

Lexical Frequency Effects on L2 Spanish Subject Pronoun Expression

Bret Linford and Naomi Lapidus Shin

1. Introduction

Investigations of frequency of linguistic forms have been central to usage-based models of grammar (e.g. Bybee 2001, 2010) and have advanced our understanding of developmental processes during language acquisition. Frequent forms are acquired earlier and processed more easily than infrequent forms (Ellis 2012, Lieven 2010). In addition to influencing the order in which forms emerge, frequency also plays a role in determining the distribution of forms in discourse. Learners first associate a frequent, prototypical form with a particular usage. Later they begin to use related, low-frequency forms in ways that match the pattern of the prototype (Ellis & Collins 2009). The outcome of this process is a probabilistically constructed grammar (Bybee 2010, Ellis 1996, Lieven 2010).

One example of how such usage-based models correctly predict developmental patterns comes from research on second language (L2) acquisition of the preterit and imperfect past tenses. Learners follow a trajectory whereby stative verbs, when used in the past tense, are paired with imperfective forms. Achievement verbs, on the other hand, are paired with preterit forms (Andersen 1986, Bardovi-Harlig 1998, Cadierno 2000, Shirai & Anderson 1995).¹ For instance, learners of Spanish typically produce the stative verb *estar* as imperfect *estaba*, but rarely as preterit *estuve*. Likewise, they produce achievement verb *darse cuenta* as preterit *me di cuenta*, but rarely as imperfect *me daba cuenta*. The evidence for this developmental pattern is in the input, as adult native speakers frequently pair stative verbs with imperfective forms and achievement verbs with preterit forms (Andersen & Shirai 1994). Furthermore, learners rely on the evidence established between the most frequent forms and their most common usage (such as *estar* conjugated as *estaba*), and they associate other members of the same lexical class (in this case all stative verbs) with the usage patterns of the prototype (Shirai & Anderson 1995).

While L2 research has clearly shown that high frequency forms are learned earlier than low frequency forms, there is still much to be discovered regarding the impact of lexical frequency on the processes underlying how grammar is learned during L2 development. For example, are there similarities between L2 learners and native speakers with respect to how frequency impacts grammar? Do frequency effects change depending on proficiency level? In the current study we further analyze the role of frequency in second language development by focusing on distributional patterns of one of the most widely studied features of Spanish, that is, subject pronoun expression (*yo bailo ~ bailo*). Our findings indicate that frequency impacts pronoun expression differently depending on learners' proficiency level. During the early stages of Spanish language acquisition, learners express pronouns significantly more often with frequent than with infrequent verbs. For more advanced learners, however, frequency does not directly influence pronoun expression; instead, it mediates other linguistic variables that impact pronoun use.

* Bret Linford, Indiana University. Naomi L. Shin, University of New Mexico. For questions, contact Bret Linford at blinford@indiana.edu. We would like to thank two anonymous reviewers, Melissa Whatley, and the editors of this volume for their helpful comments on this paper. We are also grateful to the University of Montana students that participated in this study.

¹ Salaberry (2000) proposes that before L2 learners of Spanish use the imperfect and preterit to mark aspect, the preterit is used as a default marker of past tense.

2. Variable subject pronoun expression in L2 Spanish

An integral part of successful second language acquisition is the ability to use structures in ways that match patterns found among native speakers, including patterns of use that are variable and probabilistically determined (Canale & Swain 1980, Geeslin 2011). Studies of L2 Spanish have examined the variable use of several structures, including *ser/estar*, mood, future tense, as well as subject expression (e.g. Geeslin 2006; Geeslin & Gudmestad 2008, 2011, 2012; Gudmestad 2012; Gudmestad & Geeslin 2007, 2010; Linford 2009; Lubbers Quesada 1998). With respect to subject expression, previous research has shown that many of the linguistic factors found to be significant predictors of native speakers' pronoun expression patterns have been shown to affect L2 learners' patterns as well. Some broad generalizations based on these studies are presented here, where studies of native speakers² are listed after 'NS' and studies of L2 are listed after 'L2.'

Subject personal pronouns are expressed more often:

- (1) with singular grammatical persons rather than with plural ones (NS: Bayley & Pease-Álvarez 1996, 1997; Erker & Guy 2012; Flores-Ferrán 2002; Otheguy & Zentella 2012. L2: Geeslin & Gudmestad 2008; Linford 2009).
- (2) with verb forms that have ambiguous person morphology,³ such as the imperfect, as compared to forms that are unambiguous, such as the preterit (NS: Bayley & Pease-Álvarez 1996, 1997; Cameron 1994; Erker 2005; Hochberg 1986; Holmquist 2012; Otheguy & Zentella 2012. L2: Gudmestad & Geeslin 2010, Geeslin, Linford, & Fafulas 2012).
- (3) with mental/estimative verbs, such as *creer*, less often with stative verbs, such as *estar*, and least often with external activity verbs, such as *bailar* (NS: Bentivoglio 1987, Enríquez 1984, Erker & Guy 2012, Otheguy & Zentella 2012, Travis & Torres Cacoullos 2012).
- (4) when switching reference than when maintaining reference (NS: Bayley & Pease-Álvarez 1996, 1997; Bentivoglio 1987; Cameron 1995; Flores-Ferrán 2002; Otheguy & Zentella 2012; Shin & Otheguy 2009. L2: Geeslin, Linford, & Fafulas 2012; Linford 2009, forthcoming).
- (5) when the previous mention of the referent in subject position is also expressed (NS: Abreu 2012, Cameron 1994, Cameron & Flores-Ferrán 2004, Flores-Ferrán 2005, Torres Cacoullos & Travis 2011, Travis 2007. L2: Abreu 2012, Geeslin & Linford 2012).

Other relevant variables include morphological regularity (Erker & Guy 2012), specificity of the referent (Cameron 1995, Lapidus & Otheguy 2005), as well as various extralinguistic factors such as speaker gender and socioeconomic status (Otheguy & Zentella 2012, Shin & Otheguy 2013).

While many of the variables described above impact pronoun usage among both native speakers and L2 learners, even among the most advanced learners there are still some patterns that do not match those of native speakers. For example, studies have found that, in oral production, advanced learners omit subjects at higher rates than native speakers do (e.g. Linford 2009; Geeslin & Gudmestad 2008, 2010). Experimental studies also indicate that sensitivity to pragmatic and discourse constraints on pronoun expression takes a long time to develop during L2 acquisition of Spanish (Montrul & Rodríguez-Louro 2006; Rothman 2007, 2010). While these previous studies show that learners' patterns of subject use are impacted by multiple linguistic variables, no previous research has been conducted on the influence of lexical frequency on L2 Spanish subject pronoun expression.

2.1. Lexical frequency and Spanish subject pronoun expression

To date, the few studies that have examined the role of frequency on subject pronoun expression have yielded conflicting results. In a study of Spanish in New York, Erker & Guy (2012) found that

² Included in the lists of NS studies are studies of monolinguals, as well as studies of bilinguals. Monolingual and bilingual native speakers of Spanish are similar with respect to patterns of pronoun expression (Otheguy & Zentella 2012, Torres Cacoullos & Travis 2010).

³ See Bayley & Pease-Álvarez 1997:363, Silva-Corvalán, 2001:161-163, and Torres Cacoullos & Travis 2011:253 for an alternative explanation regarding the impact of TMA on pronoun expression.

lexical frequency did not, by itself, impact subject expression. That is, highly frequent verb forms like *creo*, *sé*, and *digo* were no more or less likely to occur with an expressed subject than low frequency forms like *crezca*. Instead, frequency acted as a mediator of other predictor variables: Morphological regularity, Semantic class, and Person significantly predicted pronoun expression among the most frequent verb forms, but not among the less frequent ones. In addition, some predictor variables, such as Switch-reference and TMA, significantly impacted both frequent and infrequent verb forms, but had a stronger impact on the frequent forms. The authors conclude that lexical frequency functions as an activator or amplifier of linguistic predictors of pronoun expression. However, another recent study presents evidence against such a role of frequency. Bayley, Holland & Ware (2013) examined oral interviews gathered from 29 Spanish speakers from California and Texas and found that frequency had an independent impact on subject expression: infrequent verbs slightly favored expression, while frequent verbs slightly disfavored expression. Also, in the majority of cases, frequency had neither an amplifying nor activating effect on the other linguistic variables included in the analysis. As discussed in the conclusion of Bayley et al. (2013), the conflicting results in these two studies underscore the need for more studies of frequency effects on variable syntactic structures. The current study adds to this growing body of research by examining the impact of lexical frequency on subject pronoun expression in L2 Spanish.

3. The current study

In the current study we ask the following research questions:

1. Does lexical frequency of verb forms directly influence second language learners' Spanish subject pronoun expression?
2. Does lexical frequency of verb forms mediate other linguistic variables that impact second language learners' Spanish subject pronoun expression?
3. Does lexical frequency of verb forms affect L2 learners differently depending on their language proficiency?

3.1. Participants and tasks

The participants in this study were 12 native speakers of English, recruited and interviewed by the primary investigator. All were undergraduate university students learning Spanish as a second language and were enrolled in either a second- or fourth-year Spanish class at the time the study took place. Two tasks, a grammar test and a semi-directed sociolinguistic interview, were completed in a quiet office the university. The goal of the grammar test was simply to ensure that second- and fourth-year students differed significantly in terms of proficiency levels. This test consisted of 25 multiple-choice items examining various grammatical structures. Participants had to choose among three possible options to complete sentences that were embedded into narratives. The following is an example of a test item:

1. *Creo que es muy interesante _____ de los hábitos alimenticios de la gente.*
 a. *hablo* b. *hablar* c. *hablando*

The highest possible score on the test was 25 (1 point per item). The second task was the semi-directed, digitally-recorded, sociolinguistic interview. Questions for the interview were designed to elicit a) sociodemographic information, such as age, place of origin, experience with the Spanish language, as well as b) longer stretches of narratives about important memories, future plans, and opinions about social and political topics. The interviews lasted approximately twenty minutes. All interviews were transcribed by the primary investigator.

The participants were categorized into two proficiency levels based on Spanish course enrollment, with Level 1 participants enrolled in second-year Spanish courses, and Level 2 participants in fourth-year courses. The division of the participants into these two groups was further justified by the grammar test scores, which were significantly different for Levels 1 and 2 [$t(10) = 8.12$, $p < .001$].

Another clear difference between the two levels was that all the Level 2 students had spent at least five months living either in Mexico or Spain, while the Level 1 learners had spent limited experience abroad. Table 1 presents details for each group.

Table 1. Participants by proficiency levels

	Level 1	Level 2
N participants	8	4
Year of university study	2nd year	4th year
Mean test score	41% (s = 9.2)	82% (s = 5.2)
Range of test scores	28% - 52%	76% - 88%
Gender	4 men, 4 women	2 men, 2 women
Average age	23 years	22 years
Average time abroad	2.1 months	6.3 months

3.2. Data and coding

All finite verbs that occurred with a subject pronoun (but could have occurred without one), or that were occurred without a pronoun (but could have occurred with one), were extracted for analysis. In other words, only *variable* contexts were included, that is, contexts where expression or omission of a pronoun is possible. (For a detailed discussion of the envelope of variation, see Otheguy & Zentella 2012). A total number of 980 tokens occurred in such variable contexts and thus were included in the study. Each verb was then coded for the following four linguistic variables, drawing from methods outlined in Erker & Guy (2012) and Otheguy & Zentella (2012):

1. *Person* (2 levels: 1sg, 3sg): Only first and third person singular verbs (765 out of a total of 980 tokens, i.e. 80 percent) were included in the analyses of this variable because several participants employed a restricted range of grammatical persons resulting in insufficient examples of 2sg, 1pl, and 3pl verb forms to run statistical analyses. Given that the two categories included, 1sg and 3sg,⁴ only differ in person, not number, we label this variable ‘Person.’
2. *Switch-reference* (2 levels: no-switch, switch): This variable refers to whether the subject of the verb is the same as, or different from, the subject of the previous verb. When the referent is different, the verb is coded as ‘switch,’ as in *Él se cayó y después ella gritó*. When the referent is the same across two subjects, the verb is coded as ‘no-switch,’ as in *Él se cayó y después (él) gritó*. All 980 tokens were included in the analyses of this variable.
3. *Semantic class of verb* (3 levels: mental activity, state, external activity): All 980 tokens were included in the analyses of this variable. The three categories of Semantic class are the following:
 - a. Mental activity verbs: e.g. *saber, creer, suponer* and *recordar*
 - b. Stative verbs: e.g. *estar, ser, tener*
 - c. External activity verbs: Verbs that describe a physical, social, or behavioral activity, e.g. *jugar, salir*
4. *Tense, mood, aspect (TMA)* (3 levels: present, preterit, imperfect): Only three TMA types were included in the analyses of this variable: simple present indicative (e.g. *canto*), preterit (e.g. *canté*), and imperfect indicative (e.g. *cantaba*). These three TMA forms represented 94 percent of the corpus (923 tokens).

In addition to the linguistic variables described above, each verb token was coded for lexical frequency. Two different measures of frequency were used in order to broaden our understanding of how frequency functions in second language acquisition.

5. *Local lexical frequency* (binary): Following Erker & Guy (2012:536), this measure was operationalized by determining the relative lexical frequency of the verb tokens within the corpus itself (including all the interviews). Based on the Zipfian distribution,⁵ a verb token was coded as

⁴ Our 3sg category consists almost entirely of *él* and *ella*. There was only one case of *usted*. The indefinite pronoun *uno* was not produced at all.

⁵ The Zipfian distribution (Zipf 1935) demonstrates that more frequent linguistic forms are exponentially more frequent than forms that are less frequent.

frequent if it represented at least one percent of the verb tokens in the corpus. All other verb tokens were coded as infrequent. For example, in the current study, *sé* was categorized as frequent because it occurred 76 times in the corpus (over seven percent of the verb tokens), while *sabía* was considered infrequent because it was only produced twice, and therefore represented less than one percent of the corpus.

6. *Repetition* (binary): In addition to the measure of frequency described above, we wanted to know if second language learners might be sensitive to a temporary, in-the-moment, frequency effect. We hypothesize that learners' own repetition of a verb can simulate a frequency effect. The idea is that repetition of a verb increases the strength of the mental representation of that verb, and, consequently, decreases processing difficulties associated with its production (see Ellis 2012:2). Once a verb is repeated and processing difficulties are reduced, the learner can entertain the various linguistic predictors that affect subject pronoun expression, such as Switch-reference. To test this hypothesis, each verb token was categorized into one of two groups based on how many times it was previously produced by the speaker in the interview. Verb tokens categorized into group '0' were tokens that had not been previously produced in the interview (hence, zero repetitions). If it was the third time or more that the speaker produced the token, then that token was categorized as '3+'. For example, the first time a speaker produced the verb token *sé*, it was coded as '0', while the third production or more of the token *sé* was placed in the 3+ group.⁶ A total of 176 verb tokens did not fit into either of these two categories, leaving a total number of 804 tokens to be analyzed for this measure (about 82 percent of the corpus).

3.3. Statistical analyses

For both levels of L2 proficiency we first examine overall rates of pronoun expression. Then we investigate the direct impact of Local lexical frequency and Repetition on pronoun expression via bivariate (chi-square) and multivariate (binary logistic regression) analyses. Finally, we use bivariate analyses (chi-square) to assess whether or not Local lexical frequency had a mediating effect on the other linguistic predictors of pronoun use among the advanced learners.

4. Results

The results are divided into three sections. Overall rates of subject pronoun expression by level of L2 proficiency are presented in Section 4.1. The results for both bivariate and multivariate analyses exploring whether lexical frequency has a direct impact on pronoun expression are presented in Section 4.2. Lastly, Section 4.3 hones in on the advanced learners and presents bivariate analyses examining interactions between Local lexical frequency and the other linguistic predictor variables.

4.1. Subject pronoun expression by L2 proficiency level

The distribution of subject pronoun expression and omission by proficiency level are displayed in Table 2. The columns labeled 'N' present the number of verb tokens in each context, while the columns labeled '%' show the percent of the total these tokens represent for each level of proficiency.

Table 2. Subject pronoun expression by proficiency level

L2 Level	Pronoun unexpressed		Pronoun expressed		Total
	N	%	N	%	
1	317	72	123	28	440 (100%)
2	472	87	68	13	540 (100%)

⁶ We focus on a) no previous mentions of the verb versus b) two or more previous mentions in order to clearly categorize tokens as infrequent (produced for the first time) versus frequent (produced for the third time or more).

As illustrated in Table 2, Level 1 learners expressed pronouns at a higher rate (28 percent) than Level 2 learners (13 percent). In terms of pronoun rates, our Level 2 learners closely resemble the advanced L2 speakers in Geeslin & Gudmestad (2008), who expressed subject pronouns at a rate of 11 percent.⁷ It is important to note that in Mexico and Spain, the two countries where Level 2 learners had spent time, the pronoun rate tends to be approximately 20 percent (Cameron 1994:31; Cantero Sandoval 1978:261, footnote 2; Enríquez 1984; Otheguy & Zentella 2012:72; Shin 2012:134; Shin & Otheguy 2013; Solomon 1999). This finding, that subject pronoun omission rates increase with Spanish proficiency and, at advanced levels, surpass native speakers' tendencies, corroborates previous research on L2 development of subject expression (e.g. Geeslin & Gudmestad 2008, Pérez-Leroux & Glass 1999).

4.2. Lexical frequency as a direct influence on pronoun expression

In this section we explore Research Question 1, that is, whether or not frequency has a direct impact on pronoun expression. Table 3 shows the percent of pronoun expression for each proficiency level with the verbs divided into categories of Local lexical frequency (infrequent, frequent), as well as Repetition (0, 3+). Here the columns labeled 'N' present the total number of verb tokens in each context, while the columns labeled '% pro' show the percent of *expressed* subject pronouns in each context. For example, according to the Local lexical frequency measure, Level 1 learners produced a total of 62 infrequent verbs, 23 percent of which occurred with an expressed pronoun.

Table 3. Subject pronoun expression rates by frequency (two proficiency levels)

L2 Level	Local lexical frequency				Repetition frequency			
	infrequent		frequent		0		3+	
	N	% pro	N	% pro	N	% pro	N	% pro
1	62	23	61	37	63	24	37	42
2	34	11	34	15	39	14	20	12

For Level 1 learners both measures of frequency had a significant impact on pronoun expression; they expressed more pronouns with the most frequent verbs according to the Local lexical frequency measure [$\chi^2(1)=11.09$, $p<.01$], as well as the Repetition measure [$\chi^2(1)=10.41$, $p=.001$]. For Level 2 learners, however, verb form frequency was not significant [Local lexical: $\chi^2(1)=1.57$, $p=.26$; Repetition: $\chi^2(1)=.39$, $p=.53$]. Thus the results indicate that at earlier stages of L2 acquisition, lexical frequency has a direct impact on pronoun expression such that pronouns are expressed more often with frequent verb forms than with infrequent ones. This effect dissipates in later stages of language development.

Now we turn to binary logistic regression analyses conducted in order to further investigate the direct impact of frequency on pronoun expression. For each proficiency level we performed two separate regressions in which the dependent variable was an expressed versus omitted subject pronoun. Included in each regression were five independent variables: Person, Switch-reference, TMA, Semantic class, as well as one measure of frequency; in one analysis the frequency was Local lexical frequency, in the other it was Repetition. Note that these analyses were restricted to verb tokens that satisfied the categories for all five independent variables. Table 4a presents the results for the analysis that included Local lexical frequency, and Table 4b presents the results for the analysis that included the Repetition measure. The column with the title *p* tells us whether or not the each variable reaches statistical significance as a predictor of pronoun use, with two asterisks indicating significance at the $<.01$ cut-off value, and one asterisk indicating significance at the $<.05$ cut-off value, while *ns* means 'nonsignificant.'

⁷ Geeslin & Gudmestad (2008) present the percentages of pronoun expression including all subject forms (e.g. lexical noun phrases, demonstrative pronouns, etc.) of both human and non-human referents.

Table 4a. Predictors of subject pronoun use including Local lexical frequency, Level 1 and Level 2 learners

Level 1 (N Vbs = 303), $R^2 = .08$			Level 2 (N Vbs = 288), $R^2 = .21$		
Variables	Wald	<i>p</i>	Variables	Wald	<i>p</i>
1. Local lexical freq.	6.36	*	1. Switch Reference	9.96	**
2. Switch Reference	6.31	*	2. Person	7.69	**
3. TMA	4.43	ns	3. Semantic Class	5.43	ns
4. Semantic Class	2.42	ns	4. TMA	4.16	ns
5. Person	.24	ns	5. Local lexical freq.	1.21	ns

Table 4b. Predictors of subject pronoun use including Repetition, Level 1 and Level 2 learners

Level 1 (N Vbs = 303), $R^2 = .07$			Level 2 (N Vbs = 288), $R^2 = .21$		
Variables	Wald	<i>p</i>	Variables	Wald	<i>p</i>
1. Switch Reference	5.05	*	1. Switch Reference	10.43	**
2. Repetition	4.71	*	2. Person	7.02	**
3. TMA	4.26	ns	3. Semantic Class	5.99	*
4. Semantic Class	1.03	ns	4. TMA	4.07	ns
5. Person	.40	ns	5. Repetition	.08	ns

For Level 1, Switch-reference and both frequency measures significantly predicted pronoun use. No other variables were significant. For Level 2, Switch-reference and Person were significant in both analyses, and Semantic class was significant in the analysis that included Repetition. Neither Local lexical frequency nor Repetition significantly impacted pronoun expression among Level 2 learners. The results here present two important findings. First, as proficiency in Spanish increases, learners become increasingly sensitive to the various linguistic factors that influence pronoun expression among native speakers. This is evidenced by the fact that Level 2 learners were sensitive to more factors than Level 1 learners. Second, and most relevant to the primary focus of the current paper, the results underscore the finding that lexical frequency directly impacts pronoun expression during the early stages of L2 development, but at more advanced stages it does not have this effect.

To gain purchase on the relationship between frequency and pronoun expression among Level 1 learners, and the lack of such a relationship among Level 2 learners, we present in Tables 5 and 6 each level's three most commonly produced verb forms. Also included are raw frequencies (N) of each verb, percentages showing how much of the corpus (per level) each verb represents (%), and the percent of expressed subject pronouns with each verb form (% pro).

Table 5. Most frequent verb forms for Level 1

Level 1			
Form	N	%	% pro
<i>tengo</i>	24	6.1	54
<i>tiene</i>	20	5.1	30
<i>sé</i>	19	4.8	37

Table 6. Most frequent verb forms for Level 2

Level 2			
Form	N	%	% pro
<i>sé</i>	57	10.6	0
<i>creo</i>	24	4.4	17
<i>tenía</i>	22	4.1	14

Level 1 learners' most frequent verb form was *tengo*, which was produced 24 times. Furthermore, this form occurred with a pronoun (*yo tengo*) 13 times, and without a pronoun (*tengo*) 11 times. In other words, *yo* was expressed at a rate of 54 percent, which is quite high compared to this group's overall rate of pronoun expression (28 percent). Forms *tiene* and *sé*, which also appeared frequently in the learners' data, occurred with pronouns at higher rates than infrequent verbs did. In other words, Level 1 learners tended to express pronouns at relatively high rates with their most frequent verbs.

Now consider Level 2. As shown in Table 6, these learners produced *sé* 57 times. In fact, the form *sé*, which represents more than 10 percent of these learners' total verb tokens, never appears with *yo*. *Creo*, on the other hand, occurs with *yo* at a rate of 17 percent, which is higher than these learners' overall rate (13 percent). Thus these more advanced learners are not simply expressing pronouns with any and all frequent verbs; instead, they are employing specific patterns with each verb. Moreover, these patterns are similar to trends found for native speakers, who often express *yo* with *creo* and *sé* without a pronoun (Travis & Torres Cacoullós 2012). In summary, there appears to be a direct

relationship between verb frequency and pronoun expression among Level 1 learners such that these learners produce more pronouns with their most frequent verbs. The Level 2 learners, however, display a more nuanced and selective pattern whereby particular verbs are becoming associated with pronoun expression or omission.

4.3. Frequency as a mediator of other linguistic factors, Level 2 learners

Although frequency does not have a direct impact on Level 2 learners' pronoun expression, it is possible that it acts as a mediator of the impact of Person, Switch-reference, Semantic class, and TMA, as discussed by Erker & Guy (2012) for native speakers. The results presented in the following sections show that frequency significantly interacted with most of the linguistic factors. We primarily focus here on only one measure of frequency due to space limitations. We report analyses examining Local lexical frequency, but not Repetition.⁸ In what follows we present details of results showing these interactions.

4.3.1. Interactions between Person and frequency, Level 2 learners

In this section we examine the effects of Person (1sg vs. 3sg) on subject pronoun expression among the verb tokens categorized as infrequent and frequent based on the Local lexical frequency measure. Table 7 presents these results for the Level 2 learners. The columns labeled 'N' present the total number of verb tokens in each context, while the columns labeled '% pro' show the percent of *expressed* subject pronouns in each context. For example, there were 119 1sg infrequent verbs, 12 percent of which appeared with expressed *yo*.

Table 7. Subject pronoun expression by Person and Local lexical frequency, Level 2 learners

infrequent				frequent			
1sg		3sg		1sg		3sg	
N	% pro	N	% pro	N	% pro	N	% pro
119	12	51	29	176	14	27	33

For Level 2 learners, the variable Person significantly impacts pronoun expression (see Tables 4a and 4b), regardless of frequency [infrequent verbs: $\chi^2(1)=7.23$, $p<.01$; frequent verbs: $\chi^2(1)=8.01$, $p<.01$]. The learners express subject pronouns more often with 3sg than with 1sg verbs, and this is true for both frequent and infrequent verb tokens. The lack of interaction between Local lexical frequency and Person is consistent with Erker and Guy's (2012) results for Person/number (when 2sg is excluded) among Spanish speakers of New York, as well as Bayley et al.'s (2013) results for Spanish speakers from Texas and California.

4.3.2. Interactions between Switch-reference and verb frequency, Level 2 learners

Table 8 presents the results for Switch-reference. The data are separated in the columns by the frequency measure and the Switch-reference variable.

Table 8. Subject pronoun expression by Switch-reference and Local lexical frequency, Level 2 learners

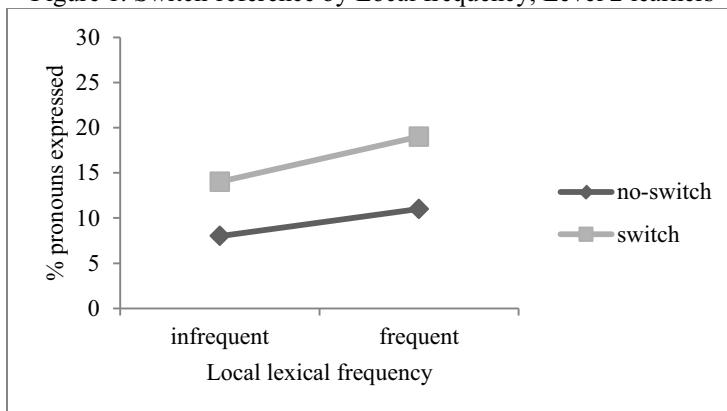
infrequent				frequent			
no switch		switch		no switch		switch	
N	% pro	N	% pro	N	% pro	N	% pro
167	8	141	14	128	11	104	19

Note. N = total verb tokens; % = percent of expressed subjects

⁸ We did, however, repeat all of the analyses reported in this section in order to investigate whether Repetition interacted with the other variables (Person, Switch-reference, Semantic class, and TMA), and we found similar results for Repetition as we did for Local lexical frequency.

The direction of the Switch-reference effect, that is, higher rates of pronoun expression in switch than in no-switch contexts, is consistent with studies of native speakers (e.g. Otheguy & Zentella 2012, Shin & Otheguy 2009), as well as previous L2 studies (Geeslin, Linford, & Fafulas 2012; Linford 2009, forthcoming). Here, we find that the impact of Switch-reference on pronoun expression is slightly stronger with frequent verbs than with infrequent verbs. The difference between switch and no-switch contexts for frequent verbs is eight percentage points, while it is six percentage points for infrequent verbs. Furthermore, the switch-reference effect more closely approaches significance for the frequent verbs [$\chi^2(1)=3.16$, $p=.076$] than for the infrequent verbs [$\chi^2(1)=2.62$, $p=.11$]. In figure 1, which displays these same results, it is evident that there is a wider gap between no-switch and switch contexts with frequent than with infrequent forms.

Figure 1. Switch-reference by Local frequency, Level 2 learners



The interpretation that there is an interaction between Switch-reference and frequency is buttressed by results for the Repetition frequency measure, in which the Switch-reference effect reaches significance with frequent verbs [$\chi^2(1)=6.33$, $p=.01$], but not with the infrequent ones [$\chi^2(1)=.09$, $p=.77$]. This result is consistent with Erker & Guy (2012), who find that Switch-reference is amplified among the most frequent verbs. Bayley et al. (2013), in contrast, found no such interaction.

4.3.3. Interactions between Semantic class and frequency, Level 2 learners

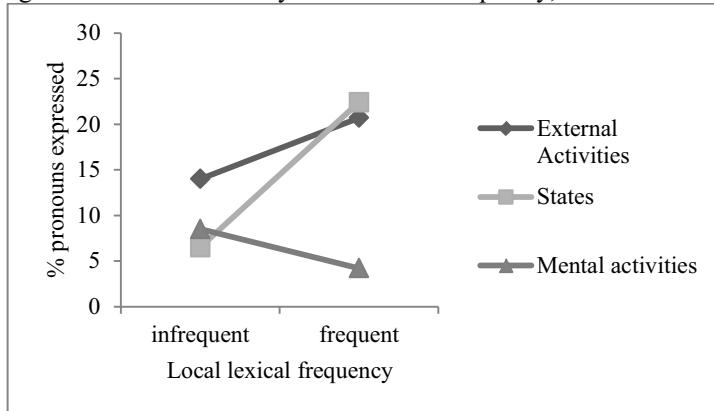
Semantic class was only significant in contexts with frequent verbs for the Level 2 learners, as demonstrated by the results presented in Table 9.

Table 9. Subject pronoun expression by Semantic class and Local lexical frequency, Level 2 learners

infrequent		frequent	
mental	external	mental	external
N	% pro	N	% pro
59	9	107	22
77	9	29	21
172	14	96	4

Semantic class had a significant impact on pronoun expression with the most frequent verbs [$\chi^2(2)=14.46$, $p=.001$], but not with the infrequent ones [$\chi^2(2)=3.503$, $p=.174$]. Figure 2 depicts the interaction between Semantic class and Local lexical frequency.

Figure 2. Semantic class by Local lexical frequency, Level 2 learners



Our finding that Semantic class is activated by frequency is consistent with Erker & Guy (2012:545) and is similar to Bayley et al. (2013), who find an amplifying effect of frequency on this variable. However, the direction of the Semantic class effect for frequent verbs in our study is unexpected; it is the mirror opposite of the findings of Erker & Guy (2012).⁹ Whereas native speakers produce pronouns most often with mental verbs, less often with stative verbs, and least often with external activity verbs, our Level 2 learners produced the highest rates of pronouns with external activity verbs, followed by stative verbs. Frequent mental activity verbs had the lowest rates of expression. This is primarily due to the verb form *sé*, a mental verb that was produced 57 times by the Level 2 learners; in all 57 cases *sé* occurred without *yo* (see Table 6).

4.3.4. Interactions between TMA and verb frequency, Level 2 learners

Finally, the results for the interactions between TMA and frequency for Level 2 learners are provided in Table 10.

Table 10. Subject pronoun expression by TMA and Local lexical frequency, Level 2 learners

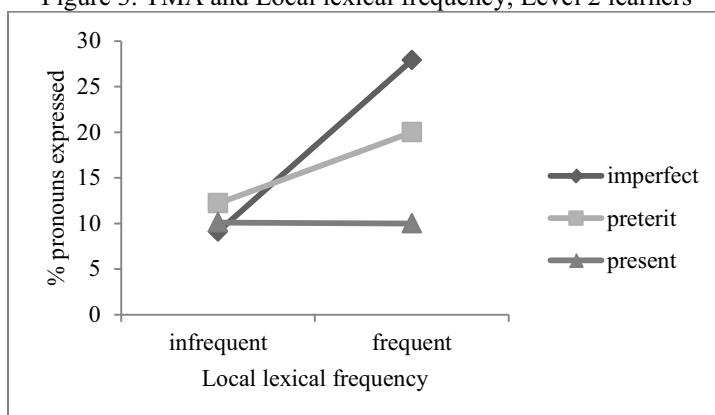
present		infrequent			frequent						
		preterit		imperfect		present		preterit		imperfect	
N	% pro	N	% pro	N	% pro	N	% pro	N	% pro	N	% pro
109	10	115	12	44	9	160	10	29	21	43	28

While TMA does not significantly impact pronoun expression when all verbs (frequent and infrequent) are grouped together (see Table 3), this variable is significant when frequent verbs are isolated [$\chi^2(1)=9.654$, $p<.01$]. This trend is similar to the results for native speakers in Erker & Guy (2012): more pronouns are expressed with imperfect forms (which have inherently ambiguous forms in some persons and numbers) than with preterit or simple present indicative forms.¹⁰ Moreover, although TMA influenced pronoun expression with both the frequent and infrequent verbs in Erker & Guy (2012), the impact was greater with the frequent verbs. The interaction of frequency and TMA in the current study is depicted in Figure 3.

⁹ It is difficult to compare our results for Semantic class with those of Bayley et al. (2013) due to differences in the operationalization of this variable.

¹⁰ As with Semantic Class, the results of TMA are not directly comparable to Bayley et al. (2013) due to differences in coding.

Figure 3. TMA and Local lexical frequency, Level 2 learners



5. Discussion

The current study was guided by three research questions investigating whether lexical frequency 1) independently influences L2 learners' Spanish subject pronoun expression, 2) mediates other linguistic variables that are known to impact Spanish subject pronoun use, and 3) affects L2 learners differently depending on their language proficiency. The answers to the three research questions are summarized in Table 11, which illustrates our findings alongside those of Bayley et al. (2013) and Erker & Guy (2012).

Table 11. Summary of research questions and findings

Research question	Our L2 study	Bayley et al. (2013)	Erker & Guy (2012)
1. Frequency directly impacts pronoun expression	✓	✓	X
2. Frequency mediates pronoun expression	✓	X	✓
3. Frequency effect dependent on proficiency	✓	--	--

In the following discussion we discuss the findings in an attempt to contribute to our general understanding of how frequency impacts the developing L2 grammar.

5.1. Frequency directly impacts pronoun expression

We found that frequency directly impacts pronoun use, but only in the beginning stages of Spanish second language acquisition. Level 1 learners, who were in their second year of Spanish classes, produced significantly higher rates of pronouns with frequent than with infrequent verbs. Bayley et al. (2013) also found a direct relationship between frequency and pronoun use. In their study, however, frequent verbs slightly disfavored pronoun expression. Thus our finding that pronouns are produced most often with frequent verbs and least often with infrequent ones appears to be characteristic of language development, not of Spanish speakers in general. One possible explanation for the L2 trend has to do with the cognitive demands associated with infrequent forms. We tentatively propose that L2 learners omit pronouns more often with infrequent verbs for the same reason they omit grammatical forms and structures in general, that is, because of the processing limitations during language production.¹¹ Such an interpretation suggests a parallel between L1 and L2 developmental

¹¹ Another possible explanation is that the patterns observed here reflect the input. It is possible that Spanish instructors' subject pronoun expression shows similar trends to those found for the Level 1 learners here. There is some evidence that instructors produce higher rates of subject pronouns in classroom settings than outside of the classroom (Dracos 2010), but we do not yet know if that tendency targets high frequency verb forms only. Further research investigating teachers' use of subject pronouns in the classroom in combination with measures of verb frequency would undoubtedly inform attempts to explain learners' developmental patterns.

processes, as children also omit subjects in contexts of heavier processing demands (Aranoff 2003, Bloom 1990). At the same time, with frequent verbs, Level 1 learners overproduce pronouns (at a rate 40 percent), a finding that concurs with studies that have shown overuse of pronouns in early stages of L2 acquisition (e.g. Montrul & Rodríguez-Louro 2006:409). In some respects, then, the L2 developmental process is reminiscent of L1 Spanish development: processing demands result in subject omission; for L2 learners this is apparent with infrequent forms, whereas for children acquiring Spanish as a first language, this is apparent in the early stages of acquisition (e.g. Grinstead 2004; Shin 2012; Shin & Erker, under review). When processing demands are lifted, learners sometimes overproduce pronouns. For L2 learners this is apparent with frequent verbs. For children, this overproduction stage occurs between the ages of eight and eleven, just as children are learning to handle all the subtleties involved in the use of referring expressions in discourse (Shin & Cairns 2012).

An interesting contribution of the current study is the new Repetition measure that was used to investigate the direct impact of frequency on pronoun expression. This measure was included to examine whether a learners' own repetition of a verb could simulate a frequency effect. The assumption was that repeating a verb temporarily increases the strength of its mental representation and, in turn, decreases the cognitive processing load associated with producing the verb form (Ellis 2012). We found that both frequency measures functioned in a similar fashion, suggesting that repetition within a conversation is similar to repetition of forms over a lifetime: Both set the stage for the routinization of linguistic patterns (Bybee 2001:6-8).

While we found a direct relationship between frequency and pronoun expression among the Level 1 learners, we did not find this same effect among the Level 2 learners. Thus the independent effect of frequency appears to dissipate as learners become more proficient in Spanish. Interestingly, concomitant with this disappearance of the direct impact of frequency, we find an increase in sensitivity to other linguistic variables that impact subject pronoun production (i.e. Switch-reference, Person, and Semantic class). As learners become more proficient, their subject pronoun use is guided by increasingly subtle linguistic cues.

5.2. *Frequency mediates other linguistic constraints on pronoun expression*

While frequency did not have a direct impact on Level 2 learners' pronoun expression, we did find that frequency mediated other linguistic variables for these learners. Overall, our findings for the Level 2 group are consistent with Erker & Guy's (2012) study of Spanish spoken in New York. Frequency very clearly mediated sensitivity to TMA and Semantic class,¹² such that pronoun expression was conditioned by these two variables with frequent, but not infrequent, verbs. Also, we found that frequency had a weaker impact on Switch-reference, which supports Erker & Guy's claim that "switch-reference is orthogonal to lexical selection" (p. 552).

Frequency did not mediate the effects of Person in our study. Regardless of whether verbs were frequent or infrequent, we found that subject pronouns were expressed more often with 3sg than with 1sg verb forms.¹³ It is important to note that we only included two grammatical persons, 1sg and 3sg. Previous research suggests that a Person-frequency effect might emerge if other persons are included, since Erker & Guy (2012:545) found that *tú* was expressed significantly more often with frequent verb forms than with infrequent ones. Future investigations of L2 Spanish pronoun expression are needed to determine whether frequency mediates advanced learners' use of *tú*, *usted*, or any plural pronouns.

To summarize, our results for the advanced learners suggest native-like trends. Their patterns of pronoun use are guided by linguistic factors such as Person, Switch-reference, TMA, and Semantic Class, as found in Otheguy & Zentella (2012), *inter alia*. Furthermore, lexical frequency does not directly affect their pronoun expression, but it does mediate the impact of other linguistic variables, as found in Erker & Guy (2012).

¹² With both Switch-reference and TMA, the direction of the effect was native-like. For Semantic class, however, we found that the lowest rates of pronoun expression were with mental activity verbs. The likeliest account for this finding is that the most frequent verb among Level 2 learners was *sé*, a mental activity verb that was never produced with the pronoun *yo* (see Table 6).

¹³ In Bayley et al. (2013) and Erker & Guy (2012) pronoun rates were higher with 1sg than with 3sg verbs.

5.3. Frequency effect changes with L2 proficiency

The final research question was whether frequency effects differ depending on L2 learners' proficiency. Our study offers an affirmative answer: Whereas frequency had a direct impact on pronoun expression in the early stages of Spanish acquisition, we found no such effect among the more advanced learners who were in their fourth year of Spanish classes. As discussed above, we think that the frequency effects found for the Level 1 learners are a response to an increased cognitive load. The production of an infrequent verb presents a challenging task for learners with low proficiency levels, triggering omission of forms that they would otherwise produce (in this case, pronouns). In contrast, frequent forms are part of these learners' routines and, therefore, the learners can handle other tasks while producing frequent forms (in this case, they can express pronouns along with production of verb forms). Thus lexical frequency can serve to bolster the production of grammatical forms. This result has a direct implication for the L2 classroom in that it suggests that grammatical constructions should be introduced first with high frequency lexical items.

6. Conclusion

The results of the current study suggest that lexical frequency plays a role in subject pronoun expression in L2 Spanish, but the effect differs according to proficiency levels. For less proficient learners, frequency functions in an independent manner whereby pronouns are expressed more often with frequent than with infrequent verbs. On the other hand, for more proficient learners, frequency activates or enhances the impact of other linguistic variables. Overall our results indicate that lexical frequency should be taken into account when introducing grammatical constructions in the L2 classroom. Learners are more likely to be successful if instructors couple new grammatical constructions with high frequency forms that are already entrenched in their linguistic routines.

References

- Abreu, Laurel. 2012. Subject Pronoun Expression and Priming Effects among Bilingual Speakers of Puerto Rican Spanish. In Geeslin & Díaz-Campos (Eds.), 1-8.
- Andersen, Roger. 1986. El desarrollo de la morfología verbal en el español como segundo idioma. In Jürgen Meisel (Ed.), *Adquisición de Lenguaje/Adquisicao da linguagem*, 115-138. Frankfurt: Vervuert.
- Andersen, Roger & Yashuhiro Shirai. 1994. Discourse motivations for Some Cognitive Acquisition Principles. *Studies in Second Language Acquisition* 16(2), 133-156.
- Aronoff, Justin. 2003. Null Subjects in Child Language: Evidence for a Performance Account. In Gina Garding & Mimu Tsujimura (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, 43-55. Somerville, MA: Cascadilla Press.
- Bardovi-Harlig, Kathleen. 1998. Narrative Structure and Lexical Aspect: Conspiring Factors in Second Language Acquisition of Tense-Aspect Morphology. *Studies in Second Language Acquisition* 20(4), 471-508.
- Bayley, Robert, Cory Holland & Kristen Ware. 2013. Lexical Frequency and Syntactic Variation: A Test of a Linguistic Hypothesis. To appear in *University of Pennsylvania Working Papers in Linguistics* 19(2).
- Bayley, Robert & Lucinda Pease-Álvarez. 1996. Null and expressed pronoun variation in Mexican-descent children's Spanish. In Jennifer Arnold, Renee Blake & Brad Davidson (Eds.), *Sociolinguistic variation: Data, theory, and analysis*, 85-99. Stanford: Center for the Study of Language and Information.
- Bayley, Robert & Lucinda Pease-Álvarez. 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9(3), 349-371.
- Bentivoglio, Paola 1987. *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela.
- Bloom, Paul. 1990. Subjectless sentences in child language. *Linguistic Inquiry* 21, 491-504.
- Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Cadierno, Teresa. 2000. The acquisition of Spanish preterite and imperfect by Danish advanced language learners. *Spanish Applied Linguistics* 4(1), 1-53.
- Cameron, Richard. 1994. Switch reference, verb class and priming in a variable syntax. In K. Beals (Ed.), *Papers from the 30th regional meeting of the Chicago Linguistics Society: Volume 2: The parasession on variation in linguistic theory*, 27-45. Chicago Linguistics Society.
- Cameron, Richard. 1995. The Scope and Limits of Switch Reference as a Constraint on Pronominal Subject Expression. *Hispanic Linguistics* 6/7, 1-28.

- Cameron, Richard & Nydia Flores-Ferrán. 2004. Perseveration of subject expression across regional dialects of Spanish. *Spanish in Context* 1(1), 41-65.
- Canale, Michael & Merrill Swain. 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1, 1-47.
- Cantero Sandoval, Gustavo. 1978. Observaciones sobre la expresión innecesaria de los pronombres personales sujeto en el español de México. *Anuario de Letras* XVI, 261-264.
- Dracos, Melisa. 2010. Spanish subject pronoun use in instructional input. Paper presented at the *14th Hispanic Linguistics Symposium*, Bloomington, IN.
- Ellis, Nick. 1996. Sequencing in SLA: Phonological Memory, Chunking and Points of Order. *Studies in Second Language Acquisition* 18(1), 91-126.
- Ellis, Nick. 2012. What can we count in language, and what counts in language acquisition, cognition, and use? In Stefan Th. Gries & Dagmar Divjak (Eds.) *Frequency effects in cognitive linguistics (Vol. 1): Statistical effects in learnability, processing and change*, 7-34. Berlin: Mouton de Gruyter.
- Ellis, Nick & Laura Collins. 2009. Input and Second Language Acquisition: The Roles of Frequency, Form, and Function Introduction to the Special Issue. *The Modern Language Journal* 93(3), 329-335.
- Enriquez, Emilia. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
- Erker, Daniel. 2005. *Functional Compensation for Morphological Ambiguity in New York City Spanish*. City University of New York Graduate Center MA thesis.
- Erker, Daniel & Gregory Guy. 2012. The role of lexical frequency in syntactic variability: Variable subject Personal pronoun expression in Spanish. *Language* 88(3), 526-557.
- Flores-Ferrán, Nydia. 2002. *A sociolinguistic perspective on the use of subject personal pronouns in Spanish narratives of Puerto Ricans in New York City*. Munich: Lincom-Europa.
- Flores-Ferrán, Nydia. 2005. La expresión del pronombre personal sujeto en narrativas orales de puertorriqueños en Nueva York. In Luis Ortiz López & Manel Lacorte (Eds.), *Contactos y contextos lingüísticos: El español en los Estados Unidos y en contacto con otras lenguas* 119-129. Madrid/Frankfurt: Iberoamericana/Vervuert.
- Geeslin, Kimberly. 2006. Task Design, Discourse Context and Variation in Second Language Data Elicitation. In Carol Klee & Tim Face (Eds.), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, 74-85. Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, Kimberly. 2011. The acquisition of variation in second language Spanish: How to identify and catch a moving target. In Manuel Díaz-Campos (Ed.), *The Handbook of Spanish Sociolinguistics* (pp. 303-320). Malden, MA: Wiley-Blackwell.
- Geeslin, Kimberly & Manuel Díaz-Campos (Eds.). 2012. *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, Kimberly & Aarnes Gudmestad. 2008. 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 2007 Second Language Research Forum*, 69-85. Somerville, MA: Cascadilla Press.
- Geeslin, Kimberly & Aarnes Gudmestad. 2010. An exploration of the range and frequency of occurrence of forms in potentially-variable structures in second language Spanish. *Studies in Second Language Acquisition* 32(3), 433-463.
- Geeslin, Kimberly & Aarnes Gudmestad. 2011. Using Sociolinguistic Analyses of Discourse-Level Features to Expand Research on L2 Variation: Native and Non-Native Contrasts in Forms of Spanish Subject Expression. In Luke Plonsky and Maren Schierloh (Eds.), *Selected Proceedings of the 2009 Second Language Research Forum: Diverse Contributions to SLA*, 16-30. Somerville, MA: Cascadilla Press.
- Geeslin, Kimberly & Aarnes Gudmestad. 2012. Individual patterns of verb selection: The case of future-time expression in second-language Spanish. Paper presented at the *31st Second Language Research Forum*, Pittsburgh, PA.
- Geeslin, Kimberly & Bret Linford. 2012. A cross-sectional study of the effects of discourse cohesiveness and perseveration on subject expression. Paper presented at the *6th International Workshop on Spanish Sociolinguistics*, Tucson, Arizona.
- Geeslin, Kimberly, Bret Linford & Stephen Fafulas. 2012. The L2 development of subject form variation in Spanish: The individual vs. the group. Paper presented at the *New Ways of Analyzing Variation*. Bloomington, Indiana.
- Grinstead, John. 2004. Subjects and interface delay in child Spanish and Catalan. *Language* 80(1). 40-72.
- Gudmestad, Aarnes. 2012. Toward an Understanding of the Relationship between Mood Use and Form Regularity: Evidence of Variation across Tasks, Lexical Items, and Participant Groups. In Geeslin & Díaz-Campos (Eds), 214-227.
- Gudmestad, Aarnes & Kimberly Geeslin. 2007. A closer look at advanced learners and sociolinguistic variation: Spanish-language use in contexts of future time reference. Paper presented at the *11th Hispanic Linguistics Symposium*, San Antonio, TX.
- Gudmestad, Aarnes & Kimberly Geeslin. 2010. Exploring the Roles of Redundancy and Ambiguity in Variable

- Subject Expression: A Comparison of Native and Non-native Speakers. In Claudia Borgonovo, Manuel Español-Echevarría & Philippe Prevost, (Eds.), *Refereed Proceedings from the 12th Hispanic Linguistics Symposium*, 270-283. Somerville, MA: Cascadilla Press.
- Hochberg, Judith. 1986. Functional compensation for /-s/ deletion in Puerto Rican Spanish. *Language* 62, 609-621.
- Holmquist, Jonathan. 2012. Frequency rates and constraints on subject Personal pronoun expression: Findings from the Puerto Rican highlands. *Language Variation and Change* 24(2), 203-220.
- Lapidus, Naomi & Ricardo Otheguy. 2005. Overt nonspecific *ellos* in Spanish in New York. *Spanish in Context* 2(2), 157-174.
- Lieven, Elena. 2010. Input and first language acquisition: Evaluating the role of frequency. *Lingua* 120. 2546-2556.
- Linford, Bret. 2009. *The Acquisition of Subject Pronouns in Second Language Spanish*. The University of Montana MA thesis.
- Linford, Bret. Forthcoming. Self-reported motivation and the L2 acquisition of subject pronoun variation in Spanish. In Ryan T. Miller, Katherine I. Martin, Chelsea M. Eddington, Ashlie Henery, Nausica Marcos Miguel, Alison Tseng, Alba Tuninetti, & Daniel Walter (Eds.), *Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges Between Disciplines*.
- Lubbers Quesada, Margaret. 1998. L2 acquisition of the Spanish subjunctive mood and prototype schema development. *Spanish Applied Linguistics* 2.1-23.
- Montrul, Silvina & Celeste Rodríguez-Louro. 2006. Beyond the syntax of the Null Subject Parameter. A look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish. In Vincent Torrens & Linda Escobar (Eds.), *The acquisition of syntax in Romance Languages*, 401-418. Amsterdam: John Benjamins.
- Otheguy, Ricardo & Ana Zentella. 2012. *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. New York: Oxford University Press.
- Pérez-Leroux, Ana & William Glass. 1999. Null anaphora in Spanish second language acquisition: probabilistic versus generative approaches. *Second Language Research* 15(2), 220-249.
- Rothman, Jason 2007. Pragmatic solutions for syntactic problems: Understanding some L2 syntactic errors in terms of discourse-pragmatic deficits. In Sergio Baauw, Frank Drijkoningen & Manuela Pinto (Eds.), *Romance languages and linguistic Theory 2005. Selected papers from Going Romance. Current Issues in Linguistic Theory* 29, 299-320. John Benjamins: Amsterdam.
- Rothman, Jason. 2010. Theoretical Linguistics Meets Pedagogical Practice: Pronominal Subject Use in Spanish as a Second Language (L2). *Hispania* 93, 152-165.
- Salaberry, Rafael. 2000. *The development of past tense morphology in L2 Spanish*. Philadelphia: John Benjamins.
- Shin, Naomi Lapidus. 2012 Variable use of Spanish subject pronouns by monolingual children in Mexico. In Geeslin & Díaz-Campos (Eds.), 130-141.
- Shin, Naomi Lapidus & Helen Cairns. 2012. The development of NP selection in school-age children: Reference and Spanish subject pronouns. *Language Acquisition* 19(1), 3-38.
- Shin, Naomi Lapidus & Daniel Erker. Under review. The emergence of structured variability in morphosyntax: Childhood acquisition of Spanish subject pronouns.
- Shin, Naomi Lapidus & Ricardo Otheguy. 2009. Diminishing sensitivity to continuity of reference: Subject pronoun use in New York City. In Manel Lacorte & Jennifer Leeman (Eds.) *Español en Estados Unidos y en otros contextos: Cuestiones sociolingüísticas, políticas, y pedagógicas*, 111-136. Madrid/Frankfurt: Iberoamericana/Vervuert Verlag.
- Shin, Naomi Lapidus & Ricardo Otheguy. 2013. Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. To appear in *Language in Society* 42(4).
- Shirai, Yasuhiro & Roger Andersen. 1995. The acquisition of tense/aspect morphology: A prototype account. *Language* 71, 743-62.
- Silva-Corvalán, Carmen. 2001. *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.
- Solomon Julie. 1999. *Phonological and Syntactic Variation in the Spanish of Valladolid, Yucatán*. Stanford University PhD dissertation.
- Torres Cacoullous, Rena & Catherine Travis. 2010. Variable *yo* expression in New Mexico: English influence? In Susana Rivera-Mills & Daniel Villa Crésap (eds), *Spanish of the U.S. southwest: A language in transition*, 189-210. Madrid: Iberoamericana/Vervuert.
- Torres Cacoullous, Rena & Catherine Travis. 2011. Testing convergence via code-switching: Priming and the structure of variable subject expression. *International Journal of Bilingualism* 15(3). 241-267.
- Travis, Catherine. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19(2), 101-135.
- Travis, Catherine & Rena Torres Cacoullous. 2012. What do subject pronouns do in discourse? *Cognitive Linguistics* 23(4), 711-748.
- Zipf, George. 1935. *The Psycho-Biology of Language*. Boston: Houghton Mifflin.

Selected Proceedings of the 16th Hispanic Linguistics Symposium

edited by Jennifer Cabrelli Amaro,
Gillian Lord, Ana de Prada Pérez,
and Jessi Elana Aaron

Cascadilla Proceedings Project Somerville, MA 2013

Copyright information

Selected Proceedings of the 16th Hispanic Linguistics Symposium
© 2013 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-459-1 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 or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at 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:

Linford, Bret and Naomi Lapidus Shin. 2013. Lexical Frequency Effects on L2 Spanish Subject Pronoun Expression. In *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, ed. Jennifer Cabrelli Amaro et al., 175-189. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2933.