On the Relationship Between Subject Placement and Overt Pronouns in the Spanish of New York City Bilinguals

Rocío Raña-Risso

City University of New York

rranarisso@gradcenter.cuny.edu

Carolina Barrera-Tobón

DePaul University

cbarrer6@depaul.edu

Abstract

This paper reports on a variationist sociolinguistic analysis of the relationship between subject placement and overt pronoun rates in the Spanish of first- and second-generation Spanish-English bilinguals in New York City. The data used for the study come from a spoken corpus of Spanish in New York based on 140 sociolinguistic interviews. We show second-generation speakers exhibit a more rigid word order compared to their newly arrived first-generation peers, more often preferring subjects in the preverbal position, and we explain that this increase in word order rigidity among our second-generation can be attributed, in large part, to their increased use of and contact with English. We further posit that the difference in subject placement across generations can be explained by the different context of acquisition since the Spanish that these second-generation speakers are exposed to contains both a higher rate of overt pronouns and a higher rate of preverbal subjects.

Keywords

subject placement – pronoun expression and placement – contact-induced changes

1 Introduction

The recent waves of immigration from Latin America to the United States (henceforth US) have produced a growing interest in the field of language contact for the study of Spanish in the US. The contact situation between English and Spanish in the US has yielded several linguistic effects. Most researchers in language contact agree that the lexicon is the linguistic element most permeable to external influence in situations of language contact (Sankoff, 2002; Silva-Corvalán, 2001), and whether other linguistic domains are subject to similar influence is a central debate among theorists of language contact (Sankoff, 2002; Thomason, 1995, 2001; Thomason and Kaufman, 1988; van Hout and Muysken, 1994). Recently, several studies have looked at these effects on syntactic variables such as subject and object expression and how these relate to the pro-drop parameter in the Spanish spoken in the US (Montrul, 2004; Toribio, 2004). These studies indicate that areas where syntax borders with other domains, such as lexical-semantics, syntax-semantics, and discourse-pragmatics, may be vulnerable to change in contact situations.

If the interface between syntax and other linguistic domains were in fact vulnerable to change due to language contact, then we would expect that syntactic features that are affected by semantic-pragmatic constraints would undergo some changes. Several studies on language contact in general and Spanish in the US in particular have supported the idea that grammar external interfaces are in fact more susceptible to cross-linguistic effects (Montrul, 2004, 2008; Sorace and Serratrice, 2009; Tsimpili and Sorace, 2006; Toribio, 2004; Zapata, Sánchez and Toribio, 2005). Although several of these studies have focused on the Null subject parameter of heritage language speakers (Montrul, 2004; Sorace and Serratrice, 2009; Zapata, Sánchez and Toribio, 2005), few have looked at the relationship between pronominal expression and subject placement. In this study we compare subject placement in the Spanish of two generations of Latinos in New York City (henceforth NYC), and examine whether any generational differences we find in subject placement may be related to an increased use of subject personal pronouns. We focus on this relationship precisely because both of these variables are controlled by the pro-drop parameter. This parameter specifies that if a language can have null subjects, it can also have postverbal subjects and if a language cannot have null subjects, then it cannot freely place their subjects postverbally. Spanish conforms well to this characterization, displaying the traits of a null-subject language while English displays the traits of a non pro-drop language. As such, if we observe generational differences in both of these variables (subject placement and subject pronoun expression) these differences may indicate that the second

generation of Spanish speakers in NYC is moving towards a different setting in this parameter.

NYC is an ideal place for a study of this kind for various reasons. First, the Latino population of NYC maintains its heritage language in part due to the constant arrival of Spanish-speaking newcomers. Meanwhile the majority of NYC’s Spanish-speaking population also speaks English. This suggests that the majority of NYC Latinos are bilingual, giving rise to contact between Spanish and English in those speakers. The census data presented in the section below confirm this notion.

2 Spanish in NYC

According to the 2010 Census data, the Latino population of NYC constitutes 29.1 percent of the City’s total population. There are six ethnonational groups that conform 85 percent of the total Latino population for this year: Puerto Ricans (31 percent), Dominicans (25 percent), Mexicans (14 percent), Ecuadorians (9 percent), Colombians (4 percent), and Cubans (2 percent). These numbers show that 58 percent of NYC’s Latino population is from the Caribbean and 42 percent from the Latin American Mainland.

In order to explore the language usage trends of speakers in these six groups (the Major ethnonational Latino Groups) in NYC, we looked at the numbers of people in these six groups who reported speaking Spanish at home according to the 1990, 2000, and 2010 Census (as reported in Bergad, 2011). In Table 1 below we see the percentage of the Latino population in NYC that reported speaking Spanish at home in the last three censuses.

Although the percentage of the total Latino population that reports speaking Spanish at home is slowly decreasing (from 81.5 percent in 1990 to 76.2 percent in 2010), it can be concluded that the Latino population of NYC is still maintaining its heritage language since the decrease (5.3 percent) in the use of

TABLE 1 *NYC population in major Latino groups 5 years of age and over that speaks Spanish at home per census year.*

Census Year	Percent
1990	82
2000	79
2010	76

Spanish at home in the last 20 years has been relatively small. However, when we looked at each of the six leading groups individually, we noticed that there are differences among the groups. While use of Spanish at home has remained stable for Dominicans, Ecuadorians, and Colombians, it has significantly decreased for Cubans (from 82 percent in 1990 to 55.6 percent in 2010) and Puerto Ricans (from 79 to 65.4 percent), and increased for Mexicans (from 73 to 75.5 percent). One way of explaining these findings is by looking at the distribution of the population, where it can be noticed that while the population of Mexicans in NYC more than tripled in 20 years (from 4 to 14 percent), the population of Cubans was cut in half (from 4 to 2 percent), and the population of Puerto Ricans also decreased significantly (from 58 to 31 percent). Therefore, it seems that an increase in the number of speakers from a particular nationality correlates with maintenance of Spanish in that subgroup, while a decrease in the number of speakers correlates with assimilation to English in that subgroup.

