being native in another language besides English. The majority of these participants' second languages encoded grammatical gender. Secondly, the researchers decided to split participants into three awareness groups in order to further analyze how mindful the participants were of the implicit rule (*No Report, Noticing,* and *Understanding*). Despite these minor alterations, the study found no evidence of implicit learning in either the monolingual or the bilingual participants. The original Williams (2005) study found a relationship between the number of gendered languages one speaks (at the intermediate level or above) and his/her score on the generalization items for participants that were unaware of the animacy rule. Similarly, those who studied language-related disciplines such as linguistics, outperformed those who were not linguistic students. Faretta-Stutenberg and Morgan-Short (2011) took these observations to suggest that speaking, studying, and learning a variety of languages can have a large effect on one's ability to implicitly learn certain aspects of a new language.

The Present Study

The aforementioned studies have shown that monolinguals and second language learners can show signs of implicit learning after being exposed to patterns of an artificial grammar, even if the pattern does not follow that of the speakers' native language. It would be interesting to investigate whether second language learners would be able to use information gained from their second language in an implicit learning task of an artificial grammar. Specifically related to the study reported below, the question is whether Spanish heritage speakers' intuitive knowledge regarding certain aspects of their minority language (Spanish) would carry over to contexts where they are tested in the majority language. Would native English speakers that study Spanish as a second language be able to implicitly use grammatical gender, a feature present in Spanish but absent in English, to complete a task in an artificial grammar learning situation? Where would Spanish heritage learners fall on this spectrum? Would their knowledge of Spanish allow them to behave more like second language learners of Spanish? Would the fact that Spanish is the minority language for many heritage speakers have them perform more like English monolinguals?

As Faretta-Stutenberg and Morgan-Short (2011) suggested, it is crucial to investigate the role that prior linguistic knowledge has on implicit learning. To examine this further, I conducted a study attempting to replicate Williams' (2005) study. Furthermore, I introduced a different set of participants in this experiment: a group of second language learners of Spanish, a group consisting of English monolinguals that learned Spanish as a second language in an academic setting during adolescence, and a group of Spanish heritage language speakers, consisting of individuals that acquired Spanish at home before learning English.

Like the aforementioned studies, I used distance (near vs. far) as the grammatical feature that was explicitly taught. Participants were told that the selection of new determiners would depend on whether the entities mentioned in the sentence were near or far from the subject of the sentence. (e.g. in this artificial grammar, the English sentence

"I was terrified when I turned around and saw the lion right behind me" would require a determiner different from "The hunters saw the lion with their binoculars"). However, instead of using animacy as the grammatical rule that participants were to learn implicitly, I used the corresponding grammatical gender of each noun when translated into Spanish. I made this decision based on previous research showing that people's ideas about the "gender of objects" are strongly influenced by the grammatical gender assigned to those objects in their native language, even if the testing is conducted in a language in which nouns do not contain grammatical gender (Boroditsky, Schmidt, & Phillips, 2003). Given that Spanish heritage speakers would have had experience interacting with nouns that display grammatical gender in the home language, I expected these participants to be more likely to adopt and apply the implicit gender rule than second language learners of Spanish. Monolingual English speakers, who lack grammatical gender information, were expected to be unable to learn the rule. Second language learners of Spanish, who have some experience with a language that encodes grammatical gender, were expected to extract and apply the implicit rule more accurately than the monolinguals, but not as consistently as the Spanish heritage speakers. This expected result was based on the idea that because Spanish heritage speakers had spent more time speaking Spanish (since childhood in most cases), they would have been able to perform better than the second language learners, who began acquiring it during adolescence.

The artificial language included two determiners for nouns that are close to the subject of the sentence (e.g. "I enjoyed it when *gi* rabbit ate from my hand" versus "I saw *ro* boxes right beside me) and two for those that are far (e.g. He saw *ul* rabbit in the distance, I noticed *ne* photograph across the room). Unbeknownst to participants, for each distance pair (e.g. *gi* and *ro*), the selection of the correct determiner in a given context depended on the gender of the noun in Spanish (i.e. in the prior example, *rabbit* is marked with masculine gender, while *boxes* is marked with feminine gender, therefore requiring the determiners *gi* and *ro* respectively in this artificial grammar). The noun phrases and sentences used were adapted from Williams (2005), with some modifications. For example, in order to facilitate participants implicitly learning the role of gender, half of the animate nouns used during the training phase were animate nouns that contained biological gender information (e.g. butler, king, queen, princess). To repeat, I expected to find that the Spanish heritage speakers, the group with the greater amount of experience speaking a language that encoded grammatical gender, would be more likely to extract and use the implicit grammatical features of gender embedded in the stimuli. Second language learners were expected to perform slightly less accurately than the Spanish heritage speakers.

Chapter 2

Methods

Participants

Seventy-one undergraduates from Reed College and college students from the Chicago area, between the ages of 18 and 26 (M = 20.7, SD = 1.6) were selected for this study. The first group of 25 was composed of Spanish heritage language (HL) speakers, defined as Spanish-speaking children of Mexican immigrants, born in the United States and exposed early to Spanish in the home. The next group of 23 was formed by the Spanish L2 group, which included native English speakers that learned Spanish as a second language in an academic setting during adolescence. Monolingual English speakers, formed the final group of 23. Of the monolingual participants, six people were excluded from the final analyses: two participants had studied abroad for more than four months in Spanish-speaking countries and the other four had indicated making responses solely based on where the buttons were positioned on the response pad. The data of the remaining 17 monolinguals were used for analyses.

In order to be considered for the study, participants sent an email message with their phone number so that they could be contacted over the phone. Monolingual English speakers were scheduled in English. The Spanish L2 and Spanish HL speakers were scheduled in Spanish to verify their fluency.

Procedure

This experiment had three phases. In the *pre-training phase*, participants were presented with the four artificial determiners to be used throughout the study. The second phase, the *training phase*, presented participants with sentences containing the new determiners, and taught them an explicit rule to choose between them based on distance information. However, the determiners were also used in a way that could allow participants to extract the implicit rule (i.e. Spanish gender) governing their use. The final phase was the *testing phase*, in which participants were asked to apply the implicit rule by completing sentence fragments with the appropriate determiner-noun combination (e.g. The fire brigade had to rescue Det Noun from the top floor of the castle). These three phases are described in detail below (see Figure 10 for a summary of the experimental design).

The materials and procedures mirrored those of Williams (2005) as closely as possible. Unlike the previous studies where the choice of determiner relied on distance (as the explicit rule) and animacy (as the hidden rule), this study used distance as the explicit rule and grammatical gender as the hidden rule. Using gender instead of animacy gave the HL and L2 groups an opportunity to subconsciously learn and use a linguistic feature (gender) that they had been familiar with through their experience with the Spanish language. Because the English language has a much less prevalent gender system than Spanish, the monolingual English participants were expected to have less experience with grammatical gender. As a result, the HL and 2L groups were expected to have an easier time learning (implicitly) the role of gender in the artificial grammar, as compared to English monolinguals. Instead of including a separate yes/no recognition trial to gauge whether participants had been paying attention throughout the experiment, as in Faretta-Stutenberg and Morgan-Short (2011), I presented memory questions during the training phase but, rather than asking participants whether they had heard a particular sentence, I asked participants to identify which determiner appeared with a particular noun in a given sentence (see below). Replacing the recognition trial with memory questions shortened the length of the study (by removing an extraneous trial) and served to subtly help participants learn the role of a feature (gender) that does not exist in English.

The same four novel determiners were used as in Williams (2005): *gi* (near, masculine), *ro* (near, feminine), *ul* (far, masculine), and *ne* (far, feminine). Two examples of the sentences participants heard during the training are shown below.

1a. While I was digging in the garden *gi* rabbit came and sat next to me.

1b. The children threw sticks at *ne* squirrel in the tree.

Pre-training Phase

After providing written consent, participants studied flashcards of the four novel determiners used throughout the experiment (*gi*, *ro*, *ul*, *ne*). Each determiner was written on one side of the flashcards in different colored ink in order to help participants distinguish between the articles. The other side revealed the distance information (whether the determiner referred to a near or far object) of each determiner in English (e.g. near was written on the reverse side of *gi*).

After participants were confident that they had memorized the determiners and their corresponding meaning, they were asked to complete two tasks on the computer in order to demonstrate that they felt comfortable working with the new articles. For both tasks, incorrect responses were met with feedback informing participants of their incorrect response.

The first task displayed distance information (e.g. far) on the computer screen and participants were asked to choose the article corresponding to each distance by pushing the appropriate button (*gi*, *ro*, *ul*, *ne*) on a response pad (Figure 9). For example, if a participant was shown the word, “far”, this meant that the subject of the sentence was far away from the object in question and they should choose one of the two articles associated with far.

The second task presented participants with one of the novel determiners (e.g. *gi*) and they were asked to indicate whether it was associated with “near” or “far” by pressing the appropriate button on the response pad. After 12 consecutive correct answers, participants moved on to the training phase.

Training Phase

In the training phase, participants listened to 24 sentences with each sentence containing one of the four novel determiners. After each sentence, participants were first asked to repeat the new determiner-noun pair in that sentence. If, for example, participants heard the sentence “I pat *gi* dog”, they were expected to repeat “*gi* dog”. Second, they were instructed to form a mental image of the situation described by the sentence. Finally, they indicated whether the subject was near or far from the object of the sentence by pushing the appropriate button on the response pad. Participants were told that this process was important to gain a general idea of each sentence and to prepare for memory questions that would appear throughout the training phase.

After listening to a few training sentences, participants were presented with a sentence they had previously heard but now they saw it on the screen, with blanks where the determiner and noun had appeared (e.g. While lying in the grass, the girl was greeted by__ ___). Participants were given a choice between either the two *near* (one feminine and one masculine) or two *far* determiners (one feminine and one masculine). For example, for the aforementioned sentence, participants would be asked to choose between *gi* queen (near, masculine) and *ro* queen (near, feminine). Participants were instructed to choose the determiner-noun combination they remembered hearing in the sentence provided.

