Book chapter for Inquiries in Hispanic Contact Linguistics: Theoretical, Methodological and Empirical Perspectives

6,958 words

First person singular subject expression in Caribbean heritage speaker Spanish oral production

Abstract

Research on contact effects on subject pronoun expression in the U.S. has largely focused on Spanish in the U.S. Southwest and NYC. The conflicting results reported in this literature could be due to the differences in the distribution of varieties in these U.S. regions. To test this hypothesis, this paper offers a variationist analysis of first person singular data from eleven Caribbean heritage speakers in Florida, where the distribution of varieties is similar to that in NYC. Results are consistent with a lack of contact effects in the first person singular productions of these bilinguals, similar to the reports on the U.S. Southwest.

1. Introduction

Variationist and generativist examination of Spanish HS subject pronoun expression (SPE) reports an increase in overt third person pronominal subjects, largely attributed contact with English, irrespective of data collection method (e.g., Otheguy & Zentella 2012, Montrul 2004). Results for first person, however, are less conclusive (Torres-Cacoullos & Travis 2010, Travis 2007). It is possible that some of the differences have to do with the region in the U.S. where the data is collected: NYC vs. the Southwest. In this paper we discuss first person singular interview data from Caribbean Spanish HSs from Florida to contribute novel data to the debate of language contact in Spanish in the U.S.

In addition to person, the variationist literature indicates that Spanish SPE is restricted by a variety of linguistic factors (speech connectivity, verb form ambiguity, clause type, and semantic verb type, etc.), as well as extralinguistic factors, such as region and language contact (e.g. Carvalho, Orozco & Lapidus Shin 2015, Otheguy & Zentella 2012, Shin 2012, Silva-Corvalán 1982). Contact effects have variably been reported both in terms of SPE rates, significance of variables, and direction of effect. In this paper we further examine the effect of proficiency, a measure that has not been examined directly in previous research on oral production. All our speakers have been raised in the U.S. and consider themselves dominant in English. Additionally, they are all of Caribbean heritage. Thus, with these data we can focus on the effect of proficiency.

The aim of this paper is twofold. It examines subject pronoun expression in Florida to further understand the effect that the demographic/dialectal distribution of the communities may have on the variable expression of subject pronouns in bilingual Spanish. Additionally, it examines the role that proficiency in the heritage language has on subject pronoun expression in Spanish. With this purpose in mind, the paper is organized as follows. Section 2 summarizes previous relevant work on subject pronoun expression. Section 3 offers the motivation for the study, research questions and hypotheses. Section 4 describes the participants and research design and presents the results. Section 5 discusses the results. Lastly, section 6 offers some conclusions.

2. Literature Review

Variationist analyses of Spanish subject expression indicate that the distribution of null and overt pronominal subjects in Spanish is best explained by a combination of variables, except in those cases traditionally excluded from variable rule analyses: predicates that require an expletive subject, predicates accompanying impersonal uses of the second person singular and third person plural, reverse psychological predicates, predicates in subject relative clauses, subjects with inanimate referents, and predicates in set phrases. In each case, the null pronominal subject fails to alternate with an overt counterpart.

In variable contexts, on the other hand, the distribution has been best accounted for by a combination of variables. In general terms, null subjects tend to indicate continuity. Thus, CO-REFERENTIALITY (also referred to as switch reference), or whether the subject in the preceding sentence is the same or not, favors the use of a null subject.

(1) CO-REFERENCE OR SWITCH REFERENCE

Y yo los bañaba, y los vestía, les daba de comer, los ponía a dormir. 'And I would bathe them, dress them, feed them, put them to sleep.' [NMCOSS, 117–1A3: 248.-2512] (Travis 2007)

In (1) the pronominal subject is expressed in the first instance and omitted afterwards where it is co-referential. Similarly, TENSE, ASPECT, MOOD (TAM) CONTINUITY favors null subjects, as exemplified in (2).

(2) Coding of language-internal variables: Continuity TAM CONTINUITY

> *Mañana voy. Yo dejé diez paquetes allá.* 'I will go tomorrow. I left ten packets there.' [Colombia, cooking: 100-101] (Travis 2007)

The subject in (2) is null in the first clause and expressed in the second, where there is a change in TAM. Some of the verb forms in Spanish, namely the first and third person singular forms, are the same in several tenses (imperfect, conditional, present subjunctive and related compound forms). These ambiguous forms have sometimes been attested with more overt pronominal subjects, with disambiguating purposes. Consider the following example:

(3) Coding of language-internal variables: VERB FORM AMBIGUITY En la noche ella iba a mi lado y yo estaba temblando

'At night she used to go by my side and I wasn't shaking.' (Silva-Corvalán 1994)

In (3), both verbs could refer to a first or a third person singular. In this case, to disambiguate, both appear with an overt pronominal subject. While overt subjects are favored with verbal forms that are *ambiguous* and not with *unambiguous* ones (Silva-Corvalán 1994) this result is not always attested (Casanova 1999, Morales 1997, Ranson 1991). Unlike in example (3), often times there is enough information in the context for the referent not to be ambiguous even if the verb form is so, as in (2) above. The distribution of overt subjects is also relevant to the establishment of the speaker's position on an idea. As a result, SEMANTIC VERB TYPE affects the distribution. For instance, verbs that express opinion or estimative verbs favor overt subjects (Enríquez 1984, Morales 1997, Otheguy et al. 2007, Silva-Corvalán 1982, 1994, Travis 2007). For example, Morales (1996) shows that the subjects of verbs like *pensar* 'to think' may be produced even in topic continuation contexts, as in (4).

(4) Coding of language-internal variables: VERB TYPE
 Parece que ellos piensan que es signo de cultura.
 'It seems as if they think that is it a sign of culture.' (Morales 1996)

More recently the classification of verb type has been reduced to external actions, mental processes and stative verbs (see Orozco 2015 for a fuller description). In general, mental and stative predicates favor null subjects while external actions favor overt pronominal subjects.

These factors (cf. Carvalho et al. 2015 for a more complete list of factors) do not all affect the distribution equally. Switch reference, for instance, is significant and highly ranked (i.e., larger magnitude of effect) across studies while verb form ambiguity tends to be lower ranked or not significant at all. Subject pronoun rates, on the other hand, vary significantly from region to region. Research on monolingual varieties in the Caribbean report significant variation from study to study (Alfaraz 2015, Bentivoglio 1987, Cameron 1995, Cameron & Flores-Ferrán 2004, Enríquez 1984, Erker & Guy 2012, Flores-Ferrán 2002, 2004, Hochberg 2986, Holmquist 2012, Martínez-Sanz 2011, Morales 1982, Orozco 2015, Orozco and Guy 2008, Ortiz López 2009, 2011, Otheguy & Zentella 2012, Shin & Otheguy 2013). The rates reported for monolingual Caribbean Spanish span from 28% in Castañer, Puerto Rico (Holmquist 2012) to 60% in San Juan, Puerto Rico (Cameron 1995). Importantly, although dialectal variation is largely attested in subject pronoun rates and restricted to lower ranked variables (those with a smaller effect size), these differences do not affect most linguistic variables. That is, this combination of variables is rather stable across varieties of Spanish.

A comparison in subject pronoun expression that has generated great interest in the literature has been between monolingual and bilingual varieties of Spanish. Spanish in contact with English has reported conflicting results: while some do not report differences (Bayley and Pease-Alvarez 1996, 1997, Flores and Toro 2000, Flores-Ferrán 2004, Silva-Corvalán 1994, Torres-Cacoullos & Travis 2010, and Travis 2007), others report an increase in overt pronominal subjects and changes in the distribution reflected in variables that are significant, their ranking, and/or their constraint ranking (Erker & Guy 2013, Erker & Otheguy 2016, Lipski 1994, 1996, Otheguy & Zentella 2012, Shin 2012, Shin & Otheguy 2013, Toribio 2004, among others). It is possible that these differences are due to the grammatical person examined in these studies (cf. Author, submitted) or differences in the communities. The data in studies that do not report a contact effect are primarily from Mexican varieties in the U.S. Southwest (with the exception of Flores and Toro 2000, Flores-Ferrán 2004) while the data in studies that report a contact effect are from several varieties of Spanish or Caribbean Spanish in NYC (with the exception of Toribio 2004). The demographics of the U.S. Southwest are rather different than those in NYC. According to the 2010 Census, in New Mexico 46.3% of the population is Hispanic while in NYC 22.5% of the population is Hispanic. Within the Hispanic population, only 1.44% of the Hispanic population in New Mexico is of Caribbean origin (including Cuba, Dominican Republic, Puerto Rico, and Venezuela) while it is 57.8% in NYC. In this paper we examine data from Florida, whose Hispanic population is 22.5% of the total population. Within the Hispanic population, 55.3% is of Caribbean origin. Thus, in this paper we examine a novel community (Caribbean speakers in Florida), which is similar to NYC in the distribution of varieties to determine if the demographics, and the resulting dialect contact, can shed some light on the conflicting previous results on SPE in U.S. Spanish.

3. Motivation and research questions

The goal of this study is to explore subject expression in first person singular (*yo* 'I'/Ø) through spontaneous speech samples from Spanish HSs of Caribbean heritage living in Florida in order to provide novel data to better understand differences in varieties of U.S. Spanish. Similarly to Otheguy and Zentella's (2012) contact situation NYC, in Florida there are speakers of Caribbean and non-Caribbean Spanish with a slightly higher population of Caribbean speakers. In New Mexico, in contrast, the Caribbean population constitutes only 1.44% of the Hispanic population. While Otheguy & Zentella (2012) report contact effects in terms of rates of overt pronominal subjects and the strength of the variables regulating the distribution, Torres-Cacoullos & Travis (2010) do not. Since Torres-Cacoullos & Travis (2010) only included first person singular subjects in order to be able to make comparisons without added variability.

Our first research question examines rates of overt pronominal expression and the effect that language contact has on it. It is well established in the literature that monolingual Caribbean speakers produce more overt pronominal subjects than monolingual speakers from other Spanish-speaking countries. Irrespective of region of origin, Otheguy & Zentella (2012) report an increase in rates of overt pronouns in speakers born and raised in NYC as compared to those born in a Spanish-speaking country who migrated to the U.S. in adulthood and those that were born in a Spanish-speaking country and who are more established in the U.S., while Torres-Cacoullos & Travis (2010) do not find this effect in their comparison of bilinguals in NM with different degrees of English proficiency (and presence of code-switching in their speech). We divided our participants, who were all raised in the U.S. (and who were either born in the U.S. or migrated to the U.S. by age 11) and dominant in English, according to their proficiency in Spanish, as measured by the version of the DELE commonly used in the field of SLA. Since our participants are of Caribbean heritage and the linguistic composition of Florida is more similar to that in NYC, we anticipate our participants to show an effect of language contact. All speakers are expected to exhibit higher rates than those reported in the literature on Caribbean monolingual varieties but more so in the case of lower proficiency speakers.

Our second research question inquires about the patterns of subject expression across these two groups. Torres-Cacoullos & Travis (2010) report minor differences between those speakers with higher English proficiency and those with lower English proficiency. Similarly, Otheguy & Zentella (2012) report similarities in the variables that are significant, with minor differences. Specifically, those groups that are more established in the U.S. exhibit a weakening of the effect of some variables and some differences in the direction of effect in the variable Person. Since we only examine first person singular subjects in this paper, we anticipate little difference in the variables that regulate subject expression and the direction in both groups of speakers. Differences, however, may exist in the magnitude of effect of these variables, as in Otheguy & Zentella (2012). Regarding the variables included in this study, then, the anticipated direction of effect (if there is an effect) is for (i) contexts of different reference to favor overt pronominal subjects more than those of same reference; (ii) continuity of TAM to favor pronominal omission; (iii) mental and stative predicates to favor pronominal omission; and (iv) ambiguous verb forms to favor overt pronominal subjects. In order to test these hypotheses we analyzed data from sociolinguistic interviews carried out in Spanish with speakers of Caribbean heritage in Florida.

4. The current study

4.1 Participants

Eleven participants of Caribbean heritage volunteered for this study. All participants were instructed HSs enrolled in university Spanish courses at the time of the study. There were 10 females and 1 male. There were five participants of Cuban heritage, three of Puerto Rican heritage, two of Venezuelan heritage and one of Dominican heritage. As an independent measure of proficiency, we utilized a customized version of the Diploma de español como lengua extanjera (DELE) commonly used in the field of second language acquisition and heritage speaker bilingualism (Montrul & Slabakova, 2003; Cuza, 2013) and presented online through qualtrics. The test consists of 50 questions divided in two different sections: a multiple-choice vocabulary section and a cloze passage that tests grammatical knowledge. The mean score was 27 (range: 14–39). We calculated the median (29/50) and split the participants in two groups: the advanced group, who scored above 29 (n=5) and the non-advanced, who scored 29 or below (n=6). Although none of the participants scored within the range commonly referred to as advanced in the literature (40-50), most of these were perceived as native-like by the author.

They belonged to different generations: two were born in Puerto Rico and moved to the Mainland U.S. before age 11 (Generation 1.5), three were second-generation HSs, two were born in the U.S. with one parent born in a Spanish-speaking country and one parent who was a second-generation HS (Generation 2.5), and four third-generation HSs. The following table summarizes this information.

Generation	1.	5	2		2.5		3	
Proficiency	Advanced	Non	Advanced	Non	Advanced	Non	Advanced	Non
-		advanced		advanced		advanced		advanced
Number of	2	0	1	2	0	2	1	3
participants								

Table 1: Participant information

Given the diversity attested in the participant pool, in addition to group results we will offer some information on individual data.

4.2 Methodology

We conducted 11 semi-structured sociolinguistic interviews. The interviews lasted between 30 to 60 minutes and were always recorded in a quiet office. All interviews were digitally recoded using a Marantz PDM 661 and a Shure SM 10-A head-mounted microphone. The interviewers used a Power Point with questions to guide the interview and to assure that all participants spoke about similar topics and in similar tenses. Participants were told that they should consider the interview as an informal conversation and that we would simply be talking about some events in their past, present and future. The interviewer initiated the conversation asking participants to talk about their birth city and family. Then, participants were asked to talk about their childhood friends and teachers, and to describe interesting events from their childhood. Next, participants

were asked to talk about their current experience at the university (classes, friends, teachers) as well as about their current jobs. Finally, the interview focused on future events by asking participants about their plans for the next holiday (e.g., Thanksgiving/Spring break), their plans for their future career and, finally, participants had to describe life in the year three thousand.

The interview data was transcribed and coded for several linguistic variables: SUBJECT FORM, SWITCH REFERENCE, TAM CONTINUITY, VERB AMBIGUITY, and VERB TYPE. We also coded two extra-linguistic variables: PROFICIENCY and GENERATION. Only tokens in first person singular, topic continuation, main clauses, and with animate referents were included (with a total of 2,054 tokens). Table 2 describes the variable and constraints considered for analysis.

Variable type	Variable	Constraint
Linguistic variables	SUBJECT FORM	Null Overt Pronominal
	SWITCH REFERENCE	Same referent Different referent
	TAM CONTINUITY	Same TAM Different TAM
	VERB TYPE	State Mental External
Extralinguistic variables	PROFICIENCY	Advanced Non-advanced
	GENERATION	1.5 Second 2.5 Third

Table 2: Linguistic and extralinguistic variables and associated constraints

Next, we illustrate the above constraints from examples from the interviews:

(5) SUBJECT FORM

a. Overt Pronominal

<u>Yo</u> nací en Plantation en la Florida, en el sur de la Florida. Entonces por ahí todos hablan español.

'I was born in Pantation, in Florida, in South Florida. So, everyone speaks Spanish over there.' b. Null

 $\underline{\mathcal{O}}$ Entonces **tengo** a dos hermanos que también hablan español y aprendimos español de mi mamá.

'I have two brothers who also speak Spanish and we learned Spanish from my mom.'

In (5a), the speaker produces the first person singular pronoun *yo*, while in (5b) she does not. This distribution tends to be regulated by switch reference.

(6) SWITCH REFERENCE

a. Same referent *Tenía clase en el college o me quedaba en la escuela*'I either had class at the college or I stayed in school.'

b. Different referent *Mi quinto periodo era history creo que era y después mi séptimo periodo*'My fifth period was history, I think that is what it was, and then I had my seventh period.'

As shown in (6), the speaker uses a null subject with *me quedaba* in (a) whose subject has the same referent as the previous verb (*tenía*), and with *creo* in (b), even though the previous subject has a different referent. Another variable that expresses speech connectivity is TAM continuity:

(7) TAM CONTINUITY

a. Same TAM No dormí mucho esos cuatro años pero **aprendí** como escribir ensayos. 'I didn't sleep much those four years but I learned how to write essays.'

b. Different TAM *Como yo hice dual enrollment no tenía clase* 'Since I did dual enrollment I didn't have class.'

As can be seen in (7), the speaker omits the subjects with *aprendi*, which is in the same TAM as the previous verb, *dormi*. Some previous literature has reported the use of overt pronominal subjects to disambiguate.

(8) VERB AMBIGUITY

a. Non-ambiguous *Me identifico mucho como Cubano cuando voy a Miami.*'I identify a lot as a Cuban when I go to Miami.'

b. Ambiguous*Les dije que era puertorriqueño.*'I told them I was Puerto Rican.'

In (8), the speaker uses a null subject both with the morphologically unambiguous verb form *voy* and with the ambiguous verb form *era*. Lastly, speakers tend to use more overt pronominal subjects with external action predicates than mental or stative predicates.

(9) VERB TYPE

a. ExternalSo *fui* para casa siempre y con el bus.'So I always went home and on the bus.'

b. Stative

Y leí, *vi televisión, leí más. No me gusta jugar deportes y nunca hací eso.* 'And I read, watched tv, read some more. I don't like playing sports and I never did that.'

c. Mental

Ella, mi mamá, no **quería** que yo guiaba todos los días porque era muy temprano y estaba muy cansado y ella se preocupó.

"As to her, my mom, she didn't want me to drive every day because it was very early and I was very tired and she got worried."

In (9), the speaker uses null subjects with external actions (*fui*) as with stative (*lei*) and mental (*queria*) predicates. All first person singular tokens with animate referents in main clauses were coded and included in the analysis.

4.3 Results

The purpose of this study was to examine the distribution of overt pronominal subjects in the Spanish of Caribbean HSs in Florida. The group of speakers belonged to different generations and proficiency levels. There was collinearity between the variables generation and proficiency, such that all speakers in generation 1.5 were advanced and none of speakers in generation 2.5 were advanced. A separate run for the variable generation revealed that it is a significant variable.

Table 3: First person singular subject expression in Caribbean Spanish-English bilinguals in Florida: Generation

Multivariate regression analysis of the contribution of external factors to the probability of producing a first person singular overt pronominal subject vs. a null subject; factor groups selected as significant in gray background

sciected as significant in g	ay background			
	Factor weight	%	Ν	% of data
GENERATION				
2.5	0.78	51.2	196	18.6
Second	0.69	39.8	274	33.5
Third	0.29	10.6	51	23.5
1.5	0.23	8	40	24.3
Range	55			
Total N			2,054	
Corrected Mean			.227	
Log likelihood			-1030.585	
Significance			.000	

The different behavior for generation 1.5 and the second generation as well as between the generation 2.5 and the third generation prevented the recoding into two generations. Generation 2.5 and the second generation favored the overt form while generation 1.5 and third generation disfavored overt pronominal subjects. Given the low number of participants in each generation, the results for generation have to be taken with caution. The results for proficiency, on the other

hand, are more informative. The participants with lower proficiency produced significantly more overt pronominal subjects than those in the higher proficiency group.

Table 4: First person singular subject expression in Caribbean Spanish-English bilinguals in Florida: Proficiency

Multivariate regression analysis of the contribution of external factors to the probability of producing a first person singular overt pronominal subject vs. a null subject; factor groups selected as significant in gray background

	Factor weight	%	Ν	% of data
PROFICIENCY				
Lower	0.62	37	345	45.4
Higher	0.40	19.3	216	54.6
Range	55			
Total N			2,054	
Corrected Mean			.264	
Log likelihood			-1164.089	
Significance			.000	

For the lower proficiency level, in addition to producing an average rate of overt pronominal subjects higher than that of the higher proficiency group, there was more variation within the group. Within the higher proficiency group, the range of overt pronominal subjects is 6% to 30.3%, while it is 3% to 92.9% in the lower proficiency group.

 Table 5: Subject pronoun rates per participant

Percentage of use of overt pronominal subjects by participant in each proficiency group

	0/0	Ν	% of data
HIGHER PROFICIENCY			
Participant 2	16.9	10	2.9
Participant 3	18.8	15	3.9
Participant 4	30.3	143	23
Participant 5	25.3	23	4.4
Participant 10	6	25	20.4
LOWER PROFICIENCY			
Participant 1	3	4	6.5
Participant 6	7	14	9.7
Participant 7	61.7	137	10.8
Participant 8	92.9	105	5.5
Participant 9	25	26	5.1
Participant 11	36.6	59	7.8

When compared with previous studies, the number of overt pronominal subjects is lower than in other Caribbean varieties. Although we cannot explore this hypothesis further here it is possible that it is the result of convergence with other varieties of Spanish. It is also possible that differences do exist between different varieties within the Caribbean region. Participant 7, for instance, was the only participant from the Dominican Republic. Additionally, Participant 8 obtained the lowest proficiency score (14/50 in the DELE) and had difficulty maintaining a fluid conversation during the interview. Given the diversity attested in our participant pool and the low number of participants per cell, it is possible that the attested differences in proficiency are due to a variety of factors, including dialectal variation within the Caribbean region and differences in proficiency within each of the two proficiency groups, an issue that invites future research. In previous literature differences in rates across regions or groups of speakers did not necessarily imply differences in their grammars as instantiated in the variables regulating the distribution and the direction of effects (Carvalho et al. 2015). In order to examine possible differences across the two proficiency groups, we ran separate regressions for the higher and the lower proficiency groups.

The grammars of the higher proficiency group were regulated by two factors: switch reference and verb type. TAM continuity and verb form ambiguity were not selected as significant.

background				
	Factor weight	%	Ν	% of data
SWITCH REFERENCE				
Different referent	0.70	32.7	108	29.8
Same referent	0.41	13.7	108	70.2
Range	29			
VERB TYPE				
Stative	0.60	25.2	113	40.1
External	0.46	15.8	65	36.7
Mental	0.39	14.6	38	23.3
Range	21			
TAM CONTINUITY				
Different TAM	0.54	22	97	39.3
Same TAM	0.47	17.5	119	60.7
Range	7			
AMBIGUITY				
Ambiguous	0.55	22.4	34	13.6
Unambiguous	0.49	18.8	182	86.4
Range	6			
Total N			1121	
Corrected Mean			.175	
Log likelihood			-513.921	
Significance			.000	

Table 6. Subject pronoun expression in higher proficiency group

Multivariate regression analysis of the contribution of internal and external factors to the probability of producing a first person singular overt pronominal subject vs. a null subject in the higher proficiency group; factor groups selected as significant in gray background

Although previous research has claimed that bilingualism may affect the interface between syntax and pragmatics (Sorace 2011), for this group of bilinguals switch reference is a strong predictor of subject pronoun expression, with overt pronominal subjects being favored in contexts of different reference and nulls in contexts of same reference. With respect to verb type, recall that previous studies do not always return this variable as significant but when they do stative and mental verbs favor overt pronominal subjects while external action predicates favor the omission of the pronoun. In this group, stative and external action predicates exert the same effect as in previous studies. Mental predicates, on the other hand, disfavor the use of overt pronominal subjects. It is possible that this is due to the specific verbs included in the category (Orozco 2015). This variable remains highly ranked in the lower proficiency group while switch reference does not reach significance. Instead, TAM continuity is selected as significant. Verb form ambiguity is not significant in this group either.

T-1-1-	7.	C1-14			:	:	41	τ	····· · · · · · · · · · · · · · · · ·	
Table	11	Subject	pronoun	expres	ssion	1n	the.	Lower	proficiency	/ group
		·- ·· · J · · ·	F						r	0

Multivariate regression analysis of the contribution of internal and external factors to the probability of producing a first person singular overt pronominal subject vs. a null subject in the higher proficiency group; factor groups selected as significant in gray background

	Factor weight	%	Ν	% of data
VERB TYPE				
Mental	0.64	51	106	22.3
Stative	0.50	36.8	148	43.1
External	0.41	28.2	91	34.6
Range	23			
TAM CONTINUITY				
Same TAM	0.53	40.1	224	59.8
Different TAM	0.45	32.3	121	40.2
Range	8			
SWITCH REFERENCE				
Different referent	0.55	41.7	126	32.4
Same referent	0.48	34.7	219	67.6
Range	7			
AMBIGUITY				
Ambiguous	0.53	40	66	17.7
Unambiguous	0.49	36.3	279	82.3
Range	4			
Total N			933	
Corrected Mean			.366	
Log likelihood			-598.184	
Significance			.028	

For the lower proficiency group, verb type returns results consistent with previous research; mental and stative predicates favor overt pronominal subjects while predicates expressing external actions do not. TAM continuity exhibits the opposite trend attested in previous studies; same TAM favors overt pronominal subjects. It is important to point out, though, the small magnitude of effect of this variable (range: 8). Switch reference exhibits the same trend as in the higher proficiency group. It seems, however, that for this group of speakers the integration of pragmatic factors and language internal factors (syntax, in this case) is a site for language contact effects.

In general, in these two groups, the difference in proficiency manifests not only in overt pronominal rates but also in the strength of the variable switch reference, which tends to be the highest ranked variable in monolingual varieties of Spanish.

5. Discussion

The purpose of this paper is to examine language contact in a novel community to better understand the effect of language contact with English on Spanish subject pronoun expression. In previous studies, there was conflicting evidence on the effect of language contact on the rate and patterns of use of overt pronominal subjects vs. their omission in Spanish. In NYC an increase in the rate of overt pronominal subjects was attested while no increase was reported in New Mexico speakers. Regarding their distribution, in New Mexico, no differences or only minor differences were attested in the "bilingual" vs. the "monolingual" group. In NYC, in contrast, there were differences in the strength of the variables.

In the current study there is no monolingual control group so we examine language contact by comparison with results of previous studies and by comparing a higher and a lower proficiency group within the sample. Although comparing rates across studies is problematic (Silva-Corvalán 2001), it seems that our participants' rates are among the lowest reported for Caribbean speakers, including monolingual speakers. In previous research the lowest percentage reported is 28% in Castañer, Puerto Rico (Holmquist 2012) and the highest is 60% in San Juan, Puerto Rico (Cameron 1996). The higher proficiency group produces overt pronominal subjects 19.3% of the time, a percentage that is lower than that for Castañer, Puerto Rico. Our lower proficiency group produces 37% overt pronouns, an average percentage among monolingual varieties of Caribbean Spanish. This percentage is similar to that in Otheguy & Zentella's (2012) Caribbean Newcomers (36%) and lower than that of the NYR group (44%). The participants in this study, however, are closer in linguistic background to those in the NYR group as they are all dominant in English. Thus, it seems that even the lower proficiency group exhibits a low rate of overt pronominal subjects, a result consistent with Torres Cacoullos & Travis (2010) that language contact with English does not result in an increase in overt pronominal subject. In these data, however, it seems that there is a lower rate for bilinguals of high proficiency in Spanish than for monolingual speakers. This result may be due to differences in the data collection process; the interlocutor's variety of Spanish (Peninsular Spanish) and data were largely obtained from same referent contexts (70.2% of the higher proficiency data and 67.6% of the lower proficiency data), or to contact with non Caribbean varieties, an issue that we leave for future research.

Overall, comparing the rates in both groups of speakers in this study with rates in monolingual studies in the Caribbean region offers no evidence of an increase in overt pronominal expression in contact with English. Comparing both groups, though, reveals a higher rate of overt pronominal subjects at a lower proficiency in Spanish. It is possible, then, that more than contact with English, the increase in overt pronominal expression in bilinguals only occurs at lower levels of proficiency in Spanish.

In addition to rates, patterns were explored through the application of four variables commonly used in the field: switch reference, TAM continuity, verb from ambiguity, and semantic verb type. An initial analysis including only the variable proficiency revealed that these two groups were different. Thus, separate regressions for each group were performed to examine the role of these variables in these two groups of speakers. The higher proficiency group returned switch reference and verb type as significant, with patterns similar to those attested in other communities. Switch reference does not seem to have weakened (contrary to Otheguy & Zentella's 2012 results) with a range of 29. For comparison, Alfaraz (2015) reports a range of 27 for monolingual speakers in the Dominican Republican and Orozco (2015) a range of 32 for monolingual speakers in Colombian Costeño. In the lower proficiency group, on the other hand, switch reference is not significant, with a range of 7. Thus, in this group of bilinguals there is an effect of language contact in the strength (and significance) of the variable switch reference, as predicted by Otheguy & Zentella (2012) and in contrast with the results of Torres-Cacoullos & Travis (2010). It is possible, though, that all the participants in the latter study were of higher proficiency in Spanish.

Rather surprisingly, the variable verb type was significant for both proficiency groups. In previous studies there is variability in the relevance of this factor (Orozco 2015). In fact, Orozco (2015) reveals that frequency of the verb may interact with the semantic classification of verbs such that some of the higher frequency verbs (e.g. tener 'to have' among stative verbs) behave differently from other verbs in the same class. Future research, then, will examine this variable examining the behavior of these high frequency verbs. In the meantime, it seems that this lexical variable is rather stable in bilinguals in NYC, New Mexico, and Florida.

6. Conclusion

The aim of this study was to contribute to current debates in the language contact literature on the effect of language contact, in particular with English, in the expression of overt pronominal subjects in Spanish. Previous literature on Spanish language contact with English in the U.S. seems to return conflicting results. Bayley & Pease-Álvarez (1996, 1997), Flores-Ferrán (2004), Silva-Corvalán (1994), Torres-Cacoullos & Travis (2010) and Travis (2007) reveal no increase in subject expression in Spanish in the U.S. Lapidus & Otheguy (2005a, b), Lipski (1996), Montrul (2004), Otheguy & Zentella (2012), Otheguy et al. (2007), Toribio (2004), on the other hand, report an increase. Considering some differences across these studies, this paper aimed to test the possibility that the dialectal distribution in the community had an effect: NYC, for instance, has a slightly higher presence of Caribbean than non-Caribbean Spanish speakers while New Mexico has an almost exclusively non Caribbean Spanish-speaking population. Florida exhibits similar demographics to NYC. Thus, we anticipated finding a high rate of overt pronominal subjects in our data. Our results indicate that this is not the case. In fact, both higher and lower proficiency groups report rates within (and even lower) than some of the rates reported in the literature on monolingual Caribbean varieties of Spanish. Thus, our results are more consistent with the lack of a contact effect in rates of overt pronominal subjects. Our examination does reveal, though, an effect of proficiency in Spanish, where lower proficiency speakers use overt pronominal subjects significantly more than higher proficiency speakers. Proficiency in Spanish can be difficult to measure in these participants. Since we used instructed HSs we were able to use a section of the DELE as an independent measure of proficiency. While none of the speakers rated within the advanced range (over 40/50), using a median split we obtained two differentiated proficiency groups. Thus, even though this task is problematic for several reasons (written task focused on prescriptive Peninsular Spanish), using a median split allowed us to use an independent measure of proficiency normalized for our group of speakers.

Other than rates we focused our analysis on patterns of use of overt and null pronominal subjects as instantiated in the variables returned as significant in a multivariate regression for each of the proficiency groups. The results of the higher proficiency group were also consistent with the results in Torres-Cacoullos & Travis (2010) in the variables that were found as significant (switch reference and semantic verb type) and the strength of the variables. The lower proficiency group, however, exhibited a strong weakening of variable switch reference, which was not found to be significant.

In conclusion, a contact effect manifested in a higher rate in overt pronominal subjects in bilinguals than in monolinguals and in a weakening of the effect of variables is not evident in our data, especially if we consider the higher proficiency speakers. A proficiency effect, on the other

hand, is attested. The lower proficiency speakers do produce significantly more overt pronominal subjects and display a weakening of the variable switch reference. These results indicate that the differences between studies in the U.S. Southwest and NYC cannot be attributed to the dialectal distribution of the communities.

References

- Alfaraz, G. 2015. Variation of Overt and Null Subject Pronouns in the Spanish of Santo Domingo. In Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective, edited by Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin, 3—17. Washington D.C.: Georgetown University Press
- Bayley, R., Pease-Álvarez, L., 1996. Null and expressed pronoun variation in Mexican-descent children's Spanish. In: Arnold, J., Blake, R., Davidson, B. (Eds.), Sociolinguistic Variation: Data, theory, and analysis, CSLI Publications, Stanford, pp. 85–99.
- Bayley, R., Pease-Álvarez, L., 1997. Null pronoun variation in Mexican-descent children's narrative discourse. Language Variation and Change 9, 349–371.
- Bentivoglio, P., 1987. Los sujetos pronominales de primera persona en el habla de Caracas. Central University of Venezuela, Caracas.
- Cameron, R., 1995. The scope and limits of switch reference as a constraint on pronominal subject expression. Hispanic Linguistics 6–7, 1–27.
- Cameron, R., Flores-Ferrán, N., 2004. Perseverance of subject expression across regional dialects. Spanish in Context 1 (1), 41–65.
- Carvalho, A.M., Orozco R, & Shin N.L. 2015. Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective. Washington D.C.: Georgetown University Press
- Casanova Seuma, L.,1999. El sujeto en catalán coloquial. Revista española de lingüística 29 (1), 105–131.
- Cuza, A. 2013. Cross-linguistic influence at the syntax proper: interrogative subject-verb inversion in heritage Spanish. *International Journal of Bilingualism*, 17(1), 71-96.
- Enríquez, E., 1984. El pronombre personal sujeto en la lengua española hablada en Madrid. Consejo Superior de Investigaciones Científicas, Madrid.
- Erker, D., Guy, G. R., 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. Language, 88 (3), 526–557.
- Erker, Daniel, & Ricardo Otheguy. 2016. Contact and coherence: Dialectal leveling and structural convergence in NYC Spanish. *Lingua* 172–173: 131--146
- Flores, Nydia, and Jeannette Toro. 2000. The persistence of dialect features under conditions of contact and leveling. *Southwest Journal of Linguistics* 19.31–42.

- Flores-Ferrán, N., 2002. Subject Personal Pronouns in Spanish Narratives of Puerto Rican in New York City: A Sociolinguistic Perspective. Lincom Europa, Munich.
- Flores-Ferrán, N., 2004. Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? Language Variation and Change 16, 49–73.
- Hochberg, J. 1986. Function compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62. 609-21.
- Holmquist, J. 2012. Frequency rates and constraints on subject personal pronoun expression: Findings from the Puerto Rican highlands." Language Variation and Change 24: 203–220.
- Lapidus, N., Otheguy, R., 2005a. Contact induced change? Overt nonspecific ellos in Spanish in New York. In: Sayahi, L., Westmoreland, M. (Eds.), Selected Proceedings of the 2nd Workshop on Spanish Sociolinguistics, Cascadilla Press, Somerville, pp. 67–75.
- Lapidus, N., Otheguy, R., 2005b. Overt nonspecific ellos in Spanish in New York. Spanish in Context 2, 157–174.
- Lipski, John M. 1994. Latin American Spanish. London: Longman.
- Lipski, John M. 1996. Patterns of pronominal evolution in Cuban-American bilinguals. *Spanish in contact: Issues in bilingualism*, ed. by Ana Roca and John B. Jensen, 159–86. Somerville, MA: Cascadilla.
- Martínez Sanz, C. 2011. Null and overt subjects ina variable system: The case of Domincian Spanish. Ph.D. diss., University of Ottawa.
- Montrul, S. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition* 7.125–42.
- Montrul, S., and R. Slabakova. 2003. Competence similarities between native and near-native speakers: An investigation of the preterite/imperfect contrast in Spanish. *Studies in Second Language Acquisition* 25(3), 351–98.
- Morales, A., 1997. La hipótesis funcional y la aparición del sujeto no nominal: El español de Puerto Rico. Hispania 80, 153–165.
- Orozco, Rafael. 2015. Pronominal Variation in Colombian Costeño Spanish. In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*, edited by Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin, 17–37. Washington D.C.: Georgetown University Press.
- Orozco, R. & G. Guy. 2008. El uso variable de los pronombres sujetos. ¿Qué pasa en la costa Caribe colombiana?. In *Selected Proceedings of the Fourth Workshop on Spanish Sociolinguistics*, edited by Maurice Westmoreland and Juan Antonio Thomas 70–80. Sommerville, MA: Cascadilla Proceedings Project.
- Ortiz López, l. 2009. Pronombres de sujeto en el español (L2 vs L1) del Caribe. In El español en Estados Unidos y otros contextos de contactos: sociolinguistica, ideology y pedagogía, edited by Manel Lacorte and Jennifer Leeman, 85–110. Frankfurt Main/Madrid: Vervuert/Iberoamericana.
- Otheguy, R., Zentella, A. C., 2012. Spanish in New York. Language Contact, Dialectal Leveling, and Structural Continuity. Oxford University Press, Oxford, New York.

- Otheguy, R., Zentella, A. C., Livert, D., 2007. Language and dialect contact in Spanish in New York: Towards the formation of a speech community. Language 83, 1–33.
- Ranson, D. L., 1991. Person marking in the wake of /s/ deletion in Andalusian Spanish. Language Variation and Change 3 (2),133–152.
- Shin, N. L., 2012. Variable Use of Spanish Subject Pronouns by Monolingual Children in Mexico. In: Geeslin, K., Díaz-Campos, M. (Eds.), Selected Proceedings of the 2010 Hispanic Linguistics Symposium, Cascadilla Proceedings Projects, Somerville, pp. 130–141.
- Shin, N. L., and Otheguy, R., 2013. Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. Language in Society 42:429–452.
- Silva-Corvalán, C., 1982. Subject expression and placement in Mexican-American Spanish. In: Amastae, J., Elías-Olivares, L. (Eds.), Spanish in the United States: Sociolinguistic aspects, Cambridge University Press, Cambridge, pp. 93–120.
- Silva-Corvalán, C., 1994. Language Contact and Change: Spanish in Los Angeles. Clarendon, Oxford.
- Sorace, A. 2011. Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism* 1, 1-33
- Toribio, A. J. 2004. Convergence as an optimization strategy in bilingual speech: Evidence from code-switching. *Bilingualism: Language and Cognition*, 7. 165–73.
- Torres Cacoullos, R., Travis, C. E., 2010. Variable Yo Expression in New Mexico: English influence? In: Rivera-Mills, S., Villa, D. (Eds.), Spanish of the U.S. Southwest: A Language in Transition, Iberoamericana/Vervuert, Madrid, pp. 185–206.
- Travis, C. E., 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. Language Variation and Change 19, 101–135.