Furthermore, if we classify each group by their place of birth (domestic or abroad¹) as shown in Table 2 below, we see that besides Puerto Ricans and Cubans, the population of the other four national groups in the City is mostly foreign born. This means that these national groups are continuously receiving

TABLE 2 *Distribution of the total NYC population in major ethnonational Latino groups born domestically or abroad by census year.*

Major Latino Groups	Place of Birth	1990	2000	2010
Mexican	Domestic Born	34%	32%	42%
	Foreign Born	66%	68%	58%
Ecuadorian	Domestic Born	24%	25%	33%
	Foreign Born	76%	75%	67%
Colombian	Domestic Born	21%	23%	34%
	Foreign Born	79%	77%	66%
Cuban	Domestic Born	28%	42%	55%
	Foreign Born	72%	58%	45%
Puerto Rican	Domestic Born	58%	63%	71%
	Foreign Born	42%	37%	29%
Dominican	Domestic Born	30%	31%	36%
	Foreign Born	70%	69%	64%

1 Reports produced by the Center for Latin American, Caribbean, and Latino Studies at the CUNY Graduate Center consider Puerto Ricans born on the island as foreign born, while those born in the continental United States are considered domestic born.

new immigrants, which helps maintain Spanish alive in the city. The results presented in Table 2 also help illustrate the reason why although Spanish is very much alive in the city, the number of speakers is slowly decreasing: the number of foreign-born speakers is also slowly decreasing for every single group as the years go by.

We also used census data to investigate whether those who speak Spanish at home also speak English well, or whether use of Spanish at home correlates with lack of skills in English. As illustrated by Table 3 below, we found that approximately 70 percent of the total population in the major ethnonational Latino groups that speaks Spanish at home, also reports speaking English well or very well, 20 percent reports speaking English but not well, and only 10 percent reports not speaking English. Furthermore, this situation has remained stable throughout the last 20 years. There are two interesting comments to make about these results: 1) use of Spanish at home does not seem to negatively affect the development of English language skills in the Latino population; 2) the vast majority of the Latino population appears to be bilingual.

The following conclusions have arisen about the NYC Latino population from the analysis of the 1990, 2000 and 2010 census data:

- The Latino population of NYC is maintaining its heritage language.
- 58 percent of NYC’s Latino population is from the Caribbean and 42 percent is from the Mainland.
- The ongoing arrival of Spanish-speaking newcomer immigrants helps maintain Spanish alive in the City.
- 70 percent of NYC’s Spanish-speaking population also speaks English well.

The census data presented above show that the majority of New York Latinos are bilingual. Therefore, NYC is an ideal environment to explore language contact phenomena: 70 percent of NYC’s Spanish-speakers report being balanced

TABLE 3 *Level of English skills among those who speak Spanish at home ages 5 and over in major Latino ethnonational groups.*

English skills	1990	2000	2010
Yes, speaks very well	49%	48%	50%
Yes, speaks well	23%	22%	19%
Yes, but not well	19%	20%	22%
Does not speak English	9%	10%	10%

bilinguals, giving rise to contact between Spanish and English in those speakers, possibly leading to an influence of English on the Spanish spoken in the City.

The fact that a large majority of Latinos reported speaking English well and also reported speaking Spanish at home indicates that bilingualism is a prominent characteristic among the City's Latinos. This large degree of bilingualism may give rise to several contact linguistic phenomena, and due to this large degree of bilingualism and the richness in dialectal varieties, it presents an ideal setting for a study on language and dialectal contact. Thus far, however, no comprehensive sociolinguistic study has been done on the Spanish in NYC, while such studies exist for the Spanish in Chicago and in Los Angeles (Silva-Corvalán, 1995, 2001; Torres and Potowski, forthcoming). However, researchers at the CUNY Graduate Center and the Research Institute for the Study of Language in Urban Society (RISLUS) have begun such work using the Otheguy-Zentella corpus. This corpus, described in depth in the following section, will be used to examine subject placement and subject personal pronoun expression in the Spanish of two generations of Latinos in NYC.

3 The Otheguy-Zentella Corpus

As stated previously, the data for this project were extracted from the Otheguy – Zentella corpus (henceforth referred to as 'the corpus' or 'the original corpus') which contains over a million words and consists of 140 transcripts of sociolinguistic interviews conducted in NYC, lasting between 60 and 90 minutes each. The corpus was constructed mostly in the period between 2000 and 2004 at the CUNY Graduate Center by a team of researchers from the RISLUS.² From these sociolinguistic interviews, separately stratified samples of Colombians, Dominicans, Ecuadorians, Mexicans, Puerto Ricans, and Cubans living in NYC were created. The stratification criteria used to create the sample were, within each ethnonational group, self-reports of: (a) Place of birth (Latin America or NYC), (b) age, (c) age of arrival in NYC (for informants born in the city, age of arrival is zero), (d) years of residence in NYC (for informants born in the city, years in NYC is equivalent to their age, with some exceptions to account for intervening years spent back in their parents' country of origin), (e) levels of education, (f) level of English skills, (g) social class, (h) extent of use of Spanish

² This corpus was developed with grants from the National Science Foundation (BCS 0004133), the City University of New York (09-91917), and CUNY's Professional Staff Congress (62666-00-31).

in general and in specific domains, (i) extent of use of Spanish with Latinos from one's own country, and (j) extent of use of Spanish with Latinos from countries other than one's own. For additional information regarding the data see Otheguy and Zentella (2012).