Because participants were not given any information regarding the implicit rule, they were expected to pay very close attention to the determiner-noun combinations presented during the training trial in order to answer correctly the memory questions. To help participants determine whether or not they were successfully memorizing the determiner-noun associations, they were provided with feedback whenever they answered a memory question incorrectly. The amount of attention placed on remembering which determiners went with which nouns also encouraged participants to begin considering (even if implicitly) whether there was a rule that governed the use of the different determiners. It was important to have participants consider the differences between the two near (*gi* and *ro*) and the two far (*ul* and *ne*) determiners during the training phase, as it could have been very challenging for individuals to even consider the role of a grammatical feature that does not exist in English, the language in which the experiment was conducted.

The training phase consisted of twelve sets of twelve sentences (six near and six far article-noun combinations) in an alternating order between Set 1 and Set 2. Each block was independently randomized for each participant. One memory question was presented after each block, so that each participant was asked twelve memory questions. There was also a break of at least 30 seconds between blocks. Because of the distribution of sentences across the different sets, the same noun never occurred in consecutive trials, making it more difficult for participants to make explicit comparisons between nouns. Throughout the training phase, participants were not given any information regarding the existence of a rule governing the use of one near or far determiner over the other.

Testing Phase

In the testing phase, participants were presented with 26 sentences that followed the same format as the memory questions (e.g. The fire brigade had to rescue__ ___from the top of the tree). The participants were then asked to complete the sentence by choosing which of the two determiner-noun choices (e.g. *ul* prince or *ne* prince) seemed more familiar, better, or more appropriate based on the sentences they had heard before. As in Williams (2005), the testing phase included sentences that resembled training items. For example, if the noun 'nun' was close to the subject in the training phase sentences (e.g. The ball was kicked by *ro* nun during the church picnic) the noun was also near the subject for the testing phase sentences (e.g. The doctor began to examine *ro* nun in the infirmary). More importantly, this phase also presented nouns in opposite distance relationships to the subject of the sentence that participants saw in the training phase. These items will be referred to as the *generalization* items. For example, if the noun 'princess' had been presented as being near from the subject (e.g. *ro* princess) during the training phase, it was now presented as far from the subject in the testing phase. Participants would then have to choose between two determiners that they had never seen paired with the noun (e.g. *ne* princess or *ul* princess). Crucially, the correct answers for these *generalization* items depended on the gender of the nouns in Spanish. At the end of this phase, participants were asked what criteria they had used to make their choices. If participants mentioned that grammatical gender was the determining factor, they were asked at what point they became aware of its relevance. These participants were classified as *aware* for a later analysis and they finished the study at this point.

If gender was not mentioned, participants were told that the two near and the two far determiners were not interchangeable and that there was a rule that governed the use of *gi* (near, masculine) vs. *ro* (near, feminine) and *ul* (far, masculine) vs. *ne* (far, feminine). The participants were then instructed to look for that rule while they underwent a second testing phase trial.

The second testing phase used the same sentences as the first testing phase. The second testing phase began with the same ten sentences that participants saw in training and ended with the sixteen *generalization* sentences. Unlike the previous testing phase, when participants selected the wrong determiner, they were told that they had made an incorrect selection. After the twenty-six items were presented, the participants were given a second opportunity to identify the rule that governed the use of the novel determiners.

Participants that mentioned the role of gender in the language were asked at what point they had become aware of the relevance of gender and were classified as *aware*. If gender was not mentioned at all throughout the study, participants were asked if they had ever considered that gender played a role. Participants unable to notice or state a rule were categorized as *unaware*. The groups, *aware* and *unaware*, were used to statistically analyze the awareness of the participants. Awareness, lack of awareness, and when the participants became aware (after the first or second testing phase) were variables used in the analyses of the participants.

Language Questionnaires

Before being debriefed, all of the participants completed the LEAP-Q (Marian, Blumenfeld, & Kaushanskaya, 2007), indicating their age, years of education, native language, and experience with other languages. The bilingual participants, those in the 2L and the HL groups, were also asked information about the instructional context of English and Spanish, the contexts of English and Spanish use, and a self-assessment of proficiency in both English and Spanish (beginner, intermediate, upper intermediate, or advanced). Additionally, the Spanish HL speakers completed the Spanish placement test C (Ki. Potowski, Parada, & Morgan-Short, 2012). This test measured participants' comprehension of very informal words and expressions that students would not typically encounter in a Spanish classroom setting. As a result, this exam was useful in determining which participants learned Spanish outside of a formal, classroom context and could be considered Spanish heritage speakers. These questionnaires were used as a subjective (LEAP-Q) and an objective (Spanish placement test C) measure of participants' language experiences in English and in Spanish. Participants did not complete the language questionnaires until after the experiment was completed in order to prevent them from predicting that the artificial grammar relied on a grammatical feature found in Spanish.

Chapter 3

Results

Of the 71 participants, 18 (6 HL, 4 Monolinguals, and 8 2L) reported becoming aware of the relevance of gender at the end of the first testing phase. Of the 18 aware participants, 17 indicated that they began suspecting that gender had played a role during the training phase; however, given that nouns in English do not have grammatical gender, they often second-guessed themselves. One participant reported becoming aware during the first testing phase. Of the 156 items on the training trial, all of the participants scored between 144 and 155 (M = 150.79, SD = 2.74), which suggests that they were answering the training items as accurately as possible.

The number of correct determiner choices ("correctness" determined by the 'hidden' Spanish gender rule) was submitted to a series of statistical tests as indicated below. The dependent variable was participants' accuracy on the testing phase items. The independent variables were awareness, item type, and language background. The variable "awareness" indicates whether participants reported being aware of the role of Spanish gender in the selection of a particular determiner (aware) or whether they reported not knowing the rule (unaware). The variable "item type" refers to the kind of item presented during the testing phase. These were either trained items, which were determiner-noun combinations that participants had been exposed to during the training phase, or generalization items, which were determiner-noun combinations that participants had not seen before. Dividing the item type into these two types is important, as participants could potentially perform above chance levels simply by memorizing the determiner-noun combination of the trained items.

The variable "language background" indicates the type of speaker, with HL=Heritage Language speakers, 2L= 2nd Language learners of Spanish, and Monolingual=English monolinguals. A group of independent means t-tests revealed that HL participants (N = 22, M = 1, SD = .98, range = 0-3) acquired Spanish at a significantly younger age than 2L individuals (N = 22, M = 10.86, SD = 4.66, range= 4-

19) t (44) = 9.71, p <.01. The data also suggest that HL participants (N = 22, M = 3.8, SD = 2.70, range = 0-10) also became fluent in Spanish at a significantly earlier age than the 2L individuals (N = 22, M = 19.12, SD = 1.50, range = 16-21) t (41) = 9.12, p < .05. Spanish HL participants (M = 8.86, SD = 1.03, range = 6-10) self-rated as significantly more proficient in understanding Spanish than the 2L participants on a scale from 1-10 (M = 7.00, SD = 1.74, range = 3-9) t (44) = 4.30, p < .05.

Although the following analysis will include data from both the first and second testing phases, it is important to note that the first testing phase is the crucial part of the experiment for observing participants' implicit learning abilities. During the first testing phase, participants were not explicitly informed about the presence of a rule governing the determiner selection. The unaware participants needed to rely on what they had learned implicitly in order to choose the correct determiner at statistically above chance levels. In the second testing phase, however, participants were told about the presence of a rule, were encouraged to find it, and were given feedback to aid them in the search. The second testing phase thus introduces other factors, which may account for the observed results.

First Testing Phase

An initial mixed design ANOVA measuring the effects of awareness, item type, and language background on participants' overall performance revealed main effects of awareness F (1, 59) = 22.66, p < .01 and item type F (1, 59) = 5.26, p < .05 (whether the items were previously trained or generalization items) during the first testing phase. There was also an interaction between awareness and item type F (1, 59) = 7.10, p < .01. There was no main effect of language background F (2, 59) = 1.46, *ns*. These results suggest that aware individuals performed better than the unaware individuals, and that all of the participants performed better on trained items than on the generalization items.

Given that participants' overall performance depended on awareness and item type, these factors will be analyzed separately. Awareness will be divided into aware and unaware categories. Although the mixed design ANOVA did not reveal an effect for language background, I will still analyze the different language backgrounds separately in the event that a statistically significant effect was masked due to a lack of power. Lastly, item type will be split into training items and generalization items.

The training and generalization items will also be further split into 4 groups: training items with biological gender (e.g. king, queen), training items without biological gender (e.g. squirrel, cushion), generalization items with biological gender, and generalization items without biological gender (see Figure 10 for a summary of the breakdown). Splitting items into these additional categories (those with and those without biological gender) will illustrate whether participants were using the biological gender of the nouns to perform better than chance.

I will first report participants' performance on the total testing items (which included both trained and generalization items) and the trained items very briefly. Participants may have scored significantly better than chance on these items by relying on memory for what they had heard during the training trial. The generalization items, however, are essential for testing my hypothesis and I will recount the participants' performance on these items in more detail.

Total Testing Items

When divided by awareness categories and language background, the aware and unaware participants across all of the language backgrounds (2L, HL, and Monolinguals) scored statistically above chance on the total testing items (Figure 1). It is unclear what drove the effect, however, as the testing items included old determiner-noun combinations (trained items), new determiner-noun combinations (generalization items), and nouns with biological gender. For this reason, participants' performance will be analyzed across the different item types in order to determine whether participants were able to extrapolate and apply the implicit rule when they were unable to rely on memory or biological gender encoded in English nouns.

