B O R E A L I S
AN INTERNATIONAL JOURNAL OF HISPANIC LINGUISTICS

SUBJECT EXPRESSION IN A SOUTHEASTERN U.S. MEXICAN COMMUNITY*

Philip P. Limerick
Eastern Kentucky University/New Mexico State University

ABSTRACT. Through an analysis of immigrant Spanish in Georgia, potential contact-induced language change is investigated through the lens of subject pronoun expression. Pronoun variation among Mexican speakers is examined using sociolinguistic interview data. Tokens of overt and null subject pronouns (N = 4,649) were coded for linguistic variables previously shown to constrain subject expression (e.g. person/number, tense-mood-aspect [TMA], polarity) as well as social variables (e.g. English proficiency, age), and then analysed using multivariate analyses in *Rbrul*. Results indicate an overall pronoun rate of 27%, which is slightly higher than what has been reported for monolingual Mexican Spanish. Several linguistic variables (e.g. person/number, switch-reference, morphological ambiguity, polarity) and one social variable (age) played a significant role in pronoun variation. Moreover, differential effects were revealed when compared to monolingual Mexican Spanish for variables such as TMA. These findings point in the direction of dialect contact influences and the presence of a unique variety of Mexican Spanish in the U.S.

Keywords. subject expression; pronouns; Mexican Spanish; U.S. Spanish; morphosyntactic variation

RESUMEN. Por medio de un análisis del español inmigrante en Georgia, se investiga el posible cambio lingüístico inducido por contacto desde la perspectiva de la expresión de los sujetos pronominales. Se examina la variación de pronombres entre los hablantes mexicanos utilizando datos de entrevistas sociolingüísticas. Los ejemplos de sujetos pronominales explícitos e implícitos (N = 4,649) se codificaron según las variables lingüísticas que se han demostrado restricciones en la expresión del sujeto (ej. persona/número, tiempo-modo-aspecto [TMA], polaridad) así como las variables sociales (ej. dominio del inglés, edad), y después se analizaron utilizando el análisis multivariable en *Rbrul*. Los resultados indican una tasa pronominal general de 27%, la cual es marginalmente más alta que la que se ha reportado para el español mexicano monolingüe. Varios factores lingüísticos (ej. persona/número, cambio de referencia, ambigüedad morfológica, polaridad) y un factor social (edad) jugaron un papel significativo para la variación de pronombres. Además, se revelaron efectos diferenciales al ser comparados con el español mexicano monolingüe para las variables tales como TMA. Estos hallazgos apuntan a influencias de contacto dialectal así como la presencia de una variedad única de español mexicano en los EE.UU.

Palabras clave. expresión del sujeto; pronombres; español mexicano; español estadounidense; variación morfosintáctica

* This article is a derivation of my doctoral dissertation research carried out at the University of Georgia, which has received funding from the University of Georgia Graduate School and the Willson Center for Humanities & Arts. I would like to thank Chad Howe, Sarah Blackwell, and Margaret Quesada for their invaluable feedback and support as well as Lilian Zhu and Trevor Talmadge for their assistance with interview transcription. I am also grateful to an anonymous reviewer for their helpful suggestions and insights.

1. Introduction

The present study investigates the linguistic behaviour of newly-forming bilingual communities in the U.S., particularly at the initial stages of language contact between Spanish and English. Looking to the Southeastern U.S. as a relatively new contact situation, this analysis aims to understand the processes of language shift in situations of language contact and bilingualism. Latin American immigration to the Southeast has led to recent and significant demographic shift in this region, and thus substantial Spanish-speaking populations have emerged that historically were not part of the Southeast in the way that they were in the Southwest or Northeast. Due to the lack of an established presence, Southeast Spanish is understudied in sociolinguistic research compared to varieties in areas such as California, New Mexico, and New York (but see, e.g., Smith 2006; Carter 2014; Lamanna 2012; Wilson 2013; Montes-Alcalá & Sweetnich 2014; Limerick 2017). This study will focus on Spanish spoken in Georgia, observing the Latino community in Roswell, which is part of the Atlanta metropolitan area. Roswell represents the larger demographic shift in the Southeast, with recent emergence of Spanish-speaking communities, making it an excellent test site for studying emerging speech communities (Limerick 2017).

The current investigation focuses on subject pronoun expression (SPE), which has been studied extensively among monolinguals (e.g. Cameron 1994; Travis 2005; Orozco & Guy 2008; Orozco 2015; Alfaraz 2015; Lastra & Martín Butragueño 2015), bilinguals (e.g. Silva-Corvalán 1994a,b; Flores-Ferrán 2004; Travis 2007; Shin & Otheguy 2009; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Torres Cacoullos & Travis 2015), and L2 learners (e.g. Lozano 2009; Blackwell & Quesada 2012; Geeslin, Linford, & Fafulas 2015; Geeslin & Gudmestad 2016). Whereas SPs are nearly obligatory in English, Spanish allows the SP to be omitted, as in (1), or expressed, as in (2).

(1) *Sí, ∅* ***estudié*** *para abogada y… en la ciudad de México…* [F39Mex][1]
'Yes, (I) studied to be a lawyer and…in Mexico City…'

(2) *Ah, no,* ***yo*** *lo* ***estudié*** *cuando recién llegué aquí* [M34Mex]
'Uh, no, I studied it when I first arrived here'

Preliminary research has explored SPE among ten Mexican speakers in Roswell (Limerick 2017), suggesting that bilingualism impacted their SP usage in relation to processes of grammatical simplification when compared to monolingual Spanish. Building on this research, the current study seeks to further explore potential early language contact effects in Roswell. I examine variables previously unexplored in this community and rarely studied in previous SP research (English proficiency, preferred media language, polarity, mood). Through a study of 20 first-generation Mexican immigrants, I investigate the frequency of and constraints on SPE to investigate contact-induced change.

The influence of linguistic and social factors on SP variation is analysed and compared with non-contact varieties of Spanish as well as other contact varieties by addressing the following questions: (a) What is the overall distribution of overt and null subject pronouns (SPs) in the Roswell data? How does the overt SP rate compare to that of non-contact varieties of Mexican Spanish? (b) How do language-internal factors

[1] Speaker codes indicate the gender, age, and national origin of the participants in the current study.

condition SP variation and how do such patterns compare to previous research? (c) Do language-external factors (e.g. age, English proficiency) play a role in how speakers use SPs? (d) Does bilingualism play an influencing role in SPE? (e) What do the rates/constraints of SPE tell us about the emergent nature of Roswell Spanish?

The data employed for this analysis come from sociolinguistic interviews conducted in Roswell by the author. To answer the above questions, a quantitative analysis is carried out using variationist sociolinguistic methods (Labov 1972; Tagliamonte 2012). The article is structured as follows: Section 2 provides a background of Spanish in the Southeastern U.S. and Section 3 reviews previous studies of Spanish SPE. Section 4 outlines the current study's methodology and Section 5 presents the results. A discussion of the results is presented in Section 6, and Section 7 concludes the article.

2. Spanish in the Southeastern U.S.

Within the last 30 years, the Southeastern U.S. has seen immense immigration from Latin America, with many Southern cities experiencing growth rates in their Latino populations above 500% since 1990 (Winders 2005) and several states above 100%. For instance, between 2000 and 2011, Georgia, North Carolina, and South Carolina had growth rates of 102%, 119.6%, and 154.5%, respectively (Motel & Patten 2013). Such demographic shifts are not only seen in a few cities or states, but have been evidenced in the Southeast as a whole. Overall, the Southern U.S. had the highest growth rate (57.3%) of all regions of the country in the 2010 U.S. Census.