Of special import for the present study is the stratification by exposure to English (we measure Exposure to English based on length of stay and place of birth), which allows us to observe generational differences in word order synchronically through the apparent-time hypothesis (Labov, 1963). Following Otheguy and Zentella (2012), who consider the reference lect to be the Spanish spoken by recently arrived immigrants, we analyzed three different exposure groups. The first group, the Latin American Raised (LAR) Newcomers were born in Latin America, raised there, received their education there, and arrived in the City after the age of 16. They have been in the City for less than five years. The second exposure group is the LAR established immigrants. Like the Newcomers, the established immigrants were also born and raised in Latin America, but they have spent more than five years in NYC. The final group, the New York Raised (NYR), were born in the City, or arrived before the age of three, and have two parents that are considered first-generation immigrants. Although we used the Newcomers as our reference lect, in the next section we examine several analyses of word order in both contact and monolingual Spanish contexts.

4 Subject Placement in Spanish

Traditionally, Spanish is classified as an SVO language, that is, the basic, or canonical, word order for pragmatically neutral finite declarative sentences with transitive verbs is Subject-Verb-Object. For copular verbs (the verbs *ser* 'to be', *estar* 'to be' and *parecer* 'to seem') the canonical word order is Subject-Copula-Complement (Hawkins 1983; Tavaniers 2005; Zagana 2002). Ocampo (1995) calls this word order *informational word order* because it is pragmatically neutral. This means that the most common word order is a preverbal subject and a postverbal object or adjectival complement, as in the example below.

Yo recibí eso. [012C]³

I received that.

'I received that.'

3 Numbers and letters following examples are examples from the corpus; the number corresponds to the participant number and the letter corresponds to the origin of the participant (C = Colombian, U = Cuban, D = Dominican, E = Ecuadorian, M = Mexican, P = Puerto Rican).

Departure from canonical word order is a function of numerous internal variables including discourse pragmatic variables, processing variables, and structural variables such as pragmatic function, focus, constituent length, type of verb, the informational status of the nouns, as well as the type of sentence (Arnold, Losogco, Wasow and Ginstrom, 2000; Bentivoglio and Sedano, 2001; Behagel, 1909; Bentivoglio, 2003; Butt and Benjamin, 1996; Givón, 1993; Goodall, 2004; Nava, 2007; Ocampo, 1995, 2002; Prince, 1981; Quirk, Greenbaum, Leech and Svartik, 1972; Raña Risso, 2013; Tavaniers, 2005; Zagona, 2002; Zubizarreta, 2001). Since the focus of the present analysis are the external variables that condition word order variability, we will not discuss the numerous internal variables that condition word order. For our purpose, only verbs that could appear with a nominal or pronominal subject were included in the analysis.

4.1 *Subject Placement in Contact Situations*

Several studies have focused on Spanish word order in contact situations. One of these is a sociolinguistic study by Silva Corvalán (1994) of 50 Mexican-American bilinguals in Los Angeles in which the author found evidence for what she calls *obligatory SVX order* among the second- and third-generation speakers in her corpus, in comparison to the reference lects (popular forms of Spanish in Mexico), in which word order is dependent on many variables, including the discourse-pragmatic function of the utterance. The author argues that this is an example of indirect transfer from English, which has a relatively fixed Subject-Verb-Object (henceforth SVO) order. Specifically, the author found an increased rate of preverbal subject NP's and subject personal pronouns as the speaker's contact with English increased. The author maintains that the increase in SVX order is a "consequence of processes of loss of semantic-pragmatic constraints on preverbal subject placement" (144) that can be attributed to contact with the bilingual's dominant language.

In an experimental study on the production and interpretation of sentences with intransitive verbs by 24 heritage speakers of Spanish, Zapata, Sanchez and Toribio (2005) also found evidence of convergence toward English word order. The second- and third-generation bilingual participants in their study preferred SV order in Spanish in contexts where the preference in the reference lect is for VS order (such as unaccusative verbs). In other words, the second- and third-generation speakers produced word orders that were infelicitous compared to their first-generation counterparts. The authors attribute the increased incidence of SV order in these contexts to convergence with English and a reduction in the syntactic options that allow the speakers to communicate discourse-semantic information.

4.2 *Subject Placement in NYC*

In their analyses of NYC Spanish, Barrera-Tobón (2013) and Barrera-Tobón and Raña Risse (2016) found evidence of an increasingly fixed word order among second-generation Latinos. Barrera-Tobón's (2013) analysis focuses on the variable word order and prosody of copular constructions in the Spanish of first- and second-generation Spanish-English bilinguals in NYC. Her data suggest that speakers in NYC, like their peers in Los Angeles and other parts of the US, exhibit a more rigid word order compared to their first-generation counterparts.

Similarly, Barrera-Tobón and Raña Risse (2016) also find generational differences in both nominal and pronominal subjects. Specifically, they find that second-generation speakers exhibit a more rigid word order compared to their first-generation peers, and attribute this difference to their increased use of and contact with English. However, they also find that regional differences in subject placement in the first generation are maintained, to a certain degree, in the second generation, indicating that both English contact and continuity with Latin American ways of speaking shape Spanish in NYC.

5 *Regional Differences in Subject Placement*

It has been suggested in the literature that the use of preverbal pronouns and nominal subjects allows us to group individuals based on their place of origin, with speakers from the Caribbean placing subjects before verbs more often than Spanish-speakers from other regions (Barrera-Tobón, 2013; Goodall, 2004; Lipski, 1994: 241; López-Morales, 1992: 137; Ordoñez and Olarrea, 2001; Otheguy and Zentella, 2012; Raña Risse, 2010; Raña Risse, 2013; Toribio, 2000; Zagona, 2002). In the Spanish of New Yorkers, although there is variability in the use of subject pronouns and between preverbal and postverbal subject placement, the rates of subject pronoun use and preverbal or postverbal subject placement vary considerably between communities (Barrera-Tobón, 2013; Otheguy and Zentella, 2012; Raña Risse, 2013).

In her study of pronominal subject placement in NYC, Raña Risse (2013) found that Caribbean speakers use preverbal subject pronouns at a substantially higher rate than Mainlanders, and she concluded that, for pronominal subject placement, Caribbeans and Mainlanders very likely constitute two sub-communities of Spanish-speakers in NYC.