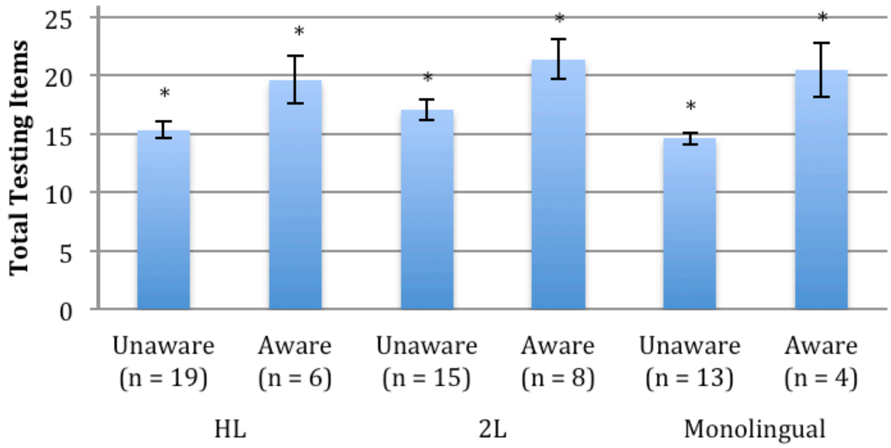

Figure 1: Mean scores for total (trained and generalization items) number of testing items (*chance* = 13).

*Significantly above chance for Figure 1 and following figures.

Trained Items

All of the participants, regardless of awareness and language background, performed better than chance on the total trained items and on the trained items without biological gender (See Table 1 and Table 2 for descriptive statistics and t-test results). For the trained items with biological gender, only the unaware HL and 2L and aware 2L and monolingual participants performed better than chance. One must take these results with caution, however, as there were only three training items with biological gender included in the stimuli. Had there been more of these items, the remaining groups may have shown a performance statistically above chance.

Generalization Items

A series of single sample t-tests showed that for the first testing phase, only the aware 2L participants ($N = 8$) scored better than chance *(chance* = 8) on the total generalization items ($M = 12.13$, $SD = 3.44$, $t\ (7) = 3.40$, $p < .05$). In contrast, none of the unaware participants scored better than chance *(chance* = 8) on the total generalization items (Figure 2).

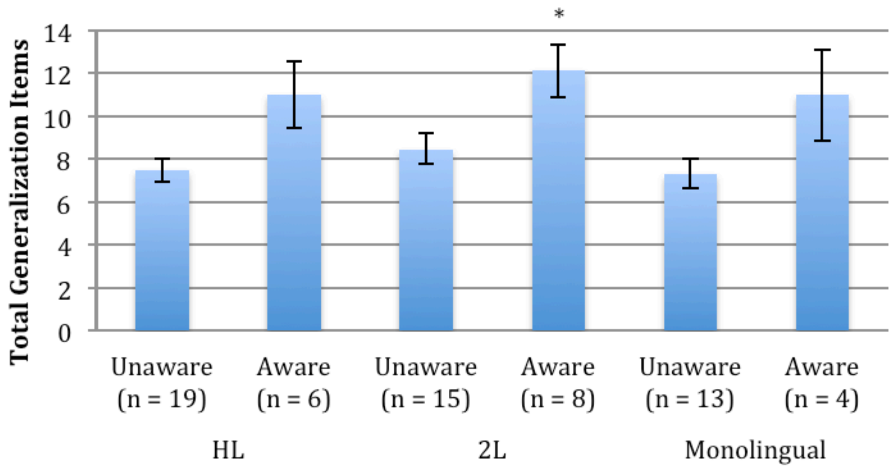

Figure 2: Mean scores for the total number of generalization items *(chance* = 8).

On generalization items with biological gender, both aware Spanish HL participants ($N = 6$, $M = 2.83$, $SD = 0.98$, $t\ (5) = 2.08$, $p < .05$) and aware 2L participant groups ($N = 8$, $M = 3.00$, $SD = 1.20$, $t\ (7) = 2.37$, $p < .05$) scored better than chance *(chance* = 2) on these items. The unaware participants across the different language backgrounds, however, did not perform better than chance on these items (Figure 3). It is plausible that the aware HL and 2L participants had discovered the rule and were correctly applying it, however it is difficult to draw conclusions based on the small number of total generalization items with biological gender (*n* = 4).

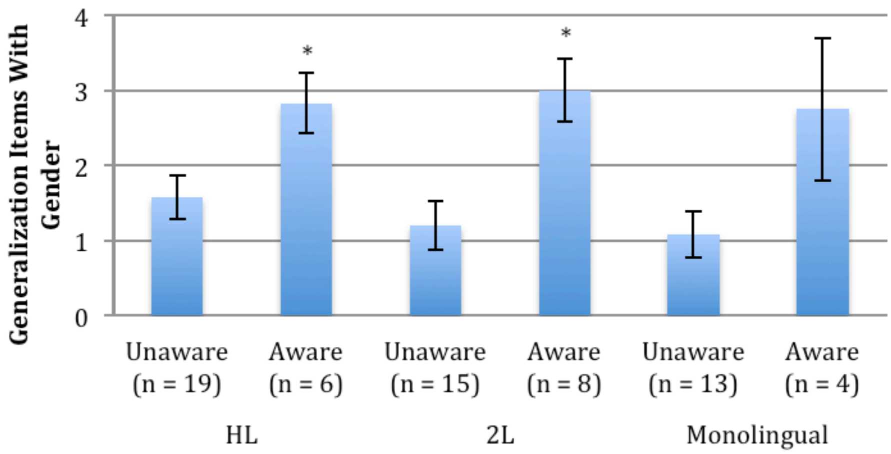

Figure 3: Mean scores for the number of generalization items with biological gender (*chance* = 2).

The aware 2L ($N = 8$, $M = 9.13$, $SD = 2.42$, $t\,(7) = 3.66$, $p < .05$) and unaware 2L ($N = 15$, $M = 7.27$, $SD = 2.37$, $t\,(14) = 2.07$, $p < .05$) participants were the only groups that performed better than chance on the non-gendered generalization items (*chance* = 6). It is surprising that the unaware 2L speakers were able to apply the implicit rule to nouns without biological gender when, as a previous analysis showed, they were unable to do so with nouns that did have biological gender. Of course, it is important to keep in mind that the participants were presented with more generalization without biological gender (n = 12) than those with biological gender (n = 4).

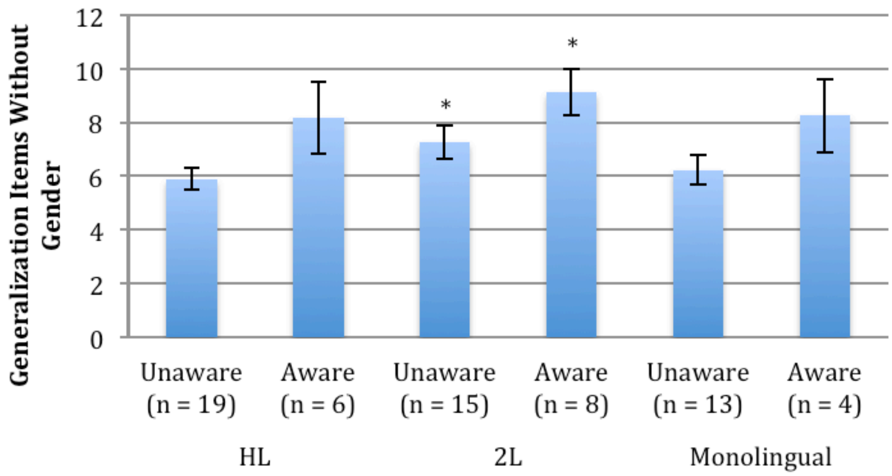

Figure 4: Mean scores for the number of generalization items without biological gender (*chance* = 6).

Second Testing Phase

Unlike the first testing phase, participants in the second testing phase were explicitly told about the presence of a grammatical rule that governed the use of the determiners, were encouraged to find it, and were given feedback to assist them. It is important to keep in mind that participants' performance on the following items may have been influenced by these added factors. After the second testing phase, 9 participants (5 HL and 4 2L) reported becoming aware of the relevance of gender.

A mixed design ANOVA measuring the effects of awareness, item type, and language background on participants' overall performance revealed a main effect of language background $F\ (2, 41) = 4.28, p < .05$. Unlike the first testing phase, there was no main effect of awareness $F\ (1, 49) = 0.93$, *ns*, item type $F\ (1, 41) = 2.57$, *ns*, nor an interaction between awareness and item type $F\ (1, 49) = 1.26$, *ns*. These results suggest that second language learners of Spanish ($M = 8.54$, $SD = 2.15$) and Spanish heritage speakers ($M = 8$, $SD = 2.33$) performed better than monolingual participants ($M = 7.08$, $SD = 2.00$).

To facilitate the comparison between the first and second testing phase, I will be splitting the data in a fashion similar to that of the previous analyses: I will be looking at the role of awareness, item type, and language background. Again, I will briefly report participants' performance on the total testing and trained items and then turn to participants' performance on the generalization items in more detail.

Total Testing Items

When divided by awareness categories and language background, the aware and unaware participants across all of the language backgrounds scored statistically above chance on the total testing items (Figure 5). To determine what drove this effect, I will look at awareness and item type separately.

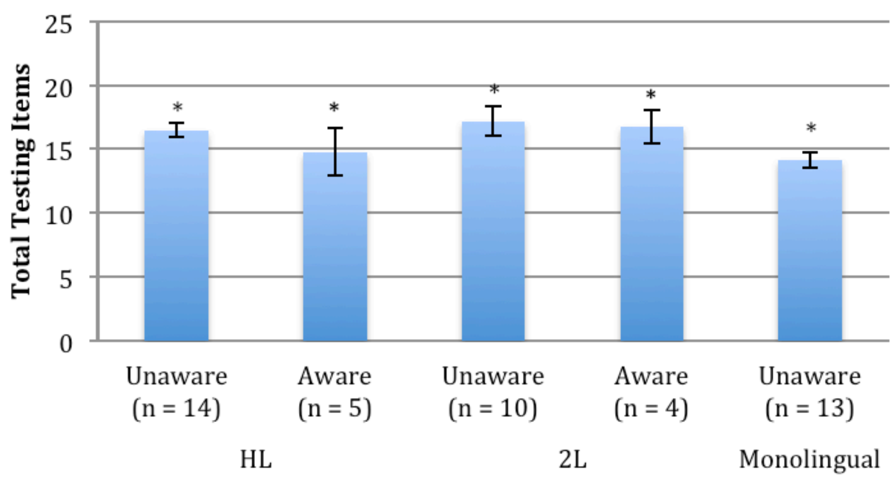

Figure 5: Mean scores for total (trained and generalization items) number of testing items (*chance* = 13).

Trained Items

All of the participants, regardless of awareness and language background, performed better than chance on the total trained items. The aware HL and 2L and the unaware 2L participants performed better than chance on the trained items with biological gender (See Table 3 and Table 4 for descriptive statistics and t-test results). For trained items without biological gender, the aware 2L and aware monolingual participants, as well as the unaware HL, unaware 2L, and unaware monolingual participants performed better than chance.

Generalization Items

None of the aware participant groups scored better than chance (*chance* = 8) on the total number of generalization items, regardless of their language background (Figure 6). The unaware Spanish HL participants (N = 14, M = 9.71, SD = 2.09, t (13) = 3.06, $p < .05$), however, did perform better than chance on these items.

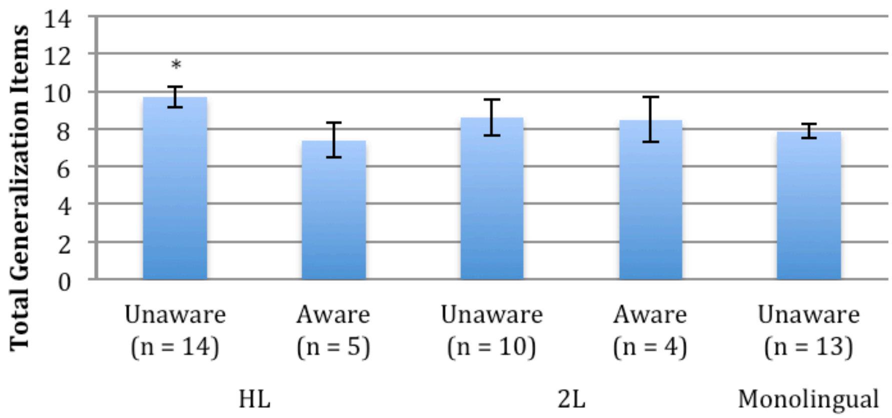

Figure 6: Mean scores for the total number of generalization items in testing phase 2 (*chance* = 8).

In the second testing phase, none of the participants scored better than chance (*chance* = 2) on the generalization items with biological gender, regardless of awareness and language background (Figure 7). One must interpret this result with great caution, as there were few generalization items with biological gender (n = 4).

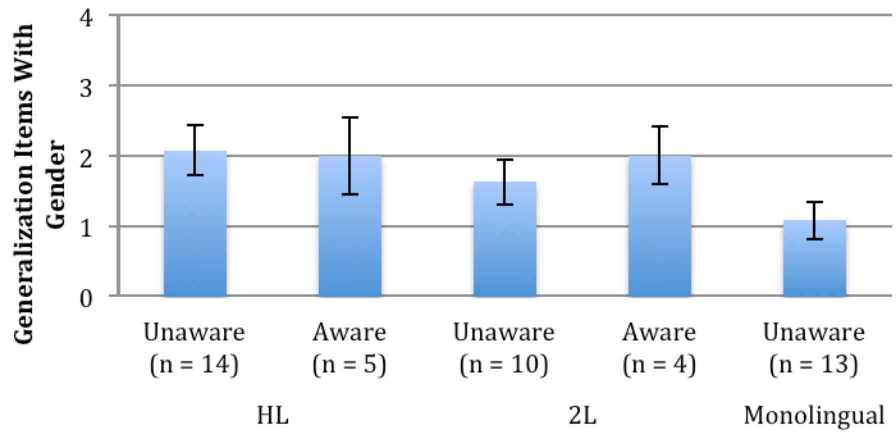

Figure 7: Mean scores for the number of generalization items with biological gender (*chance* = 2).

The aware participants did not perform better than chance (chance = 6) on the generalization items without biological gender, regardless of their language backgrounds (Figure 8). The unaware Spanish HL (N = 14, M = 7.64, SD = 1.78, t (13) = 3.45, $p <$ .05) and monolingual (N = 13, M = 6.77, SD = 1.48, t (12) = 1.87, $p <$ .05) groups, however, did score better than chance on these items. Though the 2L individuals did not perform better than chance, their scores were not far from those in the other groups (N = 10, M = 7.30, SD = 3.30, t (9) = 1.25, *ns*).

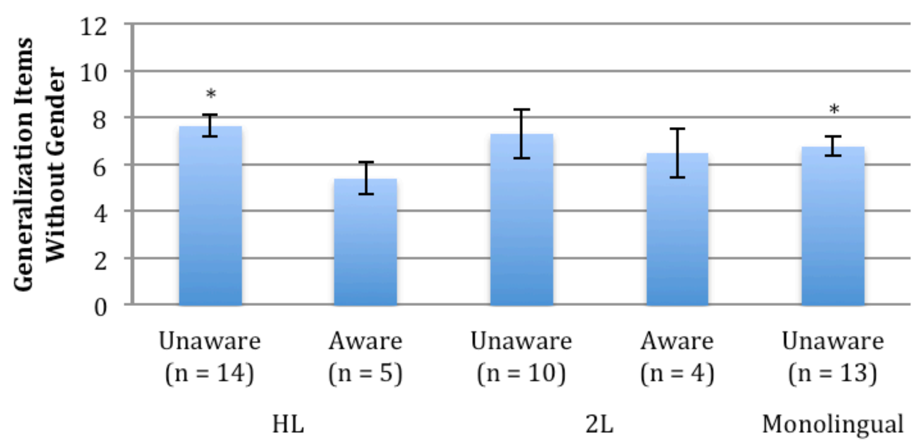

Figure 8: Mean scores for the number of generalization items without biological gender (*chance* = 6).

Chapter 4

Discussion

The present study found that second language learners of Spanish were capable of extracting and using a grammatical rule based on a feature that, although absent in the language tested, was present in their second language. Perhaps more striking is the fact that the second language learners of Spanish used the implicit grammatical rule to choose the correct determiner-noun combinations even when they were unable to justify their decisions at the end of the testing phase; that is, in spite of lacking conscious awareness of such rule. Most unaware participants reported that they had chosen between unfamiliar determiner-noun combinations simply because one option sounded right and/or seemed more familiar. In sum, in line with previous studies involving monolinguals, bilingual participants also show evidence of implicit grammatical learning. Furthermore, it extends these findings to the case where the information available to extrapolate such a grammatical rule exists only in the "non-active language" of bilinguals. I will now turn to a more careful examination of my findings, in order to explain what could have produced the observed results.

First Testing Phase

Focusing on the findings from the trained items in the first testing phase, these rules demonstrated that both aware and unaware participants (independent of specific language background) accurately remembered the determiner-noun combinations presented during the training phase. That is, all of the participants performed better than chance on sentences that contained determiner-noun combinations that were presented during the previous training phase. Note that this analysis was carried out with all of the testing items, including trained items with (e.g. queen, king) and without (e.g. table, squirrel) overt grammatical gender.

If we consider only the generalization items in the first testing phase (i.e. those items participants had not seen before in that specific determiner-noun combination), the aware second language learners of Spanish formed the only group that consistently scored better than chance. It could be the case that for the second language learners of Spanish, who learned rules governing grammatical gender in a more explicit manner, the gender distinction in Spanish is more salient. One of the first obstacles that a native English speaker learning Spanish encounters is the role of grammatical gender. Because English does not encode grammatical gender, these individuals must explicitly memorize which gender corresponds to each noun in Spanish. It could be that the second language learners of Spanish became aware of the experiment's true purpose through knowledge that the experimenter was looking for English/Spanish bilinguals for the study. If the role of gender was a feature that was particularly salient for second language learners of Spanish, it could be that they began questioning whether grammatical gender played a role from the beginning of the experiment. This notion could be supported by the fact that all but one of the aware participants began suspecting whether grammatical gender may have played a role as early as the training phase. Participants' suspicions may have been reinforced by the fact that some of the sentences throughout the training and testing phase included nouns with biological gender (e.g. queen, king, etc.). It is unsurprising that aware second language learners of Spanish would discover the rule by simply looking for supporting evidence in the stimuli. It thus becomes important to look specifically at the performance of second language learners of Spanish that were unable to identify the rule.

In this first testing phase, it was only the group of *unaware* second language learners of Spanish who performed better than chance on the generalization items that lack biological gender. Note that these were completely novel determiner-noun combinations, so it would have been impossible for the participants to have relied on memory. Perhaps enough explicit training in a second language with grammatical gender allows individuals the ability to apply that information unconsciously. This could explain why unaware second language learners were able to extract and use the grammatical gender rule while failing to identify it. If this is the case, then it becomes important to consider what kind of explicit learning strategy could have accounted for the observed results.