This influx of Latin American immigrants led to a widespread use of Spanish in Southeastern states which, during most of the twentieth century had been almost entirely English-speaking, such as in Georgia and North Carolina (Klee & Lynch 2009).

Some notable studies on Southeast Spanish include those on the imperfect/imperfect progressive distinction in North Carolina (Lamanna 2012) and the analysis of Carter (2014), which involved examining discourse concerning Spanish and U.S. Latinos in North Carolina. In Northeast Georgia, Smith (2006) studied codeswitching among Latin American immigrants. Moreover, Montes-Alcalá and Sweetnich (2014) studied language shift among Latinos in Atlanta. A few studies have examined Spanish in Roswell, Georgia, the city of focus for the present investigation (Wilson 2013; Limerick 2017). Wilson (2013), for example, examined narrative structure among Roswell speakers.

Regarding SPE in particular, Limerick (2017) examined pronoun use among ten Mexican immigrants utilizing Wilson's (2013) interview data. Limerick analysed SP patterns and considered possible English contact effects. He found that SPE was constrained by person/number, switch-reference, tense-mood-aspect (TMA),[2] lexical content, and clause type, the same factors (among many others) reported as significant in numerous other studies. Regarding potential English contact effects, little to no evidence was found as the overt pronoun rate was similar to that of monolingual Mexican Spanish reported in other studies, and, crucially, the social factors analysed (length of residency [LOR]; age of arrival [AOA]), were not significant. Nevertheless, a more indirect influence of bilingualism was suggested due to a reduced effect for the switch-reference constraint. This reduced effect was interpreted as a process of simplification on part of bilinguals whereby they were potentially losing sensitivity to switch-reference as evidenced by increased overt SP usage in same-reference contexts, a context that more strongly favours nulls in monolingual varieties. (See Section 3.2 for a discussion of investigating English contact through subject expression). To expand

[2] This is a common acronym used in the literature and will be used for the remainder of the paper.

on this research, the present study draws from a larger corpus of interviews and more closely examines language contact effects by analysing additional factors (e.g. English proficiency; preferred media language) and comparing SPE with non-contact varieties of Spanish.

3. SPE in Spanish

3.1 Pronoun Rates and Factors Conditioning SPE

It has been demonstrated that, in general, varieties from Mexico and Spain tend to have the lowest overt pronoun rates overall (16-25% overt), South American varieties such as Argentina, Colombia, and Chile exhibit mid-range rates (30-40% overt), and Caribbean speakers (e.g. Puerto Rico, Dominican Republic) have the highest pronoun rates out of all Spanish dialects (45-55%).[3] Regarding the factors that systematically guide SP variation, those shown to play a role across most dialects are person/number (Silva-Corvalán 1994a), same vs. switch-reference (Bayley & Pease-Alvarez 1996), clause type (Morales 1997), TMA (Cameron 1994), lexical content of the verb (Travis 2007), specificity of referent (Cameron 1992), reflexivity of the verb (Otheguy, Zentella, & Livert 2007), and priming (Travis 2007), as well as social factors including age and gender (Carvalho & Child 2011) (See Sections 3.3 and 3.4 for an in-depth discussion of these factors).

3.2. Pronoun Use and English contact

Contact varieties sometimes show higher overt pronoun rates than monolingual varieties, which is commonly attributed to English contact, as overt SPs are nearly obligatory in English but not in Spanish. Therefore, increased exposure to English has been argued to engender higher usage of overt pronouns in Spanish, with many studies finding higher pronoun rates among more established immigrants, the U.S. born, younger arrivals to the U.S., and those with higher English proficiency (e.g. Lapidus & Otheguy 2005; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012; Abreu 2012; Shin & Montes-Alcalá 2014). However, other researchers have not observed increased overt pronouns in U.S. contact varieties (e.g. Silva-Corvalán 1994a; Flores-Ferrán 2004; Torres Cacoullos & Travis 2010; Limerick 2017). The reason for this discrepancy has been explained in terms of differential social classes and social networks (Shin & Otheguy 2013; Shin & Van Buren 2016). Specifically, Shin and Otheguy's (2013) analysis of different nationality groups in NYC Spanish revealed that more affluent speakers, who tend to have looser social networks (Cubans and Colombians), had higher SP rates than less affluent speakers with more dense social networks (Mexicans, Dominicans, and Puerto Ricans). The latter group tends to have more tight-knit social networks and consequently maintain their linguistic patterns. Shin and Van Buren's (2016) study of US-born bilingual Mexican children in Washington/Montana corroborates this argument, finding little to no change in SPE when compared to monolingual adults. These authors posit that since the children were part of a more tight-knit community and were mostly in contact with monolingual Spanish-speakers, they retained their rates and usage patterns of SPs.

Furthermore, what constitutes evidence for language change (contact-induced or not) is debated among scholars. Poplack and Levey (2010) posit that changes in SP frequencies do not constitute evidence for change. They affirm that in order to consider

[3] See, for example, Solomon (1999) (Mexico), Ranson (1991) (Spain), Barrenechea & Alonso (1977) (Argentina), Orozco (2008) (Colombia), Cameron (1993) (Puerto Rico).

contact-induced change, analysts must consider the ranking of factors that are probabilistically shown to favour or disfavour SP occurrence, that is, *constraint hierarchies*. Torres Cacoullos and Travis (2010:189) agree by stating that 'reliance on overall rates of use to determine contact-induced change is problematic because it is well known that regional dialects vary enormously' and suggest that divergences in grammatical patterning across dialects are necessary to substantiate contact-induced change.

Other scholars attribute changes in SP usage to English more indirectly, through a process of simplification on part of the bilingual, whereby they lighten the cognitive load of managing two linguistic systems by simplifying certain morphosyntactic or discourse constraints (e.g. Silva-Corvalán 1994b; Sorace 2004). That is, speakers may exhibit a weakening of certain constraints due to bilingualism in general, which indirectly is an influence from English in the U.S. context. Specifically, speakers have been shown to simplify pragmatic constraints on SP usage. For instance, various scholars have argued that Spanish-English bilinguals show a weakened sensitivity to switch-reference whereby they exhibit an increased use of overt SPs in same-reference contexts, an environment where nulls are more strongly favoured in monolingual varieties (e.g. Silva-Corvalán 1994a; Shin & Otheguy 2009).

I present below the reported overall overt pronoun rates for monolingual varieties (Table 1) as well as Spanish-English contact varieties (Table 2). Notice the cross-dialectal variability in rates among monolingual Spanish as well as between monolingual and contact varieties.

Table 1: Pronoun rates of monolingual varieties

Variety	**PN Rate**	**Source**
Yucatan, Mexico	16%	Michnowicz 2015
Valladolid, Mexico	19%	Solomon 1999
Mexico City	22%	Lastra & Martín Butragueño 2015
Puente Genil, Spain	24%	Ranson 1991
New York Newcomers (Mexican, S. American, Caribbean)	30%	Otheguy, Zentella, & Livert 2007
Barranquilla, Colombia	36%	Orozco & Guy 2008
Buenos Aires, Argentina	36%	Barrenechea & Alonso 1977
Santiago, Dominican Republic	39%	Olloqui de Montenegro 1987
San Juan, Puerto Rico	45%	Cameron 1993

Table 2: Pronoun rates of contact varieties

Variety	PN Rate	Source
North Carolina Mexicans	17%	McKnight 2013
Georgia Mexicans	21%	Limerick 2017
New Jersey Mexicans	24%	Flores Ferrán 2007
Los Angeles Mexicans	29%[4]	Silva-Corvalán 1994a
New York born/raised (Mexican, S. American, Caribbean)	38%	Otheguy, Zentella, & Livert 2007
New York Dominicans	48%	Flores & Toro 2000

Regarding contact Spanish, one can see from a comparison of Tables 1 and 2 that SP rates are sometimes higher, but sometimes equal or lower than rates for monolingual varieties, as highlighted above. For instance, higher rates were observed for New York Dominicans (48%) than were for Dominicans in Santiago (39%). Likewise, New York born/raised Spanish-speakers showed an increase (38%) when compared to New York Newcomers (30%). In contrast, we see in other cases that pronoun rates are virtually the same in contact varieties when compared to monolingual varieties, for instance in the case of North Carolina (17%) and Georgia Mexicans (21%) when compared to Yucatan (16%) and Mexico City (22%). Next, I discuss in detail the linguistic and social factors that are relevant to the present study.

3.3 Linguistic Variables

This section will outline the main findings according to previous SPE research, particularly for the variables employed in the current study. It is important to note that several additional factors not discussed here such as reflexivity, priming, and lexical frequency also greatly influence SPE (see, e.g., Travis 2005; Carvalho & Child 2011; Erker & Guy 2012 for further discussion).

3.3.1 Person/number

Person/number of the verb highly influences SP variation and has been shown to be the strongest predictor of variable SPE cross-dialectally (Orozco 2015). Specifically, first-person singular and third-person singular verbs tend to favour overt pronouns (e.g. Silva-Corvalán 1994a; Flores-Ferrán 2002; Shin 2012; Lastra & Martín Butragueño 2015). In fact, most studies have found that all singular forms in general are more likely to appear with overt SPs compared to plural forms (Orozco 2015).

3.3.2 Switch-reference

Switch-reference has also shown a strong influence on SP variation (e.g. Bentivoglio 1987; Cameron 1994; Silva-Corvalán 1994a; Bayley & Pease-Alvarez 1997; Travis 2005; Prada Pérez 2009; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015). Specifically, where there is a switch in subject referent, the SP is often overt; when there is no switch, null SPs are preferred. This pattern is generally thought to have a functional influence that has to do with referential tracking (Shin & Otheguy 2009) and facilitating interpretation of the antecedent for the listener (Cameron 1994). Certain contact varieties have shown a weakening of this constraint (Flores-Ferrán 2002; Shin & Otheguy 2009; Michnowicz

[4] Average rate among all three generational groups from Silva-Corvalán (1994a:153).

2015). For example, Flores-Ferrán (2002) found that NYC-born Spanish-speakers exhibit an increased use of overt SPs in same-reference contexts, demonstrating their loss of sensitivity to null pronoun usage in such contexts (e.g *yo vivía aquí, **yo me iba** a Alpharetta, caminando, **yo me iba** a Sandy Springs* 'I lived here, I would go to Alpharetta, walking, I would go to Sandy Springs' [M32Mex]).

Furthermore, some researchers include an additional category, *partial-switch*, which analyses cases of switch in subject where the subject is coreferential with the immediately preceding object, as in (3).

(3) *…en México me inculcó mi mamá que **tenía** que ir a la iglesia…* [M52Mex]
'…in Mexico my mom instilled in me that I had to go to church…'

Generally speaking, previous studies have reported either a slight favouring of overt SPs, or a neutral effect for such contexts, that is, neither a favouring nor a disfavouring effect (e.g. Orozco & Guy 2008; Otheguy & Zentella 2012; Lastra & Martín Butragueño 2015; Orozco 2015).

3.3.3 Tense-mood-aspect (TMA)[5]

Certain TMAs favour overt SPs while others favour nulls. Several studies have found that imperfects and conditionals favour overt SPs while presents and preterits are more likely to appear with nulls (Silva Corvalán 1982; Cameron 1994; Travis 2007; Carvalho & Bessett 2015). It has been proposed that imperfects and conditionals favour overt SPs due to their potential ambiguity. Since their first- and third-person singular verb forms are morphologically indistinct, the use of overt SPs would serve to disambiguate the referents of such forms (see Hochberg 1986). However, other studies have found no such correlation (e.g. Enríquez 1984; Bentivoglio 1987; Ranson 1991).

An alternative explanation has been proposed by Silva-Corvalán (2001), who discusses imperfects and preterits in relation to discourse functions and suggests that more overt SPs are used with imperfects due to the backgrounded nature of imperfect aspect, and that fewer overt SPs are used with preterits since they tend to foreground events. That is, for events that are more backgrounded the focus is on the subject (hence the use of an overt SP) whereas the focus is more on the action with more foregrounded events (hence the lack of an overt SP).

3.3.4 Morphological ambiguity

Similar to TMA, many researchers have shown that, when examining verb forms that are morphologically ambiguous, these forms favour overt SPs while unambiguous forms prefer nulls (e.g. Hochberg 1986; Travis 2007; Prada Pérez 2009; Erker & Guy 2012; Lastra & Martín Butragueño 2015; Michnowicz 2015). Ambiguous verb forms are those that have indistinct person inflections in the first and third-person singular, specifically in the imperfect, subjunctive (both present and past), conditional, and pluperfect. The idea is that speakers use overt SPs to clarify the referent for the listener. However, some scholars have found the opposite effect and argue that true ambiguity is rare and that it is the context that clarifies the antecedent. For instance, Ranson (1991) found fewer overt SPs with ambiguous verb forms relative to unambiguous forms and posits that contextual markers such as previous mention of the referent or background knowledge help to clarify the speaker's intended referent when the SP is null.

[5] Although mood in particular is not discussed here (merely tense and aspect), it is addressed as a separate category (Section 3.3.6 below), following Lastra and Martín Butragueño (2015). I retain the category name of *TMA* here, however, to maintain consistency with the SPE literature.

3.3.5 Verb class

Several researchers have found that verb class can also determine how an SP is manifested (Bentivoglio 1987; Silva-Corvalán 1994a; Travis 2007; Otheguy & Zentella 2012; Orozco 2015). It has been noted that verbs of psychological activity (e.g. *creer* 'believe'), verbs of communication, and copulas tend to be expressed with overt pronouns, with psychological verbs showing the highest probability. On the contrary, motion verbs tend to disfavour overts (Bentivoglio 1987; Silva-Corvalán 1994a; Travis 2007). Regarding psychological verbs, it has been hypothesized that overts are frequently used because these verbs tend to express the point of view of the speaker and because of the implied contrastive function that is often carried out in such contexts (Silva-Corvalán 1994a). Additional categories have also been employed for verb class, such as stative and activity verbs (e.g. Orozco & Guy 2008; Erker & Guy 2012; Otheguy & Zentella 2012; Orozco 2015). These studies have generally found that stative verbs favoured overt SPs while activity verbs favoured nulls.

