

# The SLA of Direct Object Pronouns in a Study Abroad Immersion Environment Where Use Is Variable

Kimberly L. Geeslin, Lorenzo J. García-Amaya, Maria Hasler-Barker, Nicholas C. Henriksen, and Jason Killam  
Indiana University, Bloomington

## 1. Introduction

As interest in the relationship between study abroad (SA) and second language (L2) development has grown over the past few years, so has the body of research on the topic. It has been shown that rather than uniformly sustaining positive effects on language learning, the benefits of experience in the target-language environment are varied and appear to be mediated by factors such as the grammatical structure under examination, characteristics of the non-native speaker (NNS), and the type of abroad program itself (DeKeyser, 1986, 1991; Díaz-Campos, 2004; Freed, Segalowitz, & Dewey, 2004; Lafford, 1995, 2004; Segalowitz & Freed, 2004; Segalowitz et al., 2004). Despite this growth in research on the SA context, little has been done to connect work on SA with work on another area of second language acquisition (SLA) research, namely the SLA of variable structures. We know that native-like variation is an essential part of Communicative Competence (Canale & Swain, 1980) and that studies of frequency and the conditions that predict the variants under investigation have contributed to our knowledge of NNS grammars (Adamson & Regan, 1991; Geeslin, 2003; Geeslin & Guijarro-Fuentes, 2006; Gudmestad, 2006; Mougeon & Dewaele, 2004; Mougeon, Rehner & Nadasdi, 2004). Nevertheless, most of the research on SA has focused on properties that are relatively consistent across native-speaker populations. The current study examines the SLA of direct object (DO) pronouns, a grammatical structure known to vary geographically throughout the Spanish-speaking world (García & Otheguy, 1974; Klein-Andreu, 2000), by NNSs participating in an abroad immersion program in León, Spain. In addition to connecting two growing bodies of research (on the effects of SA and on the SLA of variable structures), the current study contributes to an extensive body of research on the SLA of object pronouns by adding longitudinal data to a database that is largely cross-sectional.

This paper is organized as follows. Section 2 is a review of the existing research on the effects of SA on language learning and on the SLA of variable structures, specifically those in Spanish. Section 3 provides information about the current research design, including participants and methodology. Section 4 presents the results and compares NNS data to NS data. Section 5 discusses the findings and makes connections with previous research on SA and the SLA of variable structures. Finally, Section 6 concludes and provides options for future research.

## 2. The research context

### 2.1. *The effects of SA on language learning*

It is commonly believed that time spent living in a SA environment is beneficial for language acquisition. Early studies, such as Carroll (1967), for example, demonstrated positive gains for a large number of NNSs (N=2,782), all college seniors majoring in French, German, Italian, Russian or Spanish in 203 institutions across the US. Participants were tested on listening, speaking, reading, and

writing, and overall results revealed that even those learners with short stays in the target environment improved their L2 skills. Subsequent research has attested to the complexity of the effect of SA on language learning. For example, we know that not all abilities develop equally (see Freed (1995) and Freed & Collentine (2006) for reviews). Whereas areas such as oral and written fluency, lexical abilities, narrative abilities, and sociolinguistic awareness tend to show improvement after a SA experience (Collentine, 2004; Dekeyser, 1986; García-Amaya, 2008; 2009; Lord, 2009; Segalowitz & Freed, 2004), areas such as grammatical precision often show mixed results or even an advantage for at-home (AH) learners (Collentine & Freed, 2004; Dekeyser 1986, 1991). Likewise, studies on pronunciation and the effects of SA have shown mixed results. Three studies in particular have shown different results. While Díaz-Campos (2004) did not find a significant difference between SA and AH groups, Stevens (2001) and Díaz-Campos (2006) found that SA was superior in pronunciation than AH. Similarly, Simões (1996), who examined one SA group, reported that participants did improve their pronunciation abroad. Other variables, such as number of years of study, for example, have proven to be better predictors of improvement rather than whether NNSs have been abroad (Díaz-Campos & Lazar, 2003).

Although it is clear that various elements of a NNS's competence may be affected differentially by the SA experience, Lafford & Collentine (2006) point out several reasons why some of the variation in the current database may actually be a reflection of the research methodology employed. For example, it has been shown that NNSs vary in the degree to which they interact with their hosts and, thus, the quantity of input they receive may differ (Díaz-Campos, 2004; Segalowitz & Freed, 2004; Wilkinson, 2002). Likewise, because most participants in research on SA are at the intermediate level, variability may be a result of the fact that they have not reached a minimum proficiency threshold for clear gains to be demonstrated. This would explain why studies such as Isabelli & Nishida (2005), whose participants were more advanced, were able to show clear gains in morphosyntax (i.e., subjunctive and complex sentence structure). In addition to these two factors, Lafford & Collentine (2006) note that the average length of stay is relatively short in most studies, that individual NNS characteristics have not been widely examined (see DeKeyser (1986) for an exception), and that there is an extensive use of instruments that do not measure the elements of communicative competence most likely to improve during SA. In other words, although we continue to believe that SA is beneficial, the variation in the findings across studies points to a need for continued research on the topic. Although the current project makes no comparison between AH and SA learners, we build on current research by adding to the existing database a longitudinal study with highly motivated NNSs in a contact-intensive immersion program. Moreover, we extend this body of research to a grammatical structure that is known to vary from one target environment to another. The examination of structures that vary is especially valuable given the fact that one's ability to vary speech according to the social context is an essential, but understudied, component of communicative competence. Thus, in order to understand the process of acquiring native-like language use, one must know how structures that vary across and between native speakers are incorporated into developing interlanguage grammars.