Visualization is considered to be a very effective memorization tool, especially when one must memorize a list of words (Gleitman, Gross, & Reisberg, 2011). It could be the case that the second language learners of Spanish formed mental pictures of Spanish nouns that formed a relationship between the noun and its grammatical gender. When learning the noun ‘LIBRO’ (‘book’), for example, a student may imagine a book with a moustache, because this noun is marked with masculine grammatical gender. Asking second language learners of Spanish to visualize the scenes described by the sentences, as was done in the training phase, may have prompted them to access any mental images or associations that could have helped them memorize the grammatical gender of a particular noun in a previous Spanish class. That second language learners were able to use information (mental images from a previous Spanish course and/or the knowledge that certain nouns belong to a certain grammatical gender category in Spanish) from their second language is unsurprising, given that prior research has found that certain aspects of a bilingual’s two languages, such as the phonological information, can be activated simultaneously by merely reading words printed in one of the two languages (Van Leerdam, Bosman, & De Groot, 2009; Van Wijnendaele & Brysbaert, 2002; as cited by De Groot, 2011).

Initially, we expected the Spanish heritage speakers (HL), exposed to Spanish from early childhood and therefore more experienced with a language that encodes grammatical gender, to outperform the other participants. However, the lack of implicit learning in this group may be unsurprising, given that prior research has found that HL speakers produce gender errors in nouns with nontransparent word endings (Montroul, 2010a). Similarly in our study, the English nouns lacked any cues to aid HL participants choose the correct gendered article.

It is somewhat puzzling that the monolingual participants scored as high as the Spanish heritage speakers in a number of generalization items in the first testing phase. In fact, the performance of some of the monolingual participants suggests that these individuals may have known more about a language that encoded grammatical gender than they indicated. Therefore, it is possible that the monolingual group's overall performance included scores from individuals that correctly applied knowledge gained from experience studying a language that encodes grammatical gender. Although we only included people that self-reported as monolingual, some participants may have had more experience with a gendered language than desirable for this study. In fact, the results suggest that the monolingual group's performance was carried by a few exceptional "monolinguals" (i.e. those who had some experience with languages with grammatical gender). It would be difficult for English monolinguals without any background in a language that encoded grammatical gender to score above chance on new items that did not contain biological gender, as they would have had to choose a gender for each noun arbitrarily. The presence of these exceptional monolinguals may have undermined any expected difference between the monolingual and both bilingual groups.

Given that taking a course in a foreign language is a graduation requirement at most high schools and that Reed College, the institution from which monolinguals were recruited, also requires students to take at least a year of a mathematics, foreign language, or linguistics, it is unsurprising that the monolingual sample would include some individuals with experience with a language that encodes grammatical gender. These individuals may have studied a foreign language in high school and/or college, but perhaps not feel fluent enough in that foreign language to consider themselves bilingual. This was a difficult obstacle to prevent as it is challenging to determine whether an individual is truly monolingual or bilingual.

Second Testing Phase

Throughout the second testing phase, the difference in performance between the aware and unaware participants disappeared. This may have been due to the fact that the participants were asked to find the correct rule and they were also given feedback every time they made an incorrect response. Having the ability to know whether they had answered incorrectly might have promoted participants to engage in trial and error, as finding the rule may have appeared more important than choosing the correct determiner-noun combination with accuracy. Engaging in this process may have made the aware participants comfortable making a few incorrect responses in order to test their hypotheses and guarantee that they arrive at the correct rule. Furthermore, the generalization items were no longer novel determiner-noun combinations, as participants had seen the same items used during the first testing phase. The availability of a trial and error strategy to discover the grammatical rule, and the possibility of memory factoring into participants' performance across the generalization items, suggests that one should interpret the findings from this phase with caution. Furthermore, it may be contested whether participants were engaging in implicit learning after they were informed about the existence of a previously hidden grammatical rule and encouraged to find it.

According to the results, both aware and unaware participants in the second testing phase were able to accurately remember the determiner-noun combinations presented during the training phase. That all but the unaware Spanish heritage speakers and monolingual participants scored better than chance on the trained items with biological gender and the trained items without gender could be due to the low number of trained items with gender presented during the testing phase and the low number of aware HL participants ($N = 5$).

The lack of significant difference between aware and unaware participants observed in the first testing phase becomes more pronounced when looking at the differences between these participants on their performances across the generalization items. For the total generalization items in the second testing session, only the unaware Spanish heritage speakers performed greater than chance. For the generalization items without grammatical gender, the unaware Spanish heritage speakers and monolingual participants scored better than chance. The fact that the aware and unaware participants did not significantly differ in their performance across these generalization items could support the notion that the participants were engaging in a trial-and-error strategy, which decreased performance among the aware group during the second testing phase. It is also possible that the small number of participants that became aware of the implicit rule prevented the statistical manipulation from having enough statistical power to produce an effect greater than chance (HL, $N = 5$; 2L, $N = 4$; Monolingual, $N = 0$).

It is striking that the unaware second language learners of Spanish in the second testing phase did not perform better than chance on the generalization items that did not contain biological gender, as the unaware second language learners of Spanish did in the first testing phase. One should notice that the unaware second language learners of Spanish had the smallest number of people (*N* = 10) with the greatest amount of variability (*SD* = 3.30). It could be the case that these individuals did not appear to perform significantly greater than chance on the generalization items without biological gender due to a lack of statistical power.

Relation To Previous Studies

This study questioned whether individuals could learn and apply grammatical features that were not readily apparent, as seen in Williams (2005). This study expanded upon Williams' (2005) methodology by making the target implicit grammatical feature absent in English (apart from the few nouns with biological gender such as king and queen). This study also included individuals from various linguistic backgrounds in order to investigate the relationship between one's linguistic experience and one's ability to implicitly learn particular aspects of an unfamiliar language.

As in Williams' (2005) findings, the majority (75%) of the participants remained unaware of the implicit rule after the first testing phase. Furthermore, the unaware second language learners of Spanish selected the correct determiner-noun combination at levels significantly greater than chance, even for items they had not seen or heard before and lacked biological gender information. Also like Williams' (2005) experiment, I found a significant difference between aware and unaware learners on trained items in the second testing phase (total trained items and trained items with biological gender), where the aware participants outperformed the unaware participants.

Also mirroring Williams' (2005) results, the majority (83%) of the unaware participants remained unaware of the rule that governed the use of the determiners when they were asked to search for the implicit rule during the second testing phase, even when provided with feedback. Unlike Williams (2005), who found that unaware participants performed above chance on all generalization items, in this study only the aware second language learners of Spanish scored better than chance reliably across the various generalization item types in the first testing phase. However, it is possible that, due to the small number of aware participants in the other two groups (HL, *N* = 6; Monolingual, *N* = 4), we just lacked enough statistical power. Furthermore, in the case of the generalization items with biological gender, the monolingual participants may have failed to score significantly above chance due to the low number of these items ($n = 4$).

The results from the initial mixed design ANOVA mirrored the results found by Faretta-Stutenberg and Morgan-Short (2011), as there was no relationship between language background and performance in aware or unaware participants in the first testing phase. Put simply, when looking at the data as a whole, unaware Spanish heritage speakers showed no difference from unaware monolingual participants or second language learners of Spanish. After accounting for outside factors by dividing the participants by awareness and the stimuli by item type, however, we then found that unaware second language of learners of Spanish were able to extract and apply the implicit rule.

Limitations

It is possible that participants who, after the second test phase became aware, may have had a partial idea or suspicion that gender played a role during the first test phase, but they were unable to articulate their suspicion at that point. However, participants who were aware at the end of the second testing phase said that they were not certain of the relevance of gender prior to the instruction to search for rules. Despite this possible confounding variable, it is still important to note that in the first testing phase, the majority (71%) of 2L participants, who performed significantly above chance on generalization items (without biological gender information) neglected to mention the role of gender at all.

Another drawback of this experiment was the oversight on the part of the experimenter to counterbalance the response keys assigned to each determiner. That is, the masculine determiners were always located to the left of the feminine ones for both sets of determiners (see Figure 1) and for each participant. In fact, four monolingual participants explicitly indicated using the orientation of the buttons to choose between the generalization items in the first testing phase. It is possible that other participants may have used this trick to make their decisions and failed to report it. Despite this limitation, the fact that the aware and unaware 2L participants were the only individuals to perform better than chance on generalization items without biological gender remains impressive.

If a large quantity of participants noticed and used the pattern in the positioning of the buttons, one would expect to observe a ceiling effect. One's ability to notice a pattern in the orientation of the buttons arguably has little to do with one's proficiency in Spanish. In fact, all of the participants that reported using the alignment of the buttons to make decisions were in the monolingual group.

Initially, we decided to include items with biological gender as a way to help participants notice a grammatical feature that does not exist in English. However, including these items may have helped participants to consider grammatical gender playing a role. As previously mentioned, this may have helped the second language learners of Spanish determine which linguistic feature the experiment was investigating. Despite the prevalence of clues throughout the training and testing phases referring to grammatical gender in the stimuli, the majority (71%) of participants were still unable to express the rule after the first testing phase.

Conclusion

The present study attempted to replicate and extend the findings of Williams (2005), which examined whether grammatical rules governing form and meaning in an artificial grammar could be learned implicitly. There were a few adjustments made to the participant pool in that the participants came from different language backgrounds (Spanish heritage speakers, second language learners of Spanish, and monolinguals) and the implicit grammatical rule was a feature that did not exist in English. This study found evidence of learning without awareness in second language learners of Spanish. Given the conflicting evidence in the literature involving implicit learning, further research must be conducted in order in order to confirm and extend these results, perhaps adopting a less problematic gauge of awareness, counterbalancing the response keys, and omitting the nouns with biological gender.

Appendix A: Training Phase Items

Training Items Set 1

1. I was startled when I turned around and saw *gi* butler right behind me.
2. While I was digging in the garden *gi* rabbit came and sat next to me.
3. The children threw sticks at *ne* squirrel in the tree.
4. The subjects admired *ne* queens from a distance.
5. I spent an hour polishing *ro* table before the dinner party.
6. The archaeologists worked for hours to restore *gi* vases.
7. In the pub I asked my friend to get *ne* bottle from the other side of the room.
8. I could not sleep because of the chiming of *ul* clocks downstairs.
9. While I was out for a walk I saw *gi* king and asked for an autograph.
10. The ball was kicked by *ro* nun during the church picnic.
11. Sitting under the tree I was bothered by *ro* flies.
12. The circus performer covered herself with *ro* snakes.
13. I could hear *ul* mouse scurrying around in the roof.
14. The fire brigade had to rescue *ul* prince from the top floor of the castle.
15. I could hear *ne* princess speak from the other side of the patio.
16. I was terrified when I heard *ul* bears roaring in the distance.
17. After work I fell asleep on *gi* cushions in front of the TV.
18. I spent the night on *gi* sofa and let my guests sleep in the beds.
19. I knocked over *ro* teacup and the coffee spilled on my book.
20. I can't move in my office because *ro* boxes are piled on the floor all around me.
21. I couldn't read the title of *ul* book that was on the top shelf.
22. I had difficulty hearing *ne* news when I was on the other side of the room.
23. In the gallery we all admired *ne* photographs from the other side of the room.
24. At the fair they threw balls at *ul* plates across the range.

Training Items Set 2

1. At the banquet, the children were excited when *gi* butlers served the pastries.
2. I was amazed when *gi* rabbit ate from my hand.
3. I saw *ne* squirrels running across the forest with my binoculars.
4. The child screamed in excitement even though *ne* queen had not yet arrived.
5. The children all sat down at *ro* tables to play board games.
6. As I was passing I knocked over *gi* vase.
7. We hoped that *ne* chairs would be delivered in time for the party.
8. I looked up at *ul* clock on the church and realized that I was late.
9. My son greeted *gi* kings as we walked down the path.
10. The walkers were surrounded by *ro* nuns in the convent.
11. The little boy tried for hours to swat *ro* fly in the kitchen.
12. At the zoo none of the children wanted to touch *ro* snake.
13. We were told that *ul* mice in the roof would breed if not killed quickly.
14. Last night we heard *ul* princes arguing in the road outside.
15. I was distressed by the scream of *ne* princesses coming from the palace over the hill.
16. We were able to observe *ul* bear wandering around the hillside across the valley.
17. I told my son to put his head on *gi* cushion and go to sleep.
18. The guests entered the living room and sat down on *gi* sofas.
19. When the waitress dropped *ro* teacups on the floor she was dismissed.
20. I carefully packed my nephew's present in *ro* box before sending it to France.
21. I asked the librarian to fetch *ul* books from the stacks.
22. While working in his lab, the scientist heard *ne* news from the office upstairs.

23. The office workers threw darts across the room at *ne* photograph of their manager.
24. From the kitchen I heard the sound of *ul* plate crashing to the floor.

Appendix B: Testing Phase Items

Trained 1

1. The lady spent many hours sewing *gi* cushions.
2. The woman screamed when she saw *ul* mouse across the street.

Generalization 1

1. While lying in the field, the girl was greeted by *ro* queens.
2. While at the nursery the baby sucked on *ro* bottle.
3. The hunter wanted to catch *ul* rabbit in the distance.
4. Before the battle the soldiers checked the time on *gi* clocks.
5. The waitress was told to go and clear *ne* table.
6. The trainer struggled to control *ro* squirrel.
7. The thieves searched the whole house for *ul* vases.
8. I called *ul* butler from the other side of the mansion.

Trained 2

1. After my meal I went to the sink to wash *ro* teacup.
2. Before the ball, the hero considered courting *ne* princesses.
3. I searched everywhere for *ul* book.
4. The doctor began to examine *ro* nun in the infirmary.
5. The tourists struggled to see *ne* photographs from where they were sitting.
6. I was scared when I heard the hiss of *ro* snakes right behind me.
7. I bought a present for *ul* prince that lives across the street.
8. I packed all of my belongings into *ro* boxes right beside me.

Generalization 2

1. While waiting by the staircase, I helped *ro* queen down the steps.
2. After the meeting we collected *ro* bottles from our table.
3. In the prairie the naturalist could not get a good shot of *ul* rabbit.
4. I had to read the manual to find out how to adjust *gi* clock.
5. The customer pointed to his friends sitting at *ne* tables across the restaurant.
6. The boy in the woods played with *ro* squirrels.
7. After lunch, the workmen would fix *ul* vase at the far end of the square.
8. At the dinner party, the mayor pointed to *ul* butlers across the hall.

Appendix C: Tables

Table 1: Mean percentage correct, standard deviations, and t-values for deviation from chance (t) for unaware participants. (First testing phase)

	Total Testing (*chance*=13)	Total Trained (*chance*=5)	Total Generalization (*chance*=8)	Trained With Gender (*chance*=1.5)	Generalization With Gender (*chance*=2)	Trained Without Gender (*chance*=3.5)	Generalization Without Gender (*chance*=6)
HL							
M	15.32	7.84	7.47	2.00	1.58	5.84	5.89
SD	3.16	1.86	2.32	0.82	1.26	1.34	1.82
*t (*18)	3.19*	6.65*	-0.99	2.67*	-1.45	7.59*	-0.25
Mono							
M	14.62	7.31	7.31	1.69	1.08	5.62	6.23
SD	1.89	1.75	2.53	0.85	1.11	1.04	1.96
*t (*12)	3.07*	4.75*	-0.99	0.81	-2.98	7.31*	0.42
2L							
M	17.07	8.60	8.47	2.53	1.20	6.07	7.27
SD	3.41	1.50	2.83	0.74	1.26	0.96	2.37
*t (*14)	4.62*	9.28*	0.64	5.38*	-2.45	10.34*	2.07*

The table shows mean percentage correct, standard deviations, and t-values for deviation from chance (t).

* $p < .05$

Table 2: Mean percentage correct, standard deviations, and t-values for deviation from chance (t) for aware participants (First testing phase).

	Total Testing *(chance*=13)	Total Trained *(chance*=5)	Total Generalization *(chance*=8)	Trained With Gender *(chance*=1.5)	Generalization With Gender *(chance*=2)	Trained Without Gender *(chance*=3.5)	Generalization Without Gender *(chance*=6)
HL							
M	19.67	8.67	11.00	2.50	2.83	6.17	8.17
SD	5.09	1.86	3.85	1.22	0.98	0.75	3.31
*t (*5)	3.21*	4.82*	1.91	2.00	2.08*	8.68*	1.60
Mono							
M	20.50	9.50	11.00	3.00	2.75	6.50	8.25
SD	4.73	0.58	4.24	0.00	1.89	0.58	2.75
*t (*3)	3.17*	15.59*	1.41	.	0.79	10.39*	1.63
2L							
M	21.38	9.25	12.13	2.75	3.00	6.50	9.13
SD	4.87	1.75	3.44	0.71	1.20	1.07	2.42
*t (*7)	4.87*	6.86*	3.40*	5.00*	2.37*	7.94*	3.66*

The table shows mean percentage correct, standard deviations, and t-values for deviation from chance (t).

* $p < .05$

Table 3: Mean percentage correct, standard deviations, and t-values for deviation from chance (t) for unaware participants (Second testing phase).

	Total Testing *(chance*=13)	Total Trained *(chance*=5)	Total Generalization *(chance*=8)	Trained With Gender *(chance*=1.5)	Generalization With Gender *(chance*=2)	Trained Without Gender *(chance*=3.5)	Generalization Without Gender *(chance*=6)
HL							
M	16.50	6.79	9.71	1.57	2.07	5.21	7.64
SD	2.10	2.01	2.09	0.51	1.33	2.08	1.78
*t (*13)	6.23*	3.33*	3.06*	0.52	0.20	3.08*	3.45*
Mono							
M	14.15	6.31	7.85	1.23	1.08	5.08	6.77
SD	2.23	2.25	1.41	0.83	0.95	1.89	1.48
*t (*12)	1.87*	2.10*	-0.39	-1.17	-3.49	3.01*	1.87*
2L							
M	17.20	8.60	8.60	2.63	1.63	5.63	7.30
SD	3.71	1.51	2.99	0.52	0.92	1.30	3.30
*t (*9)	3.58*	7.56*	0.63	6.15*	-1.16	4.61*	1.25

The table shows mean percentage correct, standard deviations, and t-values for deviation from chance (t).

* $p < .05$

Table 4: Mean percentage correct, standard deviations, and t-values for deviation from chance (t) for aware participants (Second testing phase).

	Total Testing (*chance*=13)	Total Trained (*chance*=5)	Total Generalization (*chance*=8)	Trained With Gender (*chance*=1.5)	Generalization With Gender (*chance*=2)	Trained Without Gender (*chance*=3.5)	Generalization Without Gender (*chance*=6)
HL							
M	14.80	7.40	7.40	2.60	2.00	4.80	5.40
SD	4.21	2.19	2.07	0.55	1.22	1.79	1.52
*t (*4)	0.96*	2.45*	-0.65	4.49*	0.00	1.63	-0.88
2L							
M	16.75	8.25	8.50	2.50	2.00	5.75	6.50
SD	2.63	1.50	2.38	0.58	0.82	0.96	2.08
*t (*3)	2.85*	4.33*	0.42	3.46*	0.00	4.70*	0.48

The table shows mean percentage correct, standard deviations, and t-values for deviation from chance (t).

* $p < .05$

Appendix D: Figures

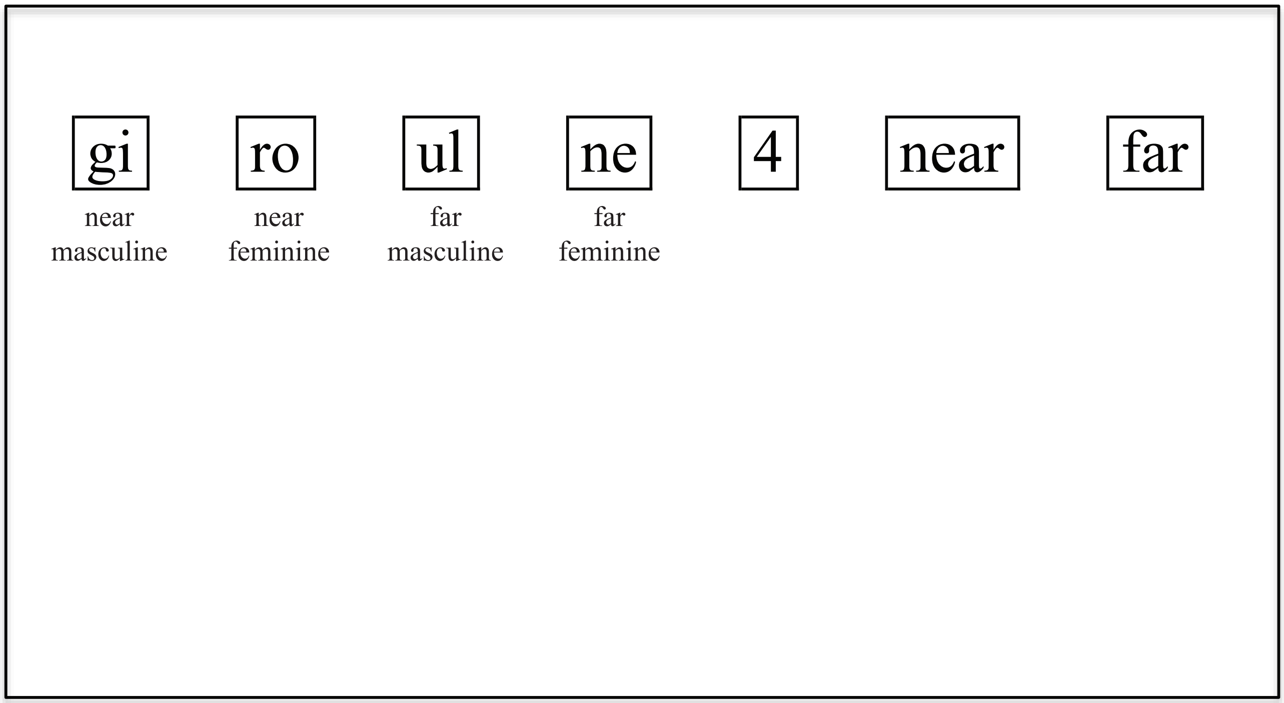

Figure 9: The button box used throughout the experiment.

The button labeled ‘4’ was used to proceed through the written instructions.

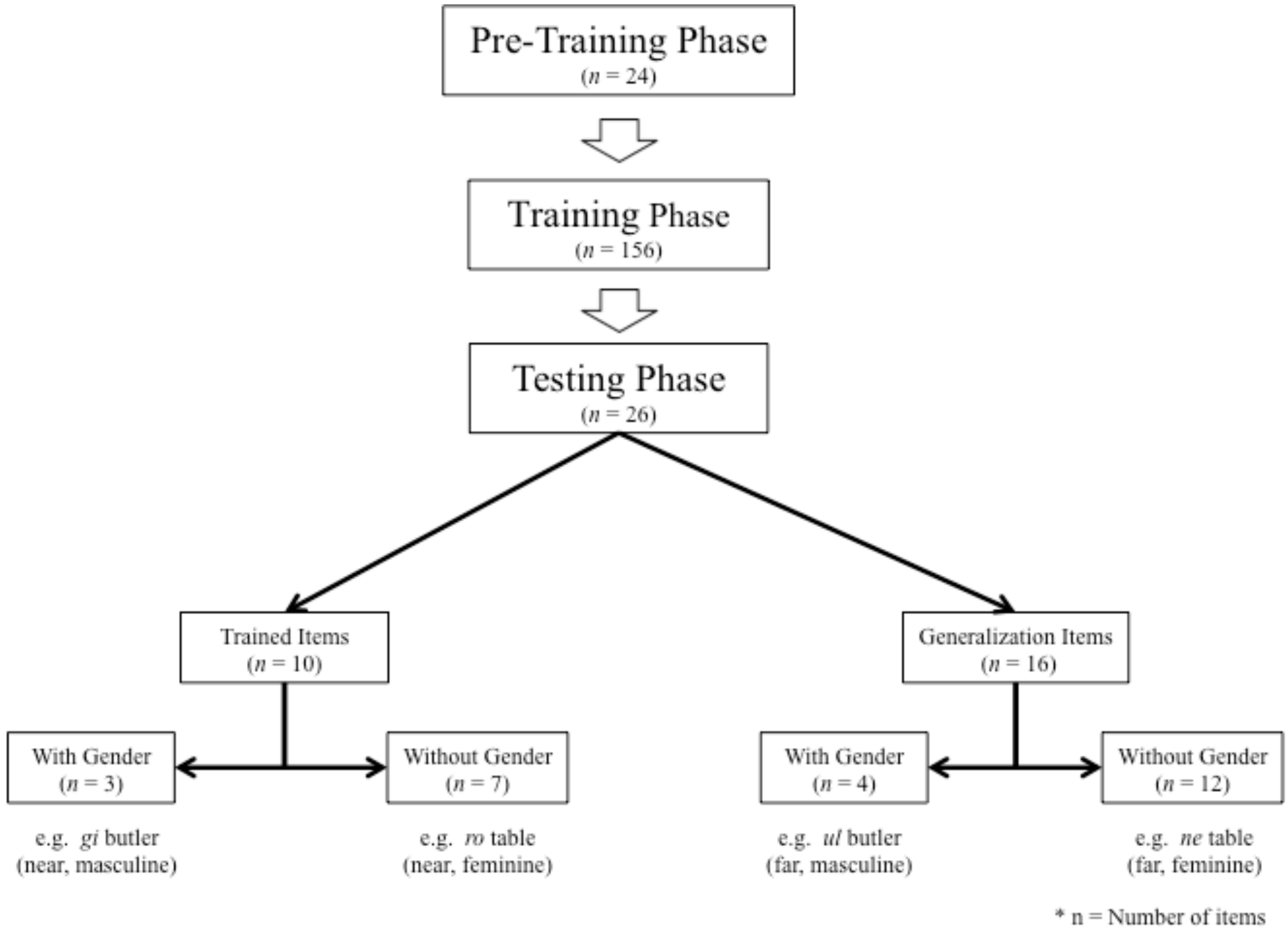

Figure 10: Summary of the experimental design.

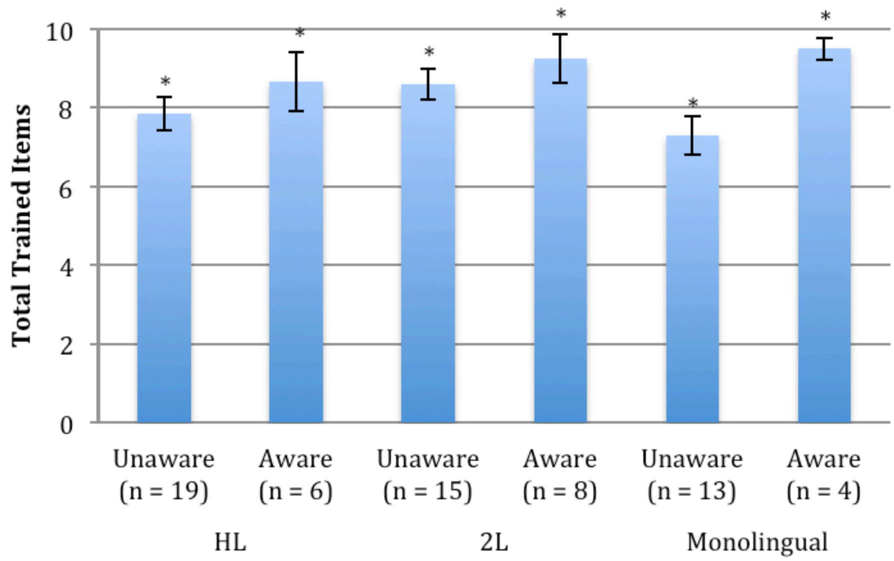

Figure 11: Mean scores for the total number of training items in testing phase 1 *(chance* = 5).

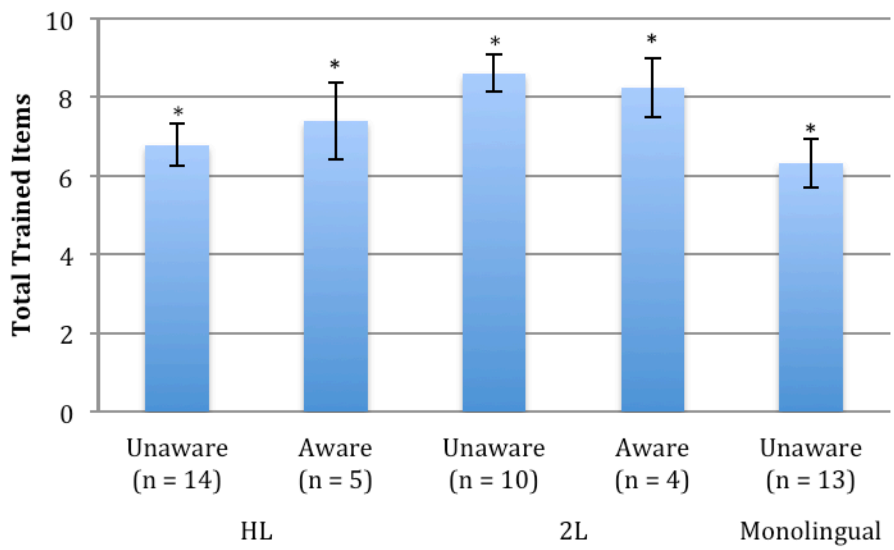

Figure 12: Mean scores for the total number of training items in testing phase 2 *(chance* = 5).

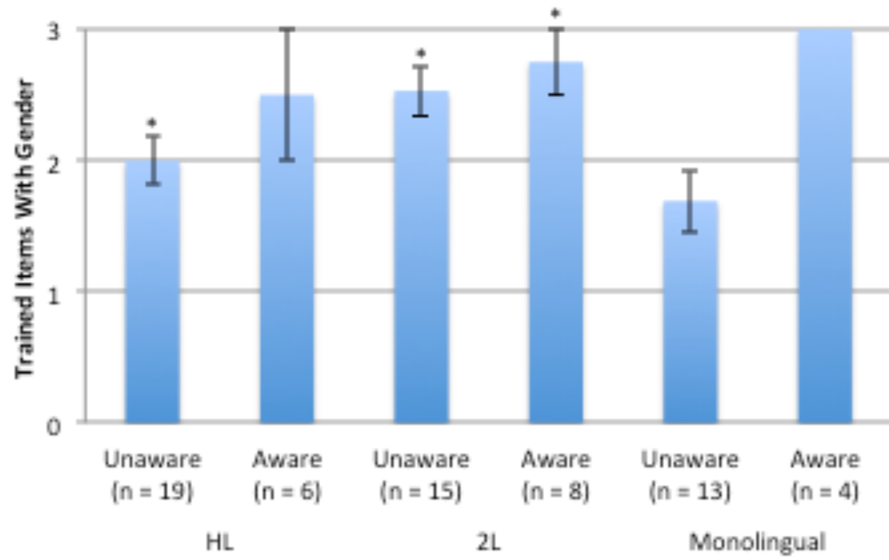

Figure 13: Mean scores for the number of training items with biological gender in testing phase 1 *(chance* = 1.5).

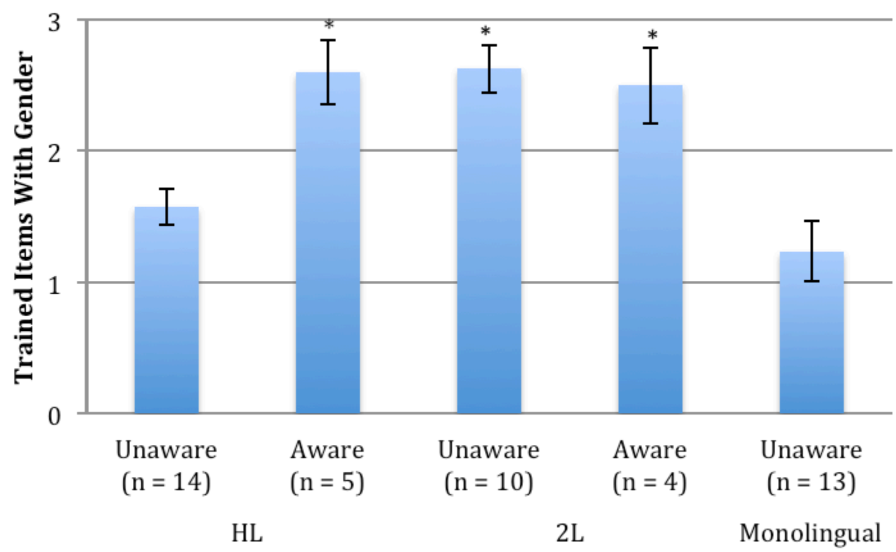

Figure 14: Mean scores for the number of training items with biological gender in testing phase 2 *(chance* = 1.5).

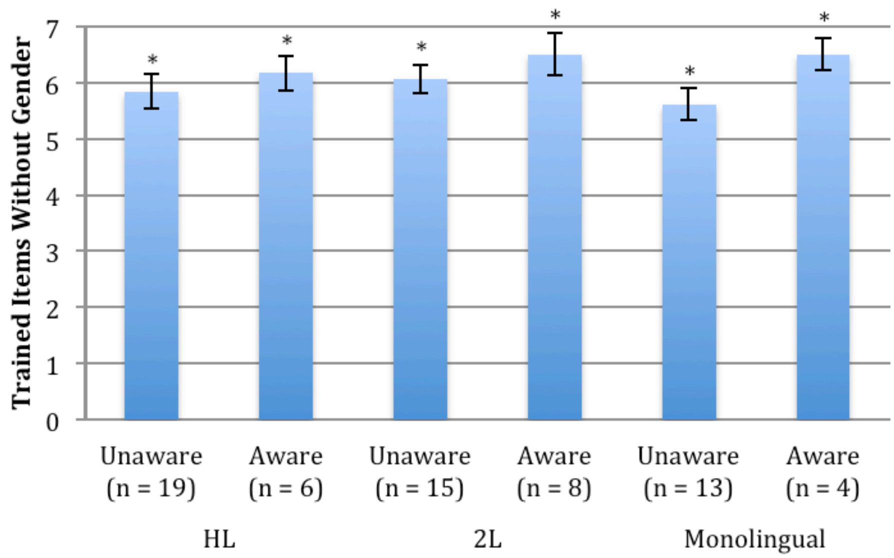

Figure 15: Mean scores for the number of training items without biological gender in testing phase 1 (*chance* = 3.5).

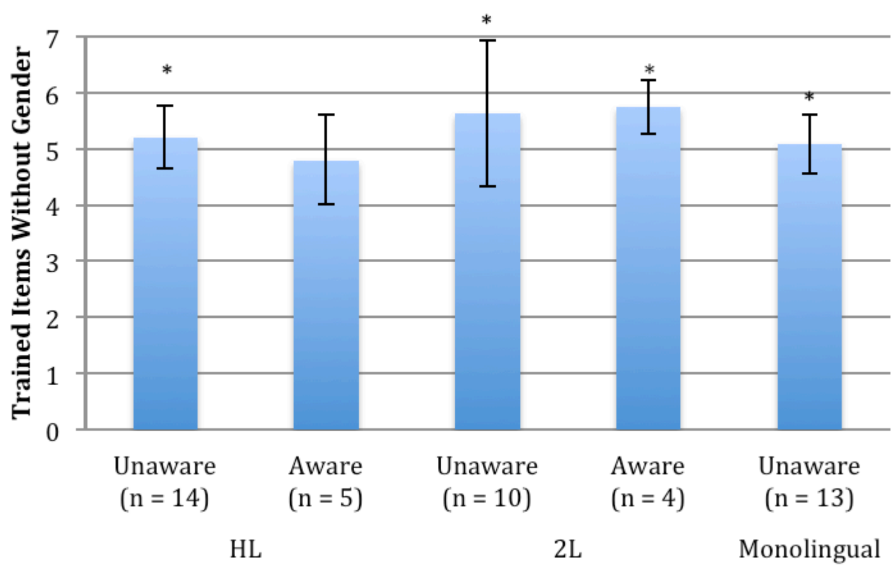

Figure 16: Mean scores for the number of training items without biological gender in testing phase 2 (*chance* = 3.5).

Bibliography

Bolger, P. A., & Zapata, G. C. (2011). Psycholinguistic approaches to language processing in heritage speakers. *Heritage Language Journal, 8*, 1-29.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind: Advances in the Study of Language and Thought* (pp. 61-79). Cambridge, Mass: MIT Press.

Cleary, A. M., & Langley, M. M. (2007). Retention of the structure underlying sentences. *Language and Cognitive Processes, 22*, 614-628.

Davidson, J., de la Fuente, I., & Montrul, S. & Foote, R. (2011). Early language experience facilitates gender processing in Spanish heritage speakers. *A supplement of the Proceedings of the 35th Annual Boston University Conference on Language Development*.

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals : An introduction*. New York: Psychology Press.

Faretta-Stutenberg, M., & Morgan-Short, K. (2011). Learning without awareness reconsidered: A replication of Williams (2005). *Selected Proceedings of the 2010 Second Language Research Forum: Reconsidering SLA Research, Dimensions, and Directions*, 18-28.

Francis, A. P., Schmidt, G. L., Carr, T. H., & Clegg, B. A. (2009). Incidental learning of abstract rules for non-dominant word orders. *Psychological research, 73*, 60-74.

Gleitman, Henry, Gross, James J., & Reisberg, Daniel. (2011). *Psychology*. New York: Norton.

Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition, 32*, 465-491.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*, 940-967.

Montrul, S. (2010a). Current issues in heritage language acquisition. *Annual Review of Applied Linguistics, 30*(2), 3-23.

Montrul, S. (2010b). How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics, 31*, 167-207.

Montrul, S. (2012). Is the heritage language like a second language? *EUROSLA Yearbook, 12*(1), 1-29.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning, 58*, 503-553.

Montrul, S., & Perpiñán, S. (2011). Assessing differences and similarities between instructed heritage language learners and L2 learners in their knowledge of Spanish tense-aspect and mood (TAM) morphology. *Heritage Language Journal, 8*, 90-133.

Potowski, K., Jegerski, J., & Morgan-Short, K. (2009). The effects of instruction on linguistic development in Spanish heritage language speakers. *Language Learning, 59*, 537-579.
Potowski, Ki., Parada, M., & Morgan-Short, K. (2012). Developing an online placement exam for Spanish heritage speakers and L2 students. *Heritage Language Journal, 9*, 51-76.
Reber, A. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855-863.
Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics, 33*, 829-856.
Roca, Ana, & Colombi, MarÌa Cecilia. (2003). *Mi lengua : Spanish as a heritage language in the United States, research and practice*. Washington, D.C.: Georgetown University Press.
Stadler, Michael A., & Frensch, Peter A. (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage.
Suárez-Orozco, M. M., & Paez, M. M. (2009). *Latinos : Remaking America*. Berkeley, CA: University of California Press.
Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition, 27*, 269-304.