In considering other external variables within each regional group, Raña Risse found that whereas Mainlanders tend to have uniform preverbal rates

(percentage of total overt subject pronouns placed before the verb) throughout the population, Caribbeans are well differentiated by education and socio-economic status. In other words, Caribbean speakers who have a higher level of education tend to place pronominal subjects preverbally more often than speakers with lower levels of education, and those with a higher socio-economic status also place pronominal subjects preverbally more often than those with lower socio-economic status.

In comparing the results from Raña Risso's (2013) investigation with those of Otheguy and Zentella's (2012), we can see a clear difference in two aspects of subject pronoun usage between the Caribbean and the Mainland: Caribbeans use more overt pronouns than Mainlanders, and they place them in preverbal position more than Mainlanders. Thus, when we study pronominal subjects, we can conclude that there are two distinct dialects of Spanish interacting in NYC, corresponding to the Caribbean and the Mainland. These two separate dialectal groupings enter the City in the speech of newcomers and provide the baseline as reference lects for subsequent developments in Spanish in New York.

The dialectal separation between a region that has higher rates of both overt and preverbal pronouns, and another region that shows lower rates in both measures can usefully be considered in the context of parameterization in formal theories of syntax, specifically with regard to the null subject parameter (Chomsky, 1981). It has been claimed in the generative literature that the null subject parameter is a cluster of properties including the possibility of null subjects and the possibility of postverbal subjects (Chomsky, 1981; Rizzi, 1982; Jaeggli and Safir, 1989). This means that if a language can have null subjects, it can also have postverbal subjects and that if a language cannot have null subjects, then it cannot freely place their subjects postverbally. Spanish, as is well known, conforms well to this characterization, displaying the traits of a null-subject language: its subjects can be null or overt, and when they are overt, they can appear before or after the verb.

The dialectal findings in this project support the idea of a cluster of properties around the null subject parameter in the two dialectal regions studied. In the Mainland, less overt pronouns are found (i.e. more null subjects), and more postverbal subject pronouns than in the Caribbean. On the other hand, in the Caribbean more overt subject pronouns and more preverbal subject pronouns are found. It seems that the more the Caribbean Spanish dialect leans towards behaving like a non-null-subject language, the less it places subjects in postverbal position. If this linguistic behavior of Caribbean speakers represents a case of change in progress, it would be safe to assume that in the distant future

Caribbean speakers would only have overt subjects, and a very limited capacity to place subjects postverbally.

Furthermore, the dialectal separation evidenced by the difference in mean rates of preverbal pronouns between the Caribbean and the Mainland is further confirmed when some of the basic socio-demographic factors are investigated. While in the Caribbean speakers of a higher socio-economic status and higher level of education favor the use of preverbal pronouns, no such role is played by those factors in the Mainland where none of these variables was statistically significant. The findings in use of preverbal pronouns in the Caribbean suggest that placing subject pronouns before verbs is a valued and approved of practice, favored by the higher socio-economic strata in the region.

The results presented by Raña Risso (2013) shed light on the variation in subject pronoun placement witnessed in NYC. She has shown that this variation can be at least partly explained by considering the origin of NYC Spanish-speakers. Moreover, she found that grouping Spanish-speakers in NYC by their immigrant generation (mirroring Otheguy and Zentella's groups) helped explain subject pronoun placement variation in the city. We consider word order in the first and second generation of Spanish-speakers next.

6 Word Order in the First and Second Generation

In view of the findings in the literature summarized above, the next step to investigate is whether the speakers' exposure to NYC, which involves contact with monolingual and bilingual speakers of English, is related to an increase in their use of preverbal pronouns, in the same way that it is connected to an increase in the use of pronouns as shown in Otheguy and Zentella's (2012) study.

Otheguy and Zentella (2012) found support for the hypothesis that English, where subject pronouns are vastly more frequent than in Spanish, is one of the forces motivating an increase in the pronoun rate in the Spanish spoken in NYC. They observed that second generation speakers (those born or raised from early infancy in New York) registered a significantly higher pronoun rate than do first generation speakers (those born in Latin America). They also noted that the speech of bilinguals, including both the Latin American-raised (LAR) established immigrants and the New York raised (NYR), had a significantly higher pronoun rate than that of the LAR newcomers. Furthermore, they found that LAR established immigrants used more pronouns than LAR newcomers, and that in turn, NYR speakers used more pronouns than LAR established immigrants. Regarding the role of English proficiency, they noticed that Latinos with excellent English proficiency used more subject pronouns

than those with lower proficiency in English. They also noted that all these findings applied to the community studied as a whole, as well as when it was divided into Latinos from Caribbean and Mainland origins.

We aim to establish a similar connection between the immigrant generation of the speakers and their rate of preverbal subjects, in particular subject pronouns as well as a connection between their level of English proficiency and their rate of preverbal subjects. In the following sections, we will determine whether groups in our sample (such as the first generation and the second generation) do in fact vary on the average occurrence of the preverbal rates. We will perform an analysis of variance (ANOVA) to this end. Although the general practice in statistics is to use ANOVA for comparing averages between three or more groups and the t-test for comparisons of two groups, we have chosen to follow Otheguy and Zentella (2012) and present all comparisons between averages as ANOVAS, even if only two groups are involved, given that the results with ANOVA are the same as the ones yielded by t-tests. When investigating an external feature (such as generation), the ANOVA takes into account the rates of all individuals in the groups being compared and determines whether differences between the groups (e.g. between first and second generation) are greater than the differences within each of the groups (within the first generation, within the second generation). The higher this between-group to within-group ratio (indicated by the *F* statistic), the more likely the groups are different. For instance, an *F* value of 2.00 indicates that variance across the groups is two times greater than variance within each of the groups. The larger the *F* value, the more chances that the measure of probability (the *p* value) will register at below the conventional .05 level for statistically significant results. To make our tables more accessible to our readers, we have limited the information provided to the number of speakers in the sample analyzed, the preverbal rate, the *F* value, and the *p* value.

6.1 *Latin American Raised Speakers and New York Raised Speakers*

Otheguy and Zentella (2012) found an increase in the use of overt pronouns from the first immigrant generation of Spanish-speakers living in New York (the LAR), and the second immigrant generation of Spanish-speakers (the NYR). We expect a similar increase in the use of both preverbal pronouns and preverbal nominal subjects and for the same reasons: (1) The NYR group has been exposed to English from birth or very early in their lives, and (2) The NYR group has been also exposed to the speech of Caribbeans whose Spanish has higher rates of preverbal subjects than that of Mainlanders. The results for the preverbal pronouns are in Table 4 and those for preverbal nominal subjects with copular verbs are in Table 5.

TABLE 4 ANOVA Preverbal pronominal rate by generation.

	N speakers	Preverbal rate
LAR (1st generation)	114	94
NYR (2nd generation)	25	97
	139	
F (1, 137) = 4.94	p < .02	

TABLE 5 ANOVA Preverbal nominal rate by generation.

	N speakers	Preverbal rate
LAR	25	64
NYR	25	87
	50	
F (1,48) = 14.169	p < .01	

As expected, the NYR group has a significantly higher rate of preverbal subject pronouns than the LAR group. While 94 percent of the verbs in the speech of the first generation appear with a preverbal pronoun, the number increases to 97 percent of verbs with a preverbal pronoun in the speech of the second generation. Variance in the use of preverbal pronouns is almost five times greater ($F = 4.94$) across the generational groups than inside each of the groups, and the four percentage-point difference between the LAR and the NYR registers a high level of statistical significance ($p < .02$).

In Table 5 we present preliminary results for the preverbal nominal rate that support the findings with the preverbal pronominal rate. Second-generation participants have a significantly higher preverbal nominal rate: while only 64 percent of copular verbs in the speech of the first generation appear with a preverbal subject, the number increases to 87 percent in the speech of the second generation. Variance in the use of preverbal subjects with copular verbs is fourteen times greater ($F = 14.169$) across the generational groups than inside each of the groups, and registers a high level of statistical significance ($p < .01$). Both the findings with the preverbal pronominal and nominal rates parallel those for the pronoun rate.

TABLE 6 ANOVA Preverbal pronominal rate by exposure.

	N speakers	Preverbal rate
LAR Immigrant Newcomers	39	94
LAR Established Immigrants	75	94
NYR	25	97
	139	
F (2, 130) = 2.45	p < .09	

6.2 Immigrant Newcomers, Established Immigrants, and the NYR

The next question we raise is whether the process of change is confined to the passage between the immigrant generations or whether instead it begins within the LAR group, among those with more years in the city. To that end, in Table 6 the LAR group is divided into immigrant newcomers (the recent arrivals) and the established immigrants (the rest of the LAR); the table also considers the NYRs to see if there is a continuum in the increase of preverbal placement.

Unlike the results of Otheguy and Zentella (2012) for pronoun rate, we do not see an increase in the preverbal rate with more exposure to NYC within the LAR group. Clearly, the shift to increased preverbality occurs only across the generations and not within the first generation. These results, submitted to a post hoc Tukey test, revealed that LAR Newcomers differed significantly from the NYR ($p < 0.01$) while the LAR established immigrants did not differ significantly from either group.

In comparing the Otheguy and Zentella (2012) and Raña Risso (2013) study, it seems that for the increase in preverbal placement to occur, the increase to more frequent occurrence of pronouns has to take place first. As discussed in the previous section, one way to interpret our findings is with reference to pronominal parametric settings in formal grammars. We speculate that the increase in overt pronouns is affecting the null subject parameter in the speakers' native language. One of the characteristics of the null subject parameter is the ability to place subjects postverbally, i.e. languages that allow subjects to be null also allow subjects to be placed after the verb. Conversely, it seems that if the speakers use fewer null pronouns, they will also use fewer postverbal pronouns. However, given the results above, it seems to be the case that the shift to fewer postverbal pronouns occurs after the shift to fewer null pronouns. In this view of the matter, the placement change may be a consequence of the

frequency change, and could be seen as evidence toward a parametric change in the speakers' grammar.

6.3 *English Proficiency Groups: English Excellent and English Less than Excellent*

We have seen above that although no changes with regard to the placement of subjects seem to occur in the first generation, the second generation, born and/or raised in NYC and bilingual, exhibits a statistically significant increase in the preverbal rate for subjects which mirrors English grammar. These results seem to indicate that those who are born or raised with Spanish and English, in a medium where English has a higher status and is prevalent, find their grammar of Spanish affected by their grammar of English. Consequently, we wanted to know whether this is just an effect of the environment and the higher status of one of the languages, or whether it is the case that, independently of where the speakers are born or raised, those who claim to have a high proficiency in English usually exhibit more preverbal subjects than those who claim to have lower or no proficiency. For the purpose of studying the effect of higher proficiency in English on the Spanish spoken in NYC, the sample is divided into two groups: those who claim to speak excellent English and everybody else. If excellent English speakers have a higher preverbal subject rate, the language contact hypothesis would be supported. The results of this investigation are presented in Table 7 for nominal subjects and Table 8 for pronominal subjects below. Also, the same division is carried out in a subset of the sample of speakers: the LAR. The goal of this inquiry is to establish whether excellent English skills (or fluent bilingualism) are a driving force behind the changes in Spanish grammar, independently of where the speakers were born and raised, that is independently of generation, exposure, etc. The results of this inquiry are presented in Table 9 below.

Both investigations yielded the expected results: speakers who claim to have excellent English proficiency also have more preverbal subjects in Spanish. Variance in the use of preverbal pronouns is more than eight times greater

TABLE 7 ANOVA Preverbal nominal rate by English skills.

	N speakers	Preverbal rate
Eng. Less than excellent	18	67
Eng. Excellent	32	80
	50	
F (1,48) = 4.11	p < .05	

TABLE 8 ANOVA Preverbal pronominal rate by English skills.

	N speakers	Preverbal rate
Eng. Less than excellent	94	94
Eng. Excellent	44	96
	138	
F (1,136) = 8.38	p < .001	

TABLE 9 ANOVA Preverbal pronominal rate by English skills LAR only.

	N speakers	Preverbal rate
Eng. Less than excellent	89	94
Eng. Excellent	24	96
	113	
F (1, 111) = 4.28	p < .04	

(F = 8.38) and for nominal subjects more than four times greater (F = 4.11) across the English proficiency groups than inside each of the groups, and registers a high level of statistical significance. The same holds true pronominally when we subdivide the sample to just contain LAR speakers and we run the same test. Speakers who claimed to have a high proficiency in English also had more preverbal pronouns, and the difference between these two groups was more than four times greater (4.28) and highly significant ($p < .02$). These results suggest very strongly that there is language contact between English and Spanish at work, and that English is one of the factors shaping the Spanish spoken in NYC.

6.4 Differences of Preverbal Rate in the Regional Subsamples

In the preceding investigations for the whole sample, it was discovered that preverbal subject rates allowed for the grouping of speakers by generation and English proficiency. Given the disparity between the findings for the preverbal subject rate and those for the use of subject pronouns, we have offered the explanation that for changes in preverbal subject rates to take place, changes in pronoun rates have to happen first, which would be aligned with the generativists' hypothesis of the null subject parameter as encompassing both the possibility of dropping subjects and the possibility of placing subjects after the verb. In this section, we will be looking at the same grouping possibilities

after subdividing the sample into the two regional groups whose validity was established by Otheguy and Zentella (2012) and Raña Risso (2013), namely Caribbeans and Mainlanders.

As exposure increases among Caribbeans from immigrant newcomers to established immigrants to NYR, the preverbal pronominal rate does show an increase. This positive finding means that, among Caribbeans, the preverbal pronominal rate differences associated with exposure differ from the trend of the whole sample (see Table 10).

TABLE 10 ANOVA preverbal pronominal rate by exposure among Caribbeans.

	N speakers	Preverbal rate
LAR Immigrant Newcomers	19	95
LAR Established Immigrants	40	97
NYR	13	98
	72	
F (2, 69) = 2.49	p < .09	

Although the results for Table 10 only approach significance, we submitted them to a post hoc Tukey test, which revealed that the differences between Caribbean LAR Newcomers and Caribbean NYR approached significance ($p = 0.07$) while the Caribbean LAR established immigrants did not differ significantly from either group. The results of Table 10 above support the language contact hypothesis among Caribbeans, and unlike the results for the whole sample, where we saw that contact was evident only in the second generation, we can see changes taking place in the first generation, with a two-point increase in the preverbal rate from immigrant newcomers to established immigrants. These findings parallel those of Otheguy and Zentella (2012) for the pronoun rate in the same subsample of speakers.

TABLE 11 ANOVA preverbal pronominal rate by English skills among Caribbeans.

	N speakers	Preverbal rate
Eng. Less than Excellent	45	95
Eng. Excellent	26	99
	71	
F (1, 69) = 14.5	p < .001	

The differences in pronoun rates of Caribbeans with different English abilities are as expected, as shown in Table 11, and they also move in the same direction as in the whole sample, yielding strong differences with clear statistical significance.

Among Caribbeans whose English competence is high, the ability to place subjects postverbally has almost disappeared. Furthermore, the difference between these two groups was more than fourteen times greater (14.5) and highly significant ($p < .001$). These findings provide strong evidence of language contact between English and Spanish in this subset of speakers. Furthermore, they are aligned with the findings of Otheguy and Zentella (2012) for the pronoun rate.

TABLE 12 ANOVA Preverbal pronominal rate by exposure among Mainlanders.

	N speakers	Preverbal rate
LAR immigrant newcomers	20	93
LAR established immigrants	35	91
NYR	12	95
	67	
$F(1, 65) = 2.14$	$p < .12$	

Preverbal pronominal rate differences for Mainlanders of different exposure groups do not follow the same pattern as for Caribbeans, since there is not an increase within the first generation of speakers (See Table 12.).

As exposure increases from Mainlander immigrant newcomers to established immigrants, the preverbal rate decreases. Then, it increases again with the NYR group to surpass that of LAR immigrant newcomers. The big increase for Mainlanders is inter-generational and not intra-generational. Regarding the decrease in preverbal rates in the LAR established immigrant group, it is interesting to note that 15 out the 35 speakers in this group registered preverbal rates that were much lower than the average rate of the whole group (from 71 to 89) and that they came mostly from Mexico, but there were also two speakers from Colombia and four from Ecuador with preverbal rates in the 71-to-92 range, which is below the mean rate for newcomers from that region, which is 93. It could be the case that LAR established immigrants from the Mainland make an unconscious attempt to differentiate themselves from Caribbean immigrants and English speakers and that this translates into placing more subjects postverbally. This behavior would be similar to the one reported by Labov (1963) in Martha's Vineyard, where the need of a group to assert its unique identity

translated into an increase in the use of the phonological patterns typically associated with the speech of local fishermen families. In that study, Labov showed that language (or dialect) contact situations do not necessarily lead to a melting-pot linguistic outcome, but rather to an increase in the differences that distinguish each group.

The differences in preverbal rate between Mainlanders with different levels of English-proficiency follow the same pattern as those of the whole sample and the Caribbean group. However, there is only one percentage point difference between Mainland speakers with high and low English skills, and the results lack statistical significance ($F(1,65) = .40, p < .52$).

In subdividing the whole sample into two subsets by region, we discovered that the factors of exposure and English proficiency affect the preverbal rate differently in each region. Whereas in the Caribbean more exposure to the U.S. and more fluency in English clearly affect the placement of pronominal subjects, the same does not hold true in the Mainland. For Mainland speakers, the increase takes place in the second generation but not in the first, and an excellent level of English fluency does not seem to be a clear indicator of increased preverbal placement of pronouns. The fact that each region behaves differently for the two-mentioned factors with regards to the preverbal rate supports the division of speakers from the two regions into two groups. Furthermore, the behavior of both regions supports the language contact hypothesis, although in one group more strongly than in the other.

Now that we know that subject placement and subject expression are related, and that external variables such as immigrant generation and English proficiency play a significant role in guiding speaker decisions on the placement of subjects, there are several avenues for future research. First, we need to examine how internal linguistic variables affect nominal and pronominal subject placement, and how much variance is accounted for by these variables in our groups of speakers. In other words, we would like to determine if linguistic variables such as pragmatic function, constituent length, and verbal transitivity, among others, play the same or different roles across regions and generations, and if they lend further support to the conclusions presented in this article. [A1] Similarly, we need determine whether we can analyze our data using a mixed-effects model with speaker as random intercept. This would allow us to compare both internal and external variables and their effect on word order.

7 Summary and Discussion

We investigated whether the placement of pronominal and nominal subjects with regard to the verb can lend support to a language contact hypothesis for

Spanish in NYC as the occurrence of pronominal subjects did in Otheguy and Zentella (2012). We believe that the answer is affirmative, and that there is a strong influence of English on the placement of subjects by speakers of Spanish in NYC.

The comparison between Otheguy and Zentella's (2012) and Raña Risso's (2013) studies demonstrates that the change towards more overt pronouns leads to an increase in preverbal pronouns. The Null Subject Parameter can be used to connect the two phenomena from the two investigations, given that it posits that languages that can drop subjects, can also freely place them before or after the verb. The results of both investigations clearly show that as groups of speakers move from a pro-drop language towards a non-pro-drop variety of that language, they begin to lose the ability to place subjects post-verbally. Weerman (1998) proposed that adult speakers have peripheral rules available to them, but not the possibility to re-set parameters, which is only available to children acquiring their first language. With that view in mind, we can interpret the results of our investigation as follows: the increased use of overt subject pronouns exhibited by the first generation of (adult) speakers because of the contact with English, responds to a new peripheral rule created by those speakers. This, in turn, provides input to NYR Latinos that is different from the input that their first-generation counterparts were exposed to (Otheguy, 2016; Pires and Rothman, 2009). If Spanish were to survive into the third and fourth generations, it is possible that those speakers' would produce input for their children that is so rich in overt preverbal subject pronouns and preverbal nominal subjects that it could trigger a parametric change in the next generation of Spanish speakers. Raña Risso (2013) proposed a Threshold Hypothesis to explain this further, whereby once the threshold is crossed for the pronoun rate, significant changes in the preverbal rate begin. Following this argument, the data can be explained in terms of the differential input for acquisition of Spanish received by the first and the second generation of speakers. The first generation received input in Latin America, which contained less overt pronouns. On the other hand, the second generation received input in NYC, which contained more overt pronouns as the first generation had significantly increased the pronoun rate. In turn, the second generation sets the parameters for the null subject differently. Therefore, we interpret the more rigid word order exhibited by second-generation speakers in our corpus as a consequence of a move toward parametric changes in their internal grammar. We propose that second-generation bilinguals do not have an *incomplete* grammar of Spanish but rather a *different* grammar than their first generation counterparts. The generational differences in nominal and pronominal subject placement found in our data are not evidence of incomplete acquisition of Spanish, but rather evidence of contact-induced diachronic change.

References

- Arnold, Jennifer, Anthony Losongco, Thomas Wasow, and Ryan Ginstrom. 2000. Heaviness versus newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76: 28–55.
- Barrera-Tobón, Carolina. 2013. *Contact-Induced changes in word order and Intonation in the Spanish of New York City Bilinguals*. Ph.D. dissertation. The Graduate Center of the City University of New York, New York: ProQuest/UMI.
- Barrera-Tobón, Carolina and Rocío Raña Risso. 2016. A corpus-based sociolinguistic study of contact-induced changes in word order in the Spanish of New York City bilinguals. In Sandro Sessarego and Fernando Tejedo-Herrero (eds.), *Spanish Language and Sociolinguistic Analysis*, 323–342. John Benjamins.
- Bergad, Laird. 2011. The Latino Population of New York City, 1990–2010. *The CUNY Center for Latin American, Caribbean, and Latino Studies' Latino Data Project Report 44*.
- Bentivoglio, Paula and Mercedes Sedano. 2000. El sujeto liviano: Una restricción de tipo funcional. *Boletín de Filología* (Universidad de Chile) 38: 9–22.
- Bentivoglio, Paula. 2003. Orden de palabras en español: Un análisis sintáctico-semántico-pragmático del sujeto. *Lexis* 1–2: 235–260.
- Butt, John and Carmen Benjamin. 1996. *A New Reference Grammar of Modern Spanish*. London: Edward Arnold / Lincolnwood, Ill.: NTC. Second Edition.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.
- Givón, Thomas. 1993. *English Grammar: A Function-Based Introduction*. Amsterdam / Philadelphia: John Benjamins Publishing Co.
- Goodall, Grant. 2004. On the Syntax and Processing of Wh-questions in Spanish. In Benjamin Schmeiser, Vineeta Chand, Ann Kelleher and Angelo Rodriguez (eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Press.
- Hawkins, John A. 1983. *Word Order Universals*. New York, NY: Academic Press.
- Jaeggli, Osvaldo and Kenneth J. Safir (eds.) 1989. *The null subject parameter*. Dordrecht: Kluwer.
- Labov, William. 1963. The social motivation of a sound change. *Word* 19: 273–309.
- Lipski, John. 1994. *Latin American Spanish*. London: Longmans.
- López-Morales, Humberto. 1992. *El Español del Caribe*. Madrid: Editorial MAPFRE.
- Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: a case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7: 125–142.
- Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam: John Benjamins.

- Morales, Amparo. 1988. Hacia un universal sintáctico del español del Caribe: El orden SVO. *Anuario de Lingüística Hispánica* 5: 139–152.
- Nava, Erika. H. 2007. Word order in bilingual Spanish: Convergence and Intonation Strategy. In Jonathan Holmquist, Augusto Lorenzino and Lotfi Sayahi (eds.), *Selected Proceedings of the 3rd Workshop on Spanish sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project.
- Ocampo, Francisco. 1995. The word order of constructions with a verb, a subject, and a direct object in spoken Spanish. In John Amastae, Grant Goodall, Mario Montalbetti and Marianne Phinney (eds.), *Contemporary Research in Romance Linguistics*. Amsterdam / Philadelphia: John Benjamins.
- Ocampo, Francisco. 2002. The word order of constructions with *ser* and *estar*, a subject NP, and an adjective in spoken Spanish. In James Lee, Kimberly Geeslin and Clancy Clemens (eds.), *Structure, Meaning, and Acquisition in Spanish: Papers from the 4th Hispanic Linguistics Symposium*.
- Ordóñez, Francisco and Antxon Olarrea. 2001. Weak Subject Pronouns in Caribbean Spanish and xP Pied-Piping. In Julia Herschensohn, Enrique Mallen, and Karen Zagona (eds.), *Features and Interfaces in Romance: Essays in Honor of Heles Contre-ras*, Amsterdam: John Benjamins, 223–238.
- Otheguy, Ricardo. 2016. The Linguistic Competence of Second-Generation Bilinguals: A Critique of ‘Incomplete Acquisition’. *Selected Proceedings from the 43rd Linguistic Symposium on Romance Languages*.
- Otheguy, Ricardo and Ana Celia Zentella. 2012. *Spanish in New York: Language Contact, Dialect Leveling, and Structural Continuity*. New York, NY: Oxford University Press.
- Pires, Acrisio and Jason Rothman. 2009. Disentangling sources of incomplete acquisition: An Explanation for competence divergence across heritage grammars. *International Journal of Bilingualism* 13(2): 211–39.
- Potowski, Kimberly and Lourdes Torres. (forthcoming). *Spanish in Chicago: Dialect Contact and Language Socialization among Mexicans and Puerto Ricans*.
- Prince, Ellen. 1981. Toward a taxonomy of given-new information. In Peter Cole (ed.), *Radical Pragmatics*. New York: Academic Press.
- Quirk, Randolph, Sydney Greenbaum, Geoffrey Leech, and Jan Svartik. 1972. *A Grammar of Contemporary English*. Essex: Longman.
- Raña Risso, Rocío. 2010. Subject Pronoun Placement as Evidence of Contact and Leveling in Spanish in New York. *International Journal of the Sociology of Language* 203. 101–114.
- Raña Risso, Rocío. 2013. *A Corpus-based Sociolinguistic Study of Subject Pronoun Placement in Spanish in New York*. Ph.D. dissertation. The Graduate Center of the City University of New York, New York: ProQuest/UMI.

- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris Publications.
- Sankoff, Gillian. 2002. Linguistic outcomes of language contact. In Jack. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, 638–669. Malden, MA: Blackwell Publishers.
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Oxford University Press.
- Silva-Corvalán, Carmen (ed.). 1995. The study of language contact: An overview of the issues. *Spanish in Four Continents: Studies in Language Contact and Bilingualism*. Washington DC: Georgetown University Press.
- Silva-Corvalán, Carmen. 2001. Lenguas en contacto y bilingüismo – el español en los EE.UU. In *Sociolingüística y Pragmática del Español*. Washington, DC: Georgetown University Press.
- Sorace, Antonella and Ludovica Serratrice. 2009. Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism* 13(2): 195–210.
- Swan, Michael. 2005. *Practical English Usage*. Oxford: Oxford University Press.
- Taverniers, Miriam. 2005. Subjecthood and the notion of instantiation. *Language Sciences* 27: 651–678.
- Thomason, Sarah. 1995. Language mixture: Ordinary processes, Extraordinary Results. In Carmen Silva-Corvalán (ed.), *Spanish in four continents: Studies in Language Contact and Bilingualism*. Washington DC: Georgetown University Press.
- Thomason, Sarah. 2001. *Language Contact: An Introduction*. Washington DC: Georgetown University Press.
- Thomason, Sarah and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Toribio, Almeida Jacqueline. 2000. Setting parametric limits on dialectal variation in Spanish. *Lingua* 10: 315–41.
- Toribio, Almeida Jacqueline. 2004. Convergence as an optimization strategy in bilingual speech: Evidence from code-switching. *Bilingualism: Language and Cognition* 7: 165–173.
- Tsimpili, Ianthi and Antonella Sorace. 2006. Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. *BUCLD Proceedings* 30: 653–664.
- U.S. Census Bureau. 2014, July 8. *State and County Quickfacts: New York, N.Y.* Retrieved October 29, 2014, from <http://quickfacts.census.gov>.
- van Hout, Roeland and Pieter Muysken. 1994. Modeling lexical borrowability. *Language Variation and Change* 6(1): 39–62.
- Weerman, Fred. 1993. The diachronic consequences of first and second language acquisition: the change from OV to VO. *Linguistics* 31: 903–931.

- Zagona, Karen. 2002. *The Syntax of Spanish*. New York, NY: Cambridge University Press.
- Zapata, Gabriela, Liliana Sánchez, and Almeida Jacqueline Toribio. 2005. Contact and contracting Spanish. *International Journal of Bilingualism* 9: 377–395.
- Zubizarreta, María Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.