## *2.2. The SLA of variable structures in Spanish*

There are still some critics of the value of studying variation in NNS language (see Ellis (1997) for a discussion), and empirical research on L2 variation continues to outnumber attempts to address how linguistic theories might incorporate such variation (see Preston (2000) and Tarone (2008) for exceptions). Nevertheless, advances in sociolinguistics are generally met with interest by L2 researchers, and research on variation in L2 has provided important insights into the process of acquisition as well as into how differences arise between NSs and NNSs of a language. We now know that NNS language varies according to social, linguistic and contextual features, and our knowledge of how particular grammatical structures vary, as well as the way in which NNS language varies in general has grown tremendously (see, for example, Adamson & Regan, 1991; Mougeon, Rehner & Nadasdi, 2004). One key insight gleaned from this body of research is the distinction between vertical or Type I variation, which is linked to second language development, and horizontal or Type II variation, which is the type of variation found in native speech as well (Rehner, 2002). The former

refers to stages prior to native-like use during which a NNS alternates between a native-like form and a non-native-like form, and this alternation provides evidence of how the grammar is restructured as acquisition proceeds (see Ellis (1997) for details). The latter is the object of study of sociolinguistics and also an essential element of Communicative Competence (Canale & Swain, 1980). The importance of this distinction will become clear in the discussion of the current study where changes over time must be linked to both the forms known to vary as development takes place regardless of the learning context and also to the variation known to exist in the target language spoken in the region where the immersion program was located.

The number of studies on the SLA of variable structures in Spanish has increased substantially in the last decades, and a wide variety of structures have now been examined. In research on the variability in the contrast between the subjunctive and indicative moods, it has been found that NNSs use the subjunctive significantly less than NSs do on the same tasks, even though the variables that predict mood choice are similar for both groups (Geeslin & Gudmestad, 2008a; Gudmestad, 2006). Research on the SLA of copula contrast has demonstrated that NNSs reach similar frequencies of use at the highest levels of proficiencies and that this use is predicted by the same linguistic variables, even though there are subtle differences in the strength of those predictors (Geeslin 2003; Geeslin & Guijarro-Fuentes, 2006). Likewise, Geeslin & Gudmestad (2008b), in an examination of several of the variables identified in sociolinguistic research to predict subject expression, found that there were qualitative differences between NS and near-NS use of the forms of subject expression, even though the range of forms produced was the same. Similar findings have recently begun to emerge for the marking of future and past time reference as well (Geeslin & Gudmestad, 2010). For example, Gudmestad & Geeslin (2007) examined the predictive power of several independent linguistic variables that predict forms of future time reference and found that only certain factors were good predictors of NNSs use, whereas additional factors were significant predictors of NS use. In combination, these studies demonstrate that NNSs are capable of acquiring native-like constraints on variation, even if their use of these factors is not yet native-like. Moreover, this research demonstrates the complexity of the phenomenon and the degree to which such potentially-variable structures are useful in exploring NS and NNS differences.

Despite the increase in studies on variation in L2 Spanish, and the breadth of structures examined, we know of only one study that addresses the relationship between variation in the L2 and the variety of Spanish to which the NNS is exposed. In a preliminary examination on geographically-linked variation (e.g., /s/-weakening and use of /θ/), Geeslin & Gudmestad (2008c) demonstrated that this type of variation is not acquired at the earliest stages of development and that despite urban myths to the contrary, no single individual characteristic explained why such variants were acquired by some NNSs and not by others. One of the primary goals of the current study is to extend the body of research on L2 variation by addressing the influence of a particular target-variety on NNS development and by adding longitudinal data to a database that is nearly exclusively cross-sectional.

### 2.3. Spanish direct object pronouns and NS use

Sociolinguistic research on DO pronouns has focused on the variability that exists within and across social groups as well on as variation associated with particular geographic regions. One type of variation with DO pronouns, known as *leísmo*, may be defined as the use of the dative pronoun *le* in place of the etymological accusative pronouns *lo* and *la* with animate and inanimate referents of both genders. This type of variation is associated with Castile and León Spain, the site of the current investigation (Klein-Andreu, 2000). The variables used to design the task and analyze data in the current study are taken from this body of research (all variables are also described in detail in section 3.3). One early account of *leísmo*, called the Form-Content approach (García & Otheguy, 1977), compares the use of *le*, *lo* and *la*, and the subject of the verb based on activity level. According to this account, the subject is the most active participant, followed by *le*, in turn followed by *lo* and *la*. García & Otheguy confirm a number of predictions: (1) that *le* is more likely to occur when the gender of the object is unknown than when it is known (i.e. *Le llamo hoy* can mean ‘I’ll call him today’ or ‘I’ll call her today’ in a context in which the gender of the person to be called is unknown); (2) that *le* is more likely to occur when the object is male than when it is female; and (3) that *le* is more likely to be used

when the subject of the verb is inanimate (regardless of the animacy of the object). Comparing the strength of these three variables across seven groups, there is a striking trend in the strength of the use of *le* for known male referents in Spain. For the purpose of the current study, both the animacy of the subject and the gender of the direct object will be examined. In order to address whether the referent was known to the speaker or not, we also included a variable called co-referentiality in our design. This variable distinguishes contexts where the referent of the object is also verbalized in the same utterance from those where it is not.

Klein-Andreu (2000) identifies gender and number as predictors of variation in the use of *le*. In Valladolid, near León, *leísmo* is the preferred form for human masculine referents (up to 95% depending on social class), and there is little variation between singular and plural. However, the highest social class shows variation between singular and plural female referents (34% *leísmo* for singular and 7% for plural). Based on these results, we will examine the influence of grammatical number of the object on DO pronoun choice. Finally, Flores-Cervantes (2002) has linked *leísmo* to Transitivity (Hopper & Thompson, 1980) such that predicates with lower Transitivity co-vary with the use of *le* in accusative contexts. According to Hopper and Thompson's model, telic verbs are higher in transitivity than atelic verbs. They describe transitivity as "transferring an action from one participant to another" (1980:253). The more transitive a structure is, the more the object acts as a prototypical direct object. On the other hand, atelic verbs would lead to a less transitive structure and an object that is less prototypically a direct object. If the direct object is seen as less prototypical, it loses its 'directness' and becomes more 'indirect'. Thus, the prediction is that atelic predicates, are more likely to occur with *le* than telic predicates, and this too will be tested in the current study. This body of sociolinguistic research demonstrates that variation exists in the use of direct object pronouns. Nevertheless, no study has yet to examine the SLA of direct object pronouns in a region where *leísmo* exists (i.e., where learner input is known to be variable). Such research will provide important insights as to how learners deal with variable input and incorporate this knowledge into their interlanguage grammars.

#### 2.4. The SLA of direct object pronouns

Most of the research on the SLA of DO pronouns has focused on the notion of the first noun strategy and how processing instruction might provide an effective intervention (Houston, 1997; Lee, 1987; LoCoco, 1987; VanPatten, 1984; VanPatten, 1996; VanPatten & Houston, 1998). This body of work has determined that learners of Spanish (L1 English) at the beginning level interpret the preposed pronoun as the subject both sentence-initially and sentence-internally up to 70% of the time (cf. Lee, 1987; VanPatten, 1984). Several factors attenuate the NNS's assumption that the first noun is subject, such as lack of plural 's' marker (i.e. *lo* vs. *los*) (Lee, 1987), background knowledge/topic familiarity (Houston, 1997), and the presence of contextual cues in the sentence (Malovrh, 2006; VanPatten and Houston, 1998). What these studies have in common is that they focus mostly on novice NNSs to control for the effect of prior knowledge.

Recent studies of intermediate and advanced NNS populations include Zyzik (2006) and Malovrh (2008). Zyzik (2006) studied the use of the dative pronouns *le* and *les* in accusative contexts in four oral speech formats across four groups of NNSs. Learners at the beginning level rarely used clitic pronouns and were excluded from the main results of the study. Zyzik found that overuse of the dative pronoun in accusative contexts increased from intermediate (22%) through high intermediate (37.6%) to advanced NNSs (46.2%). The use of *le(s)* for animates in this context increased as well, from 80% to 85.7% to 92.4%. In a related study, Malovrh (2008) examined use of the accusative pronouns *lo*, *la*, *los* and *las*. He found that initially, NNSs mapped *lo* as a direct object pronoun while *la*, *los* and *las* were mapped to their function as definite articles. This is precisely what is predicted by Anderson's (1984) one-to-one principle which claims that NNSs will associate one form with one meaning before making multiple form-meaning connections. NNSs at the second proficiency level began to map *los* onto its pronoun function while *la* and *las* continued to lag behind in on-line production tasks. *La* and *las* began to be mapped onto the pronoun function in the third proficiency group. Thus, Malovrh's work demonstrated that *lo* is mapped to its function as direct object before the other three pronouns. Although Zyzik and Malovrh acknowledge that *leísmo* may play a role in the acquisition of DO

pronouns by English-speaking L2 learners of Spanish, they were unable to make claims about the type of input that NNSs receive given the heterogeneity of their language learning experience. This further motivates the need to conduct research in contexts where language learning experience is more homogenous for the participant group, at least during the time of the study.

### 3. The current study

Given the preceding background information, the present study aims to answer the following research questions:

1. For learners participating in a SA immersion program, what changes in frequency of selection of the pronouns *lo(s)/la(s)* or *le(s)* are demonstrated over the course of this immersion experience?
2. What linguistic and extralinguistic variables predict this use over the course of this immersion experience?
3. How do these results compare to those of NSs in the same region?

#### 3.1. Participants

Data from 24 NSs and 33 NNSs were examined for the current investigation. The NNS participants were 33 high school students (12 male, 21 female; all aged 17) who were enrolled in a seven-week intensive summer SA program for high school honors students in León, Spain. Rather than relying on years of Spanish instruction as a measure of proficiency, a level test was administered immediately following arrival in Spain. Participants abided by a strict no-English policy while in country. Formal classroom instruction consisted of five daily courses (grammar, culture, conversation, literature, and phonetics/pronunciation) taught exclusively in Spanish by four instructors (2 NSs of Spanish and 2 advanced NNSs of Spanish). Students lived with Spanish host families and were encouraged to spend as much free time as possible with their families so as to increase their exposure to the language. Although the amount of contact hours with the target language varies individually, the nature of this program requires students to use Spanish for many more hours daily than most study abroad programs. Thus, despite the short duration, this program represents a highly contact-intensive experience, and we expect greater changes over the short period of immersion than we would in other programs where students often socialize with each other in English outside of class. The NSs who participated in the present study were a group of university students from León, Spain (N=24; 10 male and 14 female). Participants were recruited from the same school facility where the NNS participants took classes, and their ages ranged from 20-30. This second group was included in order to document the characteristics of the target variety to which NNSs were exposed.

#### 3.2. Elicitation tasks

All participants completed a background questionnaire and one or more written contextualized tasks (WCTs). There were two key differences between the NS and the NNS groups: the NNSs also completed a level test at the beginning and end of their stay abroad, and the NNSs completed a different version of the WCT at three different points in time (week 1, week 4 and week 7) during their stay abroad. In contrast, the NSs completed this task only once, using the Time 3 version of the instrument. Each of these instruments will be described below.

##### 3.2.1. Level test and background questionnaire

The level test, administered to NNSs as a pre-test immediately following their arrival in Spain and again as a post-test following completion of the final set of data elicitation instruments, was an 11-item multiple-choice instrument that covered a range of grammatical structures generally included in formal instruction of Spanish. This instrument was previously found to be a statistically reliable indicator of

proficiency (Woolsey, 2006). The background questionnaire asked participants to provide information regarding previous language experience, parents' level of education, and time spent traveling or living in another country. The background questionnaire enabled the examination of any potential effects of traditionally relevant sociolinguistic variables such as gender and socioeconomic class as well as the characterization of the demographic makeup of both groups.

### 3.2.2. *Written contextualized task*

The WCT for the current study was based on the WCT format used in Geeslin (2003) and Gudmestad (2006) to study the acquisition of Spanish copula choice and subjunctive, respectively. It was designed to ensure a balanced distribution of the combination of all categories of each of the independent variables included in the study: referent number; co-referentiality; referent gender; telicity; and subject animacy.<sup>1</sup> These independent linguistic variables were taken from the previous research on native speaker use of DO clitics and are summarized in Table 1.

Variable	Categories	Example	Criterion
Referent number	[+ singular]	<i>A Juan lo veo.</i>	Is the referent singular or plural?
	[- singular]	<i>A los chicos los veo.</i>	
Co-referentiality	[+ co-referential]	<i>A Juan lo veo.</i>	Is the referent given with a co-referential <i>a</i> + NP phrase?
	[- co-referential]	<i>Lo veo.</i>	
Referent gender	[+ male]	<i>A Juan lo veo.</i>	Is the referent male or female?
	[- male]	<i>A María la veo.</i>	
Telicity	[+ telic]	<i>A Juan lo veo.</i>	Does the predicate have endpoints (i.e., telic)?
	[- telic]	<i>A Juan lo atropello.</i>	
Subject animacy	[+ animate subject]	<i>A Juan lo veo.</i>	Is the subject animate or inanimate?
	[- animate subject]	<i>A Juan lo protege el casco.</i>	

There were three WCTs in total (i.e., one per Time). Each WCT was comprised of a dialogue between two speakers, and this created the discourse context for each item. Each item contained only animate DOs and was designed to represent a single combination of the categories of each of the independent variables included in the study.<sup>2</sup> Due to the length of the instrument, each possible combination of the categories of the independent linguistic variables appeared once only within the task. Each item contained two possible answers, differing only on the linguistic construct being tested. Learners read the discourse context and were instructed to indicate whether they preferred the first, second, or both options (see Figure 1). The coding for the sample item in Figure 1 is: [+singular]; [+co-referential]; [+male]; [+telic]; [+animate subject].

<sup>1</sup> For the current study, [+co-referential] implied that the noun phrase always preceded the verb (see sample item in Figure 1).

<sup>2</sup> This means that all items were possible contexts where *le* could be selected for an animate direct object. Given the fact that *leísmo* appears to be conditioned by several linguistic factors, however, categorical selection of *le* by NSs is not the anticipated result.

Manuel se da cuenta de que Inma no entiende la situación. Le habla sobre Vicente, un amigo que tienen en común.

Manuel: ¿Conoces a Vicente Moreno?

A. Inma: No, a Vicente Moreno no lo conozco yo.

B. Inma: No, a Vicente Moreno no le conozco yo.

Prefiero la frase A.

Prefiero la frase B.

Prefiero A y B.

Figure 1. Sample test item.

### 3.3. Procedure

The initial examination of the data involved the tabulation of each of the object pronoun forms selected at each test time. Following this analysis, the forms were collapsed as *le(s)* responses or *lo(s)/la(s)* responses, including singular and plural forms in the same categories. Additionally, the response category *both* was combined with responses that indicated the selection of non-prescriptive *le(s)*. The motivation for combining the *both* responses with the *le(s)* responses was to identify those contexts in which non-prescriptive use of *le(s)* was possible in contrast with those contexts where only the prescriptive direct object pronouns were permissible.<sup>3</sup> Once the dependent variable was collapsed, the frequency of selection of each category of the dependent variable (i.e., *le(s)* vs. no *le(s)*) was determined for each test time for the NNSs as well as for the NSs. These data were then submitted to chi-square ( $X^2$ ) tests in order to determine whether the changes across time and the differences between NSs and NNSs were statistically significant.

In addition to independent linguistic variables, the responses for each learner were coded for improvement on the level test as well as for individual characteristics such as experience with other languages and time studying Spanish. Four logistic regression analyses were performed to determine which independent variables were significant predictors in the selection of *le(s)*: one for the NSs, and three for the NNSs (i.e., Time 1, Time 2, Time 3). A regression analysis identifies the variables that best predict the selection of *le(s)* (i.e., the dependent variable) and those that do not enhance the predictive power of the model are excluded. All statistical tests were performed using the Statistical Package for the Social Sciences (SPSS) 16.0.

## 4. Results

As was previously mentioned, each participant completed a level test, a background questionnaire, and three different versions of the WCT. This section begins with a summary of the results of the level test. We then provide the results for the analysis of the frequency of forms across time for the NNSs and for the NSs, showing both the distribution of each clitic form across groups and then the same dataset with a binary dependent variable created by collapsing *le(s)* and *both* into one category and *lo(s)/la(s)* into the other. Subsequent analyses, such as the regression models used to identify the significant predictors of selection at each level, are then presented. The results for the NNSs and the NSs are discussed together throughout this section.

### 4.1. Level test

The results of the level test appear in Table 2. The information provided there summarizes the results for the L2 learners at two times: Time 1, right after the learners arrived in León; and Time 3, during the seventh week of the stay abroad. The results show an average improvement from a score of 8.24 to 9.76, and a chi-square test confirms that this difference is significant ( $X^2=17.991$ ,  $df=7$ , Cramer's  $V=0.522$ ,  $p=0.012$ ).

<sup>3</sup> See Geeslin (2003) for a similar argument in collapsing responses for copula choice WCT data.

	Time 1	Time 3
Average	8.24	9.76
Range	4-11	7-11
Standard Deviation	1.81	1.07

Note. Maximum score = 11.

#### 4.2. Frequency of selection of each clitic form

The methodology of the present study was designed to determine whether the frequency of DO clitic choice varied based on data elicitation time. Raw scores and percentages of frequency of clitic choice across three times for the NNSs and for the NSs were computed and are presented in Table 3. The data in Table 3 show that *le* and *les* are preferred by the NNSs at all times. However, the frequency of *le* and *les* diminishes at from Time 1 to Time 2 and increases again at Time 3.<sup>4</sup> In comparing the frequency of forms chosen by NNSs and NSs, the results indicate that by Time 2 the frequency of forms chosen by the NNSs approach those used by the NSs. That is, NNSs choose *le* 16.8% of the time, and the NSs choose *le* 25.5% of the time. As for the plural options, the NNSs choose *les* 19.6% of the time, and the NSs choose *les* 25.5% of the time.

	Time 1		Time 2		Time 3		NS	
	#	%	#	%	#	%	#	%
<i>Lo</i>	88	11.1	123	15.6	100	13.2	22	3.8
<i>Los</i>	79	10	107	13.5	88	11.6	33	5.8
<i>La</i>	83	10.5	109	13.8	113	14.9	104	18.2
<i>Las</i>	78	9.9	124	15.7	104	13.7	102	17.8
<i>Le</i>	205	25.9	133	16.8	161	21.2	146	25.5
<i>Les</i>	220	27.8	155	19.6	158	20.8	146	25.5
<i>Both</i>	38	4.8	39	4.9	35	4.6	19	3.3
Total	791		790		759		572	

In order to focus the contrast between *le(s)* forms and *lo(s)/la(s)* forms, it is helpful to have a binary dependent variable and, thus, the data were collapsed into two groups with the *le(s)* and *both* responses as one category and *lo(s)* and *la(s)* responses as another. This also meets the requirement for the additional statistical analyses that were conducted (see Section 4.3). The data in Table 4 show the NNS frequency of forms across time for the collapsed dependent variable as well as the same information for the NSs. The results indicate that although NNSs prefer *le(s)* at a rate of 58.6% at Time 1, they show an opposite pattern at Time 2, favoring *lo(s)/la(s)* forms at exactly the same rate (i.e., 58.6%). Finally, an increase towards *le(s)* is found at Time 3 (46.6%). As for NSs, they show a preference for *le(s)* (54.4%), compared to 45.6% for *lo(s)/la(s)*.

	Time 1		Time 2		Time 3		NS	
	#	%	#	%	#	%	#	%
<i>Lo(s)/la(s)</i>	328	41.5	463	58.6	405	53.4	261	45.6
<i>Le(s)</i>	463	58.6	327	41.4	354	46.6	311	54.4
Total	791		790		759		572	

<sup>4</sup> The totals are slightly different due to failure to respond a few items as well as attrition at Time 3.



#### 4.3. Predictors of use: the regression analysis

We turn to the results yielded by the logistic regression models in order to better understand the choice of DO clitic pronouns by the NNSs across time and also to observe how pronoun choice for NNSs compared to that of the NSs. Table 5 summarizes these results, showing the linguistic and extralinguistic variables that were selected by the statistical model as significant predictors of pronoun choice. An 'X' indicates that a variable was included in the regression, and an asterisk indicates the respective level of significance (additional statistical details are provided in the Appendix). Table 5 shows that at Time 1 there were two linguistic variables that predict NNS selection of *le(s)* (i.e., co-referentiality and telicity). At Time 2, only one linguistic variable (i.e., telicity) was included in the model. Finally, at Time 3 there were three linguistic variables included in the predictive model (i.e., co-referentiality, subject animacy, and telicity). The only non-linguistic predictor in any NNS model was improvement on the level test, which was included in the regression models at Times 2 and 3. These results further show that at Time 3 the NNS model was closer to the NS model. Specifically, the NNS model at Time 3 and the NS model included the factors subject animacy and predicate telicity. They differed, however, in co-referentiality (NNS model only) and referent gender (NS model only).

Variable	NNS Time 1	NNS Time 2	NNS Time 3	NS
Referent number				
Referent gender				X***
Co-referentiality	X***		X***	
Subject animacy			X***	X***
Telicity	X***	X***	X*	X***
Other language				
Other experience				
Improvement on level test		X*	X***	

Note. \*\*\* =  $p \leq 0.001$ ; \*\* =  $p \leq .001$ ; \* =  $p \leq 0.05$

#### 4.4. Summary of results

In sum, our longitudinal study demonstrated an initial rate of selection of the *le* forms that was similar to NS rates, but whose predictors were quite different. Over time, NNSs show a sharp decrease in selection of *le(s)* followed by a gradual increase toward the NS rates of selection. Likewise, the predictive models that represent the factors that predict the selection of *le(s)* more closely approximate the NS models, both in terms of the number of factors included in the model and the factors themselves, over time. The implications of these results will be discussed in greater depth in the section that follows.

### 5. Discussion

It will be recalled that the current study was designed to answer three research questions, the first of which asked how the frequency of selection of *le(s)* and *lo(s)/la(s)* changed over the course of the SA immersion program. Because the overarching interest in this structure stems from its connection to the SLA of variable structures, our discussion will focus on the frequency of selection of *le(s)* over time, as this indicates movement toward or away from the local norms. At Time 1, we observe near native-like rates of selection of *le(s)* (58.6% as compared to the NSs 54.4%). At Time 2, the frequency of selection of *le(s)* (41.4%) has fallen considerably from Time 1. Finally, at Time 3, we see an increase in rates of selection of *le(s)* (46.6%), but not to the levels of the NS norm. The higher frequency of selection at Time 1 may indicate that the learners' grammar is one in which the selection of *le(s)* is connected to animate objects of the verb, regardless of whether they are direct or indirect because the form *la* has already been acquired as the definite article (cf. Andersen, 1984). At Time 2,

we hypothesize that the decrease in frequency of *le(s)* indicates that learners are beginning to distinguish between direct and indirect objects according to prescriptive rules.<sup>5</sup> The final increase in frequency of selection of *le(s)* at Time 3 may indicate that learners are reaching a stage in which they are moving toward NS norms, regardless of the prescriptive rules of the language.

The second research question examined the predictors of the selection of the DO clitic forms across time, based on the results of the regression analysis. At Time 1 we see that the linguistic predictors of NNS selection of *le(s)* include the factors co-referentiality and telicity of the verb. At Time 2, the only linguistic factor included in the model is telicity. In addition, one non-linguistic factor, improvement on the grammar test, is added to the model. This implies that learners who are rapidly expanding their knowledge of Spanish on an objective measure of grammatical knowledge are also more likely to select *le(s)* than those who are making slower gains. Finally, at Time 3, the predictive model includes the linguistic factors co-referentiality, subject animacy, and telicity and the non-linguistic factor improvement on the grammar test. One notes from this analysis that telicity is an important predictor of the selection of *le(s)* at all test times.<sup>6</sup> Additionally, at Time 2, precisely when the selection of *le(s)* is least frequent, there are fewer factors included in the predictive model. This can be taken to indicate an increase in individual variability such that no single factor is a good predictor across speakers and across contexts. This is consistent with the results for frequency in that both findings indicate that this is the stage at which learners are developing a contrast in their own grammars between *le(s)* and *lo(s)/la(s)*.

The third and final research question sought to compare our NNS group to the NS group that completed the same task. This enables us to assess whether or not our learners are moving toward the NS norms. At Time 1, we see that the learners' frequency of selection of *le(s)* is similar to that of the NSs but the predictors of *le(s)* are quite different. In the NS model, the significant predictors of *le(s)* are referent gender, subject animacy, and telicity. Thus, the Time 1 model for the learners has only one factor, telicity, in common. We can conclude that at Time 1 our learners are not selecting *le(s)* in ways that are consistent with the NS norm. The differences between NSs and the NNSs become even more pronounced at Time 2, when both the frequency and the predictors differ between groups. The NNS frequency is much lower than that of the NSs and the predictive model includes only one factor in common. At Time 3, however, we see evidence of a shift toward the NSs norms in that the increase in the frequency of selection of *le(s)* coincides with an increase in the number of factors that the NS and NNS models have in common. At Time 3, we see that both telicity and subject animacy are included in both models and, thus, there is only one linguistic factor in each model that sets them apart from each other. Although we cannot separate general development on the part of the learner from development that is specific to this learning context without additional data from at-home learners, it is clear that our learners are moving toward native-like selection of *le(s)*.

Our results make two important contributions to the field. The first, described in detail in the preceding paragraphs is that we provide the first longitudinal study of DO pronouns in a SA environment where a clear account of the target to which NNSs are exposed is also provided. The result is a more detailed understanding of the path of acquisition in this learning environment than was previously available. The second contribution is that our results serve to extend the current understanding of the SLA of variable structures to a new grammatical construct. The u-shape curved in frequency of selection demonstrated in the current study is quite similar to that which has been found for the SLA of copula choice (see Geeslin (2003) for copula research and Ellis (1997) for a discussion of u-shaped development). In both cases, an initially over-generalized form first decreases in use, most

---

<sup>5</sup> The significant differences in the grammar test between the first and third test time, along with the fact that learners are spending time in the language classroom (i.e., receiving formal instruction) on this and other structures and have time to use such metalinguistic knowledge to complete the WCT suggest that this is a reasonable hypothesis.

<sup>6</sup> The results for telicity are not entirely surprising, given the predictions made by Flores-Cervantes (2002) that stated that verbs with lower telicity would favor selection of *le(s)*. Other factors, such as animacy of the subject (Garcia & Otheguy, 1977) and number of the referent (Klein-Andreu, 2000), which are also related to transitivity, show less universal effects because the former only enters into the predictive model at Time 3 and the latter is not included for any speaker group at any time.

likely when a one-to-one strategy fails to produce the range of meanings the NNS wishes to convey, and then increases more gradually toward the NS norms of frequency while at the same time increasing in terms of the appropriateness of the linguistic and social factors that predict such use. It is worth noting that while these results coincide with what is known about the SLA of copula choice, they contrast with research on the SLA of other variable structures such as the subjunctive mood (Gudmestad, 2006, 2008) and this is likely because the subjunctive is not overgeneralized at the beginning stages of acquisition and this structure takes longer to acquire. Through the addition of the current project to the existing database, new generalizations about the SLA of structures that are variable in the target are possible.

## 6. Limitations and future directions

This study has provided the first longitudinal study of the development of DO pronouns by English-speaking NNSs in a SA environment. It demonstrates the gradual progression from the overgeneralized selection of *le(s)* toward use, that while initially less frequent, becomes more target-like in terms of the predictors of the selection of *le(s)* across time. In so doing, our study builds on our existing knowledge of research conducted in the SA environment, research on DO pronouns and research on the SLA of variable structures in Spanish. Despite these contributions, there are some limitations that should be taken into account when designing future studies. Firstly, our findings are based on a single task that elicits preferences, rather than NNS-directed output. Thus, our results should be corroborated in the future using additional data elicitation tasks that provide a greater range from very constrained to very open-ended naturalistic speech. Secondly, while our research provides an account of what SLA looks like in a SA context, we are still unable to separate the effects of one target variety on acquisition from the effects of another. In the future, this study should be replicated in other learning environments where the degree of contact with the target is similar but the variety spoken in that environment differs from the one spoken in León, Spain. This will allow us to begin the process of comparing the SLA of one variety to the SLA of another. In conclusion, we hope that our project will be the first of many on this topic with the ultimate goal of understanding better the process of the second language acquisition of many different structures in many different varieties of Spanish.

## Appendix

Details from the regression analyses				
	NNS Time 1	NNS Time 2	NNS Time 3	NS
Percent predicted	61.1	60.6	60.1	76.7
Model $\chi^2$	31.961***	24.789***	51.953***	239.677***
Df	2	4	6	3
-2 log likelihood	1041.444	1046.854	996.815	548.906
Nagelkerke R <sup>2</sup>	.053	.042	.088	0.458

Note. \*\*\* =  $p \leq 0.001$ ; \*\* =  $p \leq .001$ ; \* =  $p \leq 0.05$

## References

- Adamson, H. Douglas & Regan, Vera. (1991). The acquisition of community speech norms by Asian immigrants learning English as a second language. *Studies in Second Language Acquisition*, 13, 1-22.
- Anderson, Roger. (1984). What's gender good for anyway? In Roger Anderson (Ed.), *Second languages: A cross-linguistic perspective* (pp. 77-99). Rowley, MA: Newbury House.
- Canale, Michael & Swain, Merrill. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Carroll, John B. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals*, 1, 131-51.

- Collentine, Joseph. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26, 227-248.
- Collentine, Joseph & Freed, Barbara. (2004). Learning context and its effects on second language acquisition: Introduction. *Studies in Second Language Acquisition*, 26, 153-172.
- DeKeyser, Robert. (1986). *From learning to acquisition? Foreign language development in a US classroom and during a semester abroad*. Unpublished doctoral dissertation, Stanford University.
- DeKeyser, Robert. (1991). Foreign language development during a semester abroad. In Barbara F. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 104-119). Lexington, MA: D.C. Heath.
- Díaz-Campos, Manuel. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, 26, 249-273.
- Díaz-Campos, Manuel & Lazar, Nicole (2003). Acoustic analysis of voiceless initial stops in the speech of study abroad and regular class students: Context of learning as a variable in Spanish second language acquisition. In Paula Kempchinsky & Carlos Eduardo Piñeros (Eds.), *Theory, practice and acquisition: Papers from the 6th Hispanic Linguistics Symposium* (pp. 352-370). Somerville, MA: Cascadilla Press.
- Ellis, Rod. (1997). *Second language acquisition*. Oxford: Oxford University Press.
- Flores Cervantes, Marcela. (2002). Leísmo, laísmo y loísmo: sus orígenes y evolución. México, D.F.: CONACULT-INAH, Universidad Autónoma de México.
- Freed, Barbara F. (1995). What makes us think that students who study abroad become fluent? In Barbara F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp.123-148). Philadelphia: John Benjamins.
- García, Erica & Otheguy, Ricardo. (1977). Dialect variation in leísmo: A semantic approach. In Ralph W. Fasold & Roger W. Shuy (Eds.), *Studies in language variation* (pp. 65-87). Washington, DC: Georgetown University Press.
- García-Amaya, Lorenzo. (2008). The effect of topic on rate of speech. *ELLA (Estudios de Lingüística Inglesa Aplicada)*, 8, 117-150.
- García-Amaya, Lorenzo. (2009). New findings on fluency measures across three different learning contexts. In Joseph Collentine, Maryellen García, Barbara Lafford, & Francisco Marcos Marín (Eds.), *Selected Proceedings of the 11<sup>th</sup> Hispanic Linguistics Symposium* (pp. 68-80). Somerville, MA: Cascadilla Press.
- Geeslin, Kimberly. (2003). A comparison of copula choice in advanced and native Spanish. *Language Learning*, 53(4), 703-764.
- Geeslin, Kimberly & Gudmestad, Aarnes. (2008a). Comparing interview and written elicitation task in native and non-native data: Do speakers do what we think they do? In Joyce Bruhn de Garavito & Elena Valenzuela (Eds.), *Selected proceedings of the 10<sup>th</sup> Hispanic Linguistics Symposium* (pp.64-77). Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, Kimberly & Gudmestad, Aarnes. (2008b). Variable subject expression in second-language Spanish: A comparison of native and non-native speakers. In Melissa Bowles, Rebecca Foote, Silvia Perpiñán, & Rakesh Bhatt (Eds.) *Selected proceedings of the Second Language Research Forum* (pp. 69-85). Somerville, MA: Cascadilla Press.
- Geeslin, Kimberly & Gudmestad, Aarnes. (2008c). The acquisition of a second dialect in second language Spanish? Where shall we begin? Paper presented at Current Approaches to Spanish and Portuguese Second Language Phonology, Minneapolis, Minnesota.
- Geeslin, Kimberly & Guijarro-Fuentes, Pedro. (2006). The second language acquisition of variable structures in Spanish by Portuguese speakers. *Language learning*, 56(1), 53-107.
- Gudmestad, Aarnes. (2006). L2 variation and the Spanish subjunctive: Linguistic features predicting mood selection. In Carol Klee & Timothy Face (Eds.), *Selected proceedings of the 7<sup>th</sup> Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 170-184). Somerville, MA: Cascadilla Proceedings Project.
- Gudmestad, Aarnes. (2008). *Acquiring a variable structure: An interlanguage analysis of second-language mood use in Spanish*. Unpublished doctoral dissertation. Indiana University, Bloomington.
- Gudmestad, Aarnes & Geeslin, Kimberly. (2007). A closer look at advanced learners and sociolinguistic variation: Spanish-language use in contexts of future time reference. Paper presented at Hispanic Linguistics Symposium, San Antonio, Texas.
- Hopper, Paul & Thompson, Sandra. (1980). Transitivity in grammar and discourse. *Language*, 45, 251-299.

- Houston, Tony. (1997). Sentence processing in Spanish as a second language: A study of word order and background knowledge. In William R. Glass & Ana Teresa Pérez-Leroux (Eds.), *Contemporary perspectives on the acquisition of Spanish: Vol. 2. Production, processing and comprehension* (pp. 123-134). Somerville, MA: Cascadilla Press.
- Isabelli, Casilde & Nishida, Chiyo (2005). Development of Spanish subjunctive in a nine-month study-abroad setting. In David Eddington (Ed.), *Selected proceedings of the 6<sup>th</sup> Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 78-91). Somerville, MA: Cascadilla Press.
- Klein-Andreu, Flora. (2000). *Variación actual y evolución histórica: los clíticos le/s, la/s, lo/s*. Munich: Lincom Europa.
- Lafford, Barbara. (2004). The effect of context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition*, 26, 201-26.
- Lafford, Barbara & Collentine, Joseph. (2006). The effects of study abroad and classroom contexts on the acquisition of Spanish as a second language: From research to application. In Rafael Salaberry, & Barbara Lafford (Eds.), *The art of teaching Spanish: Second language acquisition from research to praxis* (pp. 103-126). Washington, D.C.: Georgetown University Press.
- Lee, James. (1987). Morphological factors influencing pronominal reference assignment by learners of Spanish. In Terrell Morgan, James Lee, & Bill VanPatten (Eds.), *Language and language use: Studies in Spanish* (pp. 221-232.) Lanham, MD: University Press of America.
- LoCoco, Vivian. (1987). Learner comprehension of oral and written sentences in German and Spanish: The importance of word order. In Bill VanPatten, Trisha Dvorak, & James Lee (Eds.), *Foreign language learning: A research perspective* (pp. 119-129). Rowley, MA: Newbury House.
- Lord, Gillian. (2009). Language awareness and development during study abroad: A case study. *Hispania*, 92, 136-150.
- Malovrh, Paul. (2006). L2 Sentence Processing of Spanish OVS Word Order and Direct Object Pronouns: An Analysis of Contextual Constraints. In Nuria Sagarra & Almeida Jacqueline Toribio (Eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium* (pp. 169-179). Somerville, MA: Cascadilla Proceedings Project.
- Malovrh, Paul. (2008). *A multifaceted analysis of the interlanguage development of Spanish direct-object clitic pronouns observed in L2-learner production*. Unpublished doctoral dissertation. Indiana University, Bloomington.
- Mougeon, Raymond & Dewaele, Jean-Marc. (2004). Preface. *IRAL*, 42, 295-301[Special issue: *Variation in the interlanguage of advanced second language learners*].
- Mougeon, Raymond, Rehner, Katherine & Nadasdi, Terry. (2004). The learning of spoken French variation by immersion students from Toronto, Canada. *Journal of Sociolinguistics*, 8(3), 408-432.
- Preston, Dennis. (2000). Three kinds of sociolinguistics and SLA: A psycholinguistic perspective. In Bonnie Swierzbins, Frank Morris, Michael Anderson, Carol Klee, and Elaine Tarone (Eds.), *Social and cognitive factors in second language acquisition: Selected proceedings of the 1999 Second Language Research Forum* (pp. 3-30). Somerville, MA: Cascadilla Press.
- Rehner, Katherine. (2002). *The development of aspects of linguistic and discourse competence by advanced second language learners of French*. Unpublished doctoral dissertation, OISE/University of Toronto.
- Segalowitz, Norman & Freed, Barbara (2004). Context, contact, and cognition in oral fluency acquisition: learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173-199.
- Segalowitz, Norman, Freed, Barbara, Collentine, Joseph, Lafford, Barbara, Lazar, Nicole, & Díaz-Campos, Manuel (2004). A comparison of Spanish second language acquisition in two different learning contexts: study abroad and the domestic classroom. *Frontiers*, 10, 21-38.
- Simões, Antonio R.M. (1996). Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. *Hispania*, 79, 87-95.
- Stevens, John J. (2001). *The acquisition of L2 Spanish pronunciation in a study abroad context*. Unpublished doctoral dissertation, University of Southern California.
- Tarone, Elaine. (2008). Sociolinguistic approaches to second language acquisition-1997-2007. *Modern Language Journal*, 95(5), 837-848.
- VanPatten, Bill. (1984). Learners' comprehension of clitic pronouns: More evidence for a word order strategy. *Hispanic Linguistics*, 1, 57-67.
- VanPatten, Bill. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.

- VanPatten, Bill & Houston, Tony (1998). Contextual effects in processing L2 input sentences. *Spanish Applied Linguistics*, 2, 53-70.
- Woolsey, Daniel. (2006). *Second language acquisition of the verb 'estar' with adjectives: An exploration of contexts of comparison and immediate experience*. Unpublished doctoral dissertation, Indiana University, Bloomington.
- Zyzik, Eve. (2006). Learners' Overgeneralization of Dative Clitics to Accusative Contexts: Evidence for Prototype Effects in SLA. In Carol A. Klee and Timothy L. Face (Eds.), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 122-134). Somerville, MA: Cascadia Proceedings Project.

# Selected Proceedings of the 12th Hispanic Linguistics Symposium

edited by Claudia Borgonovo,  
Manuel Español-Echevarría,  
and Philippe Prévost

Cascadilla Proceedings Project Somerville, MA 2010

## Copyright information

Selected Proceedings of the 12th Hispanic Linguistics Symposium  
© 2010 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-440-9 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Geeslin, Kimberly L., Lorenzo J. García-Amaya, Maria Hasler-Barker, Nicholas C. Henriksen, and Jason Killam. 2010. The SLA of Direct Object Pronouns in a Study Abroad Immersion Environment Where Use Is Variable. In *Selected Proceedings of the 12th Hispanic Linguistics Symposium*, ed. Claudia Borgonovo, Manuel Español-Echevarría, and Philippe Prévost, 246-259. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #2421.