3.3.6 Verbal mood

The influence of mood on SPE has not been extensively studied in previous research. The subjunctive mood has been examined in relation to specific tenses and aspects, but not so much in direct relation to the indicative and imperative moods in general. Lastra and Martín Butragueño (2015) have begun studying this variable in Mexico City Spanish, finding that non-indicative moods (subjunctive and imperative), disfavoured overt SPs. Regarding the indicative mood, while showing a higher frequency of overt SPs compared to imperatives, it showed neither a statistically significant favouring nor disfavouring effect on SPE.

3.3.7 Specificity

The use of a specific referent as opposed to a more general referent (e.g. impersonal *tú*), also tends to condition whether a pronoun is overt or null. Previous research has shown that the effect of specificity on SPE can depend on the dialect. Cameron's (1992, 1993) work on San Juan and Madrid Spanish demonstrates that non-specific subject referents favoured overt SPs in San Juan while specific referents favoured overt SPs in Madrid. For Mexican Spanish, both Solomon (1999) and Michnowicz (2015) found that, like Madrid Spanish, overt SPs were favoured by specific referents.

3.3.8 Polarity

Polarity considers affirmative vs. non-affirmative clauses and has occasionally been analysed in previous research, with some researchers finding a significant effect (Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016) and others reporting a lack of effect (Travis & Torres Cacoullos 2012; Torres Cacoullos & Travis 2015).[6] For example, Lastra and Martín Butragueño (2015) found that non-affirmatives (negatives and interrogatives) disfavoured overt SPs and that affirmatives showed a neutral effect. These authors give a possible explanation for this pattern in terms of the frequent clustering together of negative clauses in their data, which tend to be coreferential. Likewise, Geeslin and Gudmestad (2016) found that negation disfavoured overt SPs while affirmative clauses favoured them.

[6] Some differences in terminology among these studies should be noted: Lastra and Martín Butragueño (2015:46) use the term *enunciative type* to refer to polarity. Other researchers use the terms *polarity* (Travis & Torres Cacoullos 2012) and *verbal negation* (Geeslin & Gudmestad 2016).

3.3.9 Style

Speech style has also been found to constrain SPE (e.g. Ávila-Jiménez 1996; Lastra and Martín Butragueño 2015). Ávila-Jiménez (1996) reported that more casual speaking favoured overt SPs in Puerto Rican Spanish. Similarly, Lastra and Martín Butragueño (2015) found that Mexico City speakers tended to use more overt SPs for more conversational styles, particularly when they had longer speech turns in the latter part of their interviews, as opposed to shorter speech turns at the beginning of the interview (less conversational), for which they used fewer overt SPs.

3.4 Social Variables

3.4.1 Age

Age is one of the main social variables that has been shown to influence SP usage. Carvalho and Child (2011) found a favouring of null pronouns among younger speakers in Uruguay (ages 16-29) while older speakers favoured overt SPs, a trend they hypothesise is influenced by the Portuguese substratum (p. 23). Similarly, Orozco and Guy (2008) observed that older Colombians favoured overt SPs while adolescents favoured nulls, which they explain as a possible change in progress attributed to greater access to education by young people compared to prior generations. They also hypothesise that these results are due to influence of the Colombian highlands dialect. Lastra and Martín Butragueño (2015) found the same effect whereby older speakers favoured overt SPs and younger speakers were more likely to produce nulls. These researchers raise the issue of a possible change in progress toward less overt pronoun use in Mexico City. Contrastively, Flores-Ferrán (2002) found that older speakers (50+) of Puerto Rican Spanish in NYC exhibited lower rates of overt SPs than younger speakers and attributes this to the tendency of older speakers to prefer a more conservative use of language following prescriptive grammar norms.

3.4.2 Gender

Numerous authors have found that females and males differ in their SP frequencies such that women use more overt SPs than men (e.g. Bayley & Pease-Alvarez 1996; Solomon 1999; Carvalho & Child 2011; Otheguy & Zentella 2012; Shin & Otheguy 2013; Alfaraz 2015), while others found gender to be nonsignificant (e.g. Orozco & Guy 2008; Lastra & Martín Butragueño 2015; Michnowicz 2015). For instance, in their study of Uruguayan Spanish, Carvalho and Child (2011) found that females favoured overt SPs whereas males favoured nulls. Similarly, Shin and Otheguy (2013) observed a higher rate of overt pronouns among women in NYC, which they attribute to women's extensive contact with U.S. born bilinguals. Bayley and Pease-Alvarez (1996) hypothesized that women use narrative discourse more often than men and thus more SPs. Moreover, Alfaraz (2015) suggests that this gender pattern, particularly in the Dominican Republic, is linked to prestige.

3.4.3 Educational level

A relatively understudied variable, educational level has occasionally been found to condition SPE. For example, Ávila-Jiménez's (1996) study of Puerto Rican Spanish found that overt SPs were disfavoured among the college educated and favoured among those with lower education levels. The few additional studies that have examined this variable (Otheguy & Zentella 2012; Lastra & Martín Butragueño 2015), while not finding a statistically significant effect, have reported trends consistent with Ávila-

Jiménez's (1996) analysis. Lastra and Martín Butragueño (2015) suggest that this pattern could indicate that null SP use has more prestige.

3.4.4 Socioeconomic status (SES)

Socioeconomic status (SES) is another understudied variable in relation to pronoun use (exceptions include Otheguy and Zentella 2012, Alfaraz 2015, and Orozco 2015). While neither Orozco (2015) nor Alfaraz (2015) have found an effect for SES, Otheguy and Zentella (2012) found that speakers with lower SESs favoured overt SPs while higher SES speakers favoured null SPs among the Mainland population in NYC.

3.4.5 Length of residency (LOR), age of arrival (AOA), and English proficiency

Important factors to consider regarding potential influence of English on Spanish are LOR in the U.S., suggesting greater exposure to English, and AOA to the U.S., indicating speakers' susceptibility to change. Drawing from Veltman (2000), Otheguy, Zentella, and Livert (2007:779) state that older arrivals who speak Spanish more frequently and fluently should show greater resistance to English influence. In addition to finding that longer LORs in NYC correlated with higher frequencies of overt SPs, the authors found that younger arrivals, being more susceptible to English influence, exhibited higher pronoun rates than older arrivals, who tend to use less English and whose Spanish is more proficient, thereby making them less susceptible to the influences of language contact.

Additionally, English proficiency has been examined in studies of U.S. Spanish, with some researchers finding a positive correlation between English proficiency and overt pronoun use (e.g. Otheguy & Zentella 2012) and others finding no such correlation (e.g. Silva-Corvalán 1994a). Another factor to explore for examining contact effects, which has not been addressed in previous studies to my knowledge, is media language. It could be the case that those who prefer English for media exhibit a higher use of overt SPs than those who prefer Spanish. This factor is explored in the current study.

4. Methodology

4.1 Data Collection

In 2015 I conducted sociolinguistic interviews with Spanish-speakers who were either living or working in Roswell, GA. The sample for the present investigation consists of 20 first-generation Mexican immigrants. Regarding the recruitment of participants, the majority took place at a local plaza containing several Latino businesses, including a grocery store, clothing stores, barber shops, restaurants, and many others. I also recruited participants at a local church and at an ESL school. The interviews lasted between 30 minutes and one hour. They were informal, conversational, and addressed topics of personal history, local community life, differences between Mexico and the U.S., and experiences adapting to life in Roswell, among others.

4.2 The Speakers

The speakers' sociodemographic backgrounds can be summarized in the following way (see Tables 3 and 4): The speakers were born in various regions of Mexico: Mexico City (8), Acapulco, Guerrero (2), the state of Guerrero (1), Juando, Mexico (1), the state of Zacatecas (1), Cuernavaca, Morelos (1), the state of Morelos (1), Tampico, Tamaulipas (1), San Juan del Río, Querétaro (1), Monterrey, Nuevo León (1), the state

of Colima (1), and the state of San Luis Potosí (1). They consist of 12 females and 8 males, and their ages range from 25 to 60. Additionally, their LORs in the U.S. range from 2 to 25 years (average = 12 years), and their AOAs range from 11 to 56 (average = 27). In terms of education levels, they range from primary school to university. The speakers have a variety of occupations, nearly half of them being small business owners. Finally, their English proficiency ranges from very poor to good. English proficiency was measured using speakers' self-ratings on a scale from 1 (very poor) to 5 (very good), with an average rating of 2.75.[7]

Table 3: Speaker demographics (females)

Speaker	**City/State of Origin**	**LOR**	**AOA**	**Education**	**Occupation**	**English proficiency**
F39Mex	Mexico City, D.F	14	25	Partial law school	Owner of clothing boutique	2
F49Mex	Juando, Mexico	7	42	Primary school	Restaurant-food prep	3
F56Mex	Acapulco, Guerrero	25	31	Secondary school	Owner of tax business	4
F34Mex	Mexico City, D.F	10	24	University	Owner of clothing store	4
F25Mex	Zacatecas (state)	13	12	High school + Cosmetology school	Restaurant worker	4
F26Mex	Monterrey, Nuevo León	12	14	High school (U.S.)	Secretary	4
F52Mex	Mexico City, D.F	2	50	High school	Owner of sewing business	1
F43Mex	San Luis Potosí (state)	24	19	High school + Cosmetology school	Cosmetologist	3
F60Mex	Colima (state)	4	56	High school	Hairstylist	1
F32Mex	Acapulco, Guerrero	7	25	University	Owner of jewelry store	2
F30Mex	Mexico City, D.F	10	20	High school + Cosmetology school	Esthetician	3
F28Mex	Guerrero (state)	2	26	Master's	Cashier	3

[7] Self-ratings are also employed by Otheguy, Zentella, and Livert (2007) and Otheguy and Zentella (2012), among others. According to Roever and Powers (2005), the use of self-reports for second-language proficiency has proven a reliable assessment when compared to direct test measures.

Table 4: Speaker demographics (males)

Speaker	City/State of Origin	LOR	AOA	Education	Occupation	English proficiency
M51Mex	Cuernavaca, Morelos	10	41	Primary school	Landscaping	1
M41Mex	Mexico City, D.F	13	28	Partial university	Owner of clothing boutique	3
M34Mex	Morelos (state)	10	24	University	Owner of computer repair shop	3
M33Mex	Mexico City, D.F	12	21	Partial University	Owner of appliance store	3
M32Mex	Tampico, Tamaulipas	16	16	Secondary school	Manager of grocery store	3
M27Mex	Mexico City, D.F	16	11	High school (U.S.)	Auto body repair	4
M43Mex	San Juan del Río, Querétaro	25	18	Partial high school	Carpenter	2
M52Mex	Mexico City, D.F	15	37	Partial secondary school (2 years)	Owner of bakery	2

4.3 The Variable Context

In order to analyse SPE, the audio data were first transcribed. I then extracted all finite verbs from the interviews in order to locate each instance of SP usage. In order to isolate only cases in which variation between an overt and null SP can occur in Spanish, I have excluded the following types of tokens that fell outside the variable context: verbs within subject headed relative clauses (e.g. *mi mami este es una gran mujer ah que* ***Ø*** *también nos* ***ha inculcado***…'my mom umm is a great woman uh that has also instilled in us…' [F39Mex]); verbs appearing with full noun phrases (e.g. *Mi esposa* ***es*** *de, nacida en Virginia* 'My wife is from, born in Virginia' [M27Mex]); existential structures (e.g. *haber*, *ser*); *hacer* + time expressions (e.g. ***Ø hace*** *diez años vine* 'ten years ago I came' [M34Mex]); verbs with inanimate referents (e.g. ***Ø son*** *buenos* 'they're good' [M30Mex, referring to *tacos*]); impersonal *se* expressions (e.g. *en la casa* ***Ø se habla*** *español* 'at home Spanish is spoken' [F43Mex]); imperatives (e.g. *siempre me dijo,* ***"Ø estudia, Ø agarra*** *un libro*…" 'He always told me, "study, get a book…"' [F52Mex]); set phrases where an overt or null SP was categorical (e.g. *¿Qué sé* ***yo****?* 'I don't know'; ***Ø*** *digamos* 'let's say'). Speakers did not alternate between an overt and null SP in any of the above cases; thus, these structures were excluded.

For all tokens within the variable context (N = 4,705), I first coded whether each verb appeared with a null or overt SP. Subsequently, I coded for the independent variables using the categorizations discussed in the following sections.

4.4 Linguistic Variables

Given that one of the primary objectives of this analysis is to investigate issues of contact-induced language change, I have chosen to include independent variables that would facilitate comparability between my results and studies of monolingual Mexican Spanish. Therefore, I have included the same linguistic variables that were observed by

Lastra and Martín Butragueño (2015) in their study of Mexico City Spanish:[8] Person/number, Switch-reference, TMA, Morphological ambiguity, Verb class, Verbal mood, Specificity, Polarity, and Style (see Table 5).

Table 5: Linguistic variables and categories

Variable	**Categories**
Person/number	first-person singular (*yo* 'I') second-person singular (*tú* 'you' [informal]) third-person singular (*él/ella* 'he/she, *usted* 'you' [formal]) first-person plural (*nosotros/nosotras* 'we' [masc./fem.]) third-person plural (*ellos/ellas* 'they' [masc./fem]).
Switch-reference	Same-reference switch-reference partial-switch: (subjects that are coreferential with preceding objects) (*yo tengo un hijo, pero **él está**, está en, en Tijuana* 'I have a son, but he is, he's in, in Tijuana' [M43Mex])
TMA[9]	Present indicative, preterit, imperfect, perfect (including present perfect and pluperfect), present subjunctive, past subjunctive, synthetic future, periphrastic future, conditional *Imperative mood was not included since it was categorically null.
Morphological ambiguity	Ambiguous: first and third-person singular verbs in the imperfect, subjunctive (present or past), conditional, or pluperfect. Unambiguous: All other verbs.
Verb class[10]	Mental processes: (e.g. *creer* 'believe', *pensar* 'think', *saber* 'know') stative verbs: (e.g. *ser* 'be', *estar* 'be', *tener* 'have') *verba dicendi* (i.e. communication, e.g., *decir* 'say', *hablar* 'speak') activity verbs (e.g. *jugar* 'play', *hacer* 'do/make', *ir* 'go')
Verbal mood	Indicative Subjunctive
Specificity[11]	Specific reference Nonspecific reference
Polarity[12]	Affirmative Non-affirmative (both negative and interrogative)
Style[13]	First part (shorter speech turns, more dialogic pairs, more interview-like)

[8] Two variables not included in the present study that were included by Lastra and Martín Butragueño (2015) are textual genre (argumentation, description, dialogue, narrative) and pronoun position (preverbal, postverbal). Textual genre was not included due to the scarcity of narrative tokens in the current data. Pronoun position, while presented as an independent variable in Lastra and Martín Butragueño (2015), is not in fact an independent variable and does not feasibly function as one. In an earlier version of their study, these researchers clarify that they analyse how other independent variables constrain pronoun position rather than employing pronoun position as an independent variable itself. This factor was not included in the current study.

[9] To facilitate comparison with Lastra and Martín Butragueño's (2015:43) study, I collapsed these nine categories into four: Present, preterit, imperfect, and Other tenses.

[10] Categories are based on Bentivoglio, Ortiz, and Silva-Corvalán (2011).

[11] Specificity was operationalized following Otheguy and Zentella (2012:255) for the factor they call 'definiteness'.

[12] I adopt Lastra and Martín Butragueño's (2015) categories for what they call 'enunciative type'.

[13] Style was operationalized following Lastra and Martín Butragueño (2015). In the current data, first part tokens were those that occurred within the first 10-15 minutes of each interview when speakers were asked about their demographic background and when the questions were less open-ended. This is when the speakers' responses were generally much shorter. After the first 10-15 minutes, the questions were more open-ended and the speakers produced much longer responses. The interview flowed more like a conversation at this point, and tokens were coded as Second part.

	Second part: longer speech turns by interviewee (more conversation-like

4.5 Social Variables

I have included the following social variables: Age, Gender, AOA, LOR, English proficiency, Preferred media language, Educational level, Socioeconomic status (SES) (see Table 6). The selection of these variables was primarily based on their inclusion in previous studies of SPE in Spanish, especially those in situations of language contact (Otheguy & Zentella 2012; Michnowicz 2015).[14] Certain variables, such as preferred media language, educational level, and SES, were also chosen due to there being relatively little research with regard to their influence on SPE. Regarding SES, I have adopted the operationalization of Otheguy and Zentella (2012:271). These researchers used a combined measure based on both education and occupation of the speakers. Specifically, speakers were given points ranging from 1 to 4 in the following way:

> Education: elementary (1), secondary (2), college (3), graduate (4)
>
> Occupation: unskilled (1), skilled blue collar, clerical (2), store owner, manager, white collar (3), professional, business owner (4).

The speakers' SES was then based on combining the point values for each of the two categories and classifying them into four SES categories: (a) 1-2 points, (b) 3-4 points, (c) 5-6 points, and (d) 7-8 points.[15]

[14] Many of the social variables are coded differently by different researchers, as one would expect given that social categories differ across societies. While I attempted to follow the categorization of other studies as closely as possible, I found it advisable to make modifications due to the nature of the current data, namely the distribution of speakers according to age, length of residency, age of arrival, etc. Some examples of categorizations in previous studies include the following: **Age**: 20-34, 35-54, 55+ (Lastra & Martín Butragueño 2015); 13-19, 20-39, 40-60, 60+ (Otheguy & Zentella 2012); 16-29, 30-49, 50-70 (Carvalho & Child 2011); **length of residency**: 0-9, 10-19, 20+ (Otheguy & Zentella 2012); < 15, 16-30, 31-45, > 45 (Flores-Ferrán 2004); **age of arrival**: 3-12, 13-19, 20+ (Otheguy & Zentella 2012).

[15] An anonymous reviewer asserted that including SES and education as separate variables is problematic in this particular study, especially since SES comprises only education and occupation (and not additional factors such as income, place of residence, etc.) as part of its operationalization. Additionally, s/he pointed out the redundancy of studying both education and occupation as part of SES. I thank the reviewer for these insights and will consider these issues in future research. For the current study, however, the methods were maintained as to facilitate comparability with previous research.

Table 6: Social variables and categories

Variable	**Categories**
Age	20-34 35-49 50+
Gender	Male Female
LOR	1-7 years 10-14 years 15-25 years
AOA	11-19 20-28 30-56
English proficiency	Self-rating, scale of 1 (very poor) to 5 (very good)
Preferred media language (television, music, radio, newspapers) (cf. Travis & Torres Cacoullos 2013)	Spanish English Both
Educational level (Otheguy & Zentella 2012)	Elementary, Secondary, College, and Graduate
SES (Otheguy & Zentella 2012)	1-2 points 3-4 points 5-6 points 7-8 points

4.6 Statistical Methods

To determine the statistical significance of the linguistic and social variables and the relative weight of each factor regarding its conditioning on SP use, mixed-effects multivariate analyses (logistic regression) were carried out using *Rbrul* (Johnson 2009) with the inclusion of the speaker as a random effect. A multivariate analysis allows us to determine the relative effect of multiple factors at once. Specifically, it will facilitate the production of ranking for the statistically significant factors based on the amount of variation explained by each factor. Moreover, the quantitative results assist in creating a ranking of values within a single factor group or variable (e.g. person/number)—i.e. values within a factor group that favour overt SPs (e.g. first-person singular) vs. those factors that disfavour SPs (e.g. third-person plural). In addition, the inclusion of the speaker as a random effect in the multivariate analysis is a way of controlling for the individual speaker and ensuring that the results obtained are generalizable to the data set as a whole, and that the patterns are not due to particular speakers skewing the results (Bayley, Greer, & Holland 2013; Shin 2014; Michnowicz 2015; Shin & Van Buren 2016).

5. Results

5.1 Pronoun Rate

Out of 4,649 total verbs, 1,239 appeared with overt SPs and 3,410 appeared with null SPs, indicating an overall overt pronoun rate of 27% for Mexicans in Roswell. When we compare this rate to other Mexican dialects (see Table 7), we see that it is marginally higher than monolingual Mexican varieties, but not far outside the range (16%-25%) of these varieties. Regarding the comparison with the benchmark variety

of Mexico City, we see a five percent rate increase in Roswell, a finding that will be further discussed in Section 6.

Table 7: Comparison of pronouns rate with monolingual Mexican varieties

Variety	**PN Rate**	**Source**
Yucatan, Mexico	16%	Michnowicz 2015
Valladolid, Mexico	19%	Solomon 1999
NYC Mexican newcomers	19%	Shin & Otheguy 2013
Mexico City	22%	Lastra & Martín Butragueño 2015
Washington/Montana Mexican monolinguals	22%	Shin & Van Buren 2016
Xalapa, Mexico	25%	Orozco 2016
Roswell Mexicans	**27%**	**Current study**

5.2 Linguistic Predictors

In the multivariate analysis, six variables were selected as significant: *Person/Number*, *Switch-reference*, *Morphological Ambiguity*, *Verb Class*, *Polarity*, and *Style*. *TMA*, *Specificity*, and *Verbal Mood* were not significant. Table 8 shows the variable hierarchy of the significant factors from strongest to weakest effect. The top-ranked factor, *Person/Number*, has the most predictive power for variable SP usage while the bottom factor, *Style*, is the least powerful. The variable hierarchy was determined based on factor weight ranges, which are an indication of the relative strength of each variable (Tagliamonte 2006, 2012).

Table 8: Variable hierarchy as compared to Mexico City[16]

Roswell, GA (present study)	**Range**	Mexico City (Lastra & Martín Butragueño 2015:43)	Range
1. Person/Number	54	1. Person/Number	32
2. Switch-reference	26	2. Switch-reference	31
3. Morphological Ambiguity	20	3. Verbal Mood	29
4. Verb Class	14	4. Polarity	23
5. Polarity	10	5. TMA	20
6. Style	8	6. Morphological Ambiguity	13
~~TMA~~		7. Style	8
~~Verbal Mood~~		~~Verb Class~~	
~~Specificity~~		~~Specificity~~	

If we compare this hierarchy to that of Mexico City, we see both similarities and differences in ranking. The similarities include the following: *Person/Number* and *Switch-reference* are ranked first and second, respectively, for both varieties. Additionally, *Style* is ranked last for both varieties. Finally, *Specificity* was not selected as significant in either study. The differences between varieties involve the following: First, *Morphological Ambiguity* is stronger in Roswell, and *Polarity* is weaker (falling closer to the bottom of the hierarchy). Secondly, some constraints were lost in Roswell, namely *TMA* and *Verbal Mood*. Moreover, other constraints emerge as significant in Roswell, namely *Verb Class*.

[16] To facilitate comparison, only common factors employed between the studies appear here.

An additional difference is revealed for *Person/Number* and its relation to *Switch-reference* upon comparing the ranges. Particularly, we see that *Person/Number*, although ranking first in both varieties, appears to be much stronger in Roswell (54) than in Mexico City (32). Further, the predictive power compared to *Switch-reference* more than doubles in Roswell (*Person/Number* = 54 vs. *Switch-reference* = 26) and these two variables barely differ in relative importance for Mexico City (32 vs. 31).

Table 9 below presents the constraint hierarchies for each factor group. The first column shows each factor group along with their particular levels, and the second column presents the factor weights (FW) for each constraint from highest to lowest probability of appearing with an overt SP. When a FW is closer to 1, this indicates a relative favouring of overt SPs. When it is closer to 0, it generally indicates a disfavouring of overt SPs (see Tagliamonte 2006:145, 156).

Table 9: Constraint hierarchy

Factor Group	**Factor Weight**	**% Overt**	**Total N tokens**	***p*-value**
Person/Number				1.71e-59
3s	.75	38%	372	
1s	.69	36%	2565	
2s	.47	14%	489	
3pl	.39	12%	387	
1pl	.21	7%	836	
RANGE	54			
Switch-reference				6.39e-41
Switch	.64	35%	1945	
Partial-switch	.49	28%	247	
Same	.38	20%	2457	
RANGE	26			
Morphological Ambiguity				5.54e-05
Ambiguous	.60	46%	533	
Unambiguous	.40	24%	4116	
RANGE	20			
Verb Class				1.95e-07
Mental	.57	37%	962	
Stative/communicative[17]	.50	25%	1896	
Activity	.43	22%	1791	
RANGE	14			
Polarity				2e-04

[17] The categories of stative and communicative were ultimately combined because they patterned very similarly, showing no statistically significant difference between them. The rate of overt SPE was 25% for stative verbs and 27% for communicative verbs.

Affirmative	.55	27%	4063	
Non-affirmative	.45	23%	586	
RANGE	10			
Style				8.45e-05
Second part (conversation-like)	.54	28%	2988	
First part (interview-like)	.46	24%	1661	
RANGE	8			
Speaker (random) Std. Dev. .40				

5.2.1 Person/number

As shown in Table 9, both third-person singular and first-person singular verbs favour overt SPs with the former constraint showing the strongest effect (FW = .75). That these two forms favour overt SPs the most is generally consistent with previous research (e.g. Silva-Corvalán 1994a; Flores-Ferrán 2002; Shin 2012; Lastra & Martín Butragueño 2015; Limerick 2017). First-person plural verbs, in contrast, strongly disfavour overt SPs (FW = .21), followed by third-person plurals. Finally, second-person singular verbs show a neutral effect, a pattern that diverges from the pan-Hispanic trend of singular forms favouring overt pronouns and plural forms favouring nulls (Bentivoglio 1987; Cameron 1992; Silva-Corvalán 1994a; Flores-Ferrán 2002; Otheguy, Zentella, & Livert 2007; Prada Pérez 2009; Carvalho & Child 2011; Abreu 2012; Alfaraz 2015; Orozco 2015). However, this pattern follows other Mexican varieties in bucking this trend, namely first-generation Mexicans in NYC who show a neutral effect (Shin & Erker 2015) and both Mexico City (Lastra & Martín Butragueño 2015) and Xalapa (Orozco 2016) speakers who disfavour overt pronouns with second-person singular verbs.

5.2.2 Switch-reference

In line with previous studies, switch-reference contexts favour overt SPs, and same-reference contexts disfavour them (e.g. Travis 2005; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015). Moreover, partial-switches are neutral (FW = .49), which is also consistent with previous research (Orozco & Guy 2008; Otheguy & Zentella 2012; Orozco 2015). Specifically, overt SPs occur in this context 28% of the time, about the midpoint between 20% with same-reference and 35% with switch-reference.

5.2.3 Morphological ambiguity

Also consonant with previous research (e.g. Hochberg 1986; Travis 2007; Prada Pérez 2009; Erker & Guy 2012; Lastra & Martín Butragueño 2015; Michnowicz 2015), ambiguous verb forms favour overt SPs with a FW of .60 while unambiguous forms promote null SPs with a FW of .40. This finding lends support to the Functional Hypothesis (Hochberg 1986), which predicts an increased use of overt SPs with ambiguous verbal morphology for first-person singular and third-person singular forms in order to compensate for the absence of explicit person marking.

5.2.4 Verb class

Regarding *Verb Class*, mental process verbs slightly favour overt SPs (FW = .57) and activity verbs favour null SPs (FW = .43). Stative/communication verbs show neither a favouring nor disfavouring, with a FW of .50. The pattern for mental verbs follows the trend in the variationist literature (e.g. Silva-Corvalán 1994a; Travis 2007; Torres Cacoullos & Travis 2010, 2011; Otheguy & Zentella 2012; Orozco 2015) and is perhaps attributable to the implied contrast expressed by the speaker (Silva-Corvalán 1994a). The finding for activity verbs also aligns with previous studies (e.g. Orozco & Guy 2008; Erker & Guy 2012; Otheguy & Zentella 2012; Orozco 2015). However, the effect of stative and communication verbs is inconsistent with previous research since these verb classes showed a neutral effect in the present study and a favouring of overt SPs in other studies (e.g. Travis 2007; Orozco & Guy 2008; Erker & Guy 2012; Otheguy & Zentella 2012; Orozco 2015).

5.2.5 Polarity

Despite its overall modest effect, the results for *Polarity* show that affirmative clauses slightly favour overt SPs and that non-affirmative forms slightly favour null SPs. This finding agrees with the results of Lastra and Martín Butragueño (2015) and Geeslin and Gudmestad (2016). It is possible that subject coreference across contiguous negated clauses explains the favouring of null SPs (Lastra & Martín Butragueño 2015).

5.2.6 Style

Albeit the weakest effect, *Style* also significantly constrains pronoun use, with second-part style (more conversational) slightly favouring overt SPs and first-part style (shorter speech turns) slightly favouring nulls. This pattern is consistent with the findings of Lastra and Martín Butragueño (2015) and Ávila-Jiménez's (1996) findings showing a favouring of overt SPs with more casual speech.

It is worth noting that the Style effect could be influenced by the data distribution. There could simply be a higher occurrence of contexts that favour overt SPs in the second part of the interview compared to the first part (e.g. switch-reference), thereby producing the result of more overt SPs in more conversational speech. After examining such distribution, I indeed found this to be true, particularly for switch-reference and mental verbs, whose occurrences more than doubled for second-part tokens (e.g. 1485 switch-reference tokens in second-part of interview vs. 707 tokens in first part).

5.2.7 The unexpected TMA finding

Aside from the significant variables, one of the most striking findings emerges when we consider one of the non-significant variables, namely *TMA*. As discussed in Section 3, *TMA* has been shown to be significant in an overwhelming number of studies in the variationist literature, and with a very clear distinction: imperfect forms, which typically favour overt SPs, and preterit forms, which typically disfavour such forms. There is generally a wide percentage difference of overt SPs in these two contexts in the literature. The result in the current data, however, shows a drastic narrowing of this distinction with a virtually identical percentage of overt SPs in both contexts: 30% with imperfect and 29% with preterit (see Table 10).

Table 10: Rate of overt SPE according to TMA with comparison to previous studies

TMA (present study) (p = 0.0992)	**% Overt**	**Total N tokens**	Lastra & Martin Butragueno (2015)	Shin & Van Buren (2016)	Orozco (2016)
Imperfect	**30%**	750	**31%**	**29%**	**29%**
Preterit	**29%**	886	**16%**	**21%**	**20%**
Present	25%	2584	21%	19%	25%
All other TMAs	28%	429	16%	24%	24%

This divergence from previous data evidences a loss of the *TMA* constraint in the Roswell data, particularly with regard to the imperfect and preterit environments, and also contradicts Silva-Corvalán's (2001) hypothesis concerning the discourse functions of foregrounding and backgrounding. These functions are not observed in the present data since the speakers did not exhibit differential pronoun behaviour according to whether they used imperfect or preterit aspect. The alternative explanation discussed previously for the *TMA* effect has to do with the potential ambiguity of the subject referent. Despite the lack of statistical significance for the *TMA* variable, the current results do not necessarily contradict such explanation since there was indeed a significant effect for the *Morphological Ambiguity* variable.

5.3 Social Predictors

For the multivariate analysis of social variables one factor, age, was found to be a significant predictor of SP variation. None of the other social variables included—*Gender, LOR, AOA, English Proficiency, Preferred Media Language, Socioeconomic Status,* and *Education*—showed a significant effect. Table 11 shows the constraint hierarchy for *Age.*

Table 11: Constraint hierarchy for Age

Factor Group	**Factor Weight**	**% Overt**	**Total N tokens**	***p*-value**
Age				0.0339
50+	.59	34%	1079	
35-49	.46	24%	1135	
20-34	.45	24%	2435	
RANGE 14				

These results indicate that older speakers (50+) display the highest rates of SP usage at a 34% rate while younger speakers (35-49 and 20-34) slightly disfavour overt SPs at a rate of 24%. This distinction between older and younger speakers is generally in line with previous studies (e.g. Orozco & Guy 2008; Carvalho & Child 2011; Alfaraz 2015; Lastra & Martín Butragueño 2015; Orozco 2015, 2016), including those of other Mexican Spanish varieties (Lastra & Martín Butragueño 2015; Orozco 2016).

5.4 Length of Residency (LOR), Age of Arrival (AOA), and English Proficiency

Although the factors of *LOR*, *AOA*, and *English Proficiency* were not selected as statistically significant, it can be informative to explore their general trends, particularly

for shedding light on issues of language contact. Table 12 below shows the SP rate for each factor.

Table 12: SP rates according to LOR, AOA, and English proficiency

Factor Group	**% Overt**
LOR	
1-7 years	30%
15-25 years	26%
10-14 years	25%
AOA	
30-56	31%
20-28	25%
11-19	23%
English Proficiency	
4 (good)	29%
1-2 (poor)	28%
3 (moderate)	24%

Overall, the trends contradict a contact hypothesis based on English transfer. First, regarding *LOR*, the highest pronoun users are those that have been in the U.S. for the shortest amount of time (1-7 years) while more established immigrants have lower pronoun rates. Secondly, regarding *AOA*, the highest pronoun users are those with the oldest ages of arrival (30-56) whereas younger arrivals have lower rates. *English Proficiency* is difficult to interpret as there is not a linear trend. Nonetheless, we see no evidence in favour of English transfer (e.g. speakers who rated their English as 1-2 and 4 have virtually identical pronoun rates). These trends differ from what researchers have found for U.S. Spanish varieties under contact-induced change (e.g. Lapidus & Otheguy 2005; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012).

6. Discussion

6.1 SP Rate Comparisons

Roswell speakers demonstrated an SP rate of 27%, which slightly diverges from most monolingual Mexican varieties. For instance, the rate found for Mexico City was 22% (Lastra & Martin Butragueno 2015) and the rate for NYC newcomers was 19% (Shin & Otheguy 2013). As stated above, however, 27% is still within the general dialectal range for Mexican Spanish; it is a relatively low pronoun rate when compared to other regional varieties, which is characteristic of Mexican Spanish. In comparing Roswell Mexicans to other contact varieties, Table 13 shows that Roswell Spanish falls

approximately in the middle of the New Jersey (24%) and Los Angeles (29%) Mexican varieties.[18]

Table 13: Comparison of pronoun rate in Roswell to other contact varieties

Variety	**PN Rate**	**Source**
New Jersey Mexicans	24%	Flores Ferrán 2007
Roswell Mexicans	**27%**	**Current study**
Los Angeles Mexicans	29%	Silva-Corvalán 1994a
New York born/raised (Mexican, S. American, Caribbean)	38%	Otheguy, Zentella, & Livert 2007

6.2. Language-internal and Language-external Variables

Overt SPs were found to be favoured in the following environments: first-person singular and third-person singular verbs; switch in subject referent; morphologically ambiguous verbs; mental process verbs; affirmative polarity environments; and conversation-like speech style. In comparing these results with previous sociolinguistic studies, they are consistent with what other researchers have found, and also parallel the Mexico City data analyzed by Lastra and Martín Butragueño (2015). Three additional linguistic variables that were included in the present analysis but that were not found to be significant were *Specificity*, *TMA*, and *Verbal Mood*. One of these nonsignificant factors, *TMA*, was particularly striking as this was an unexpected finding given its widely-replicated significance in previous research. It appears that Roswell speakers are experiencing a loss of sensitivity to *TMA*; however, further research is warranted in order to explore the motivations for this finding.

The language-external factor that played a role in SP usage for Roswell speakers was their age. Younger and middle-aged speakers slightly disfavoured the use of overt SPs (24%), while older speakers showed a favouring of overt SPs (34%). This finding corroborates previous research that has reported a distinction between younger and older speakers (Orozco & Guy 2008; Carvalho & Child 2011; Alfaraz 2015; Lastra & Martín Butragueño 2015; Orozco 2015, 2016). This observation in the current data could be an incidental reflection of the differing regional origins of the speakers according to age. That is, it could be the case that the older speakers simply hail from regions in Mexico in which overall overt SPE is higher and that the younger speakers come from lower overt SPE regions. Further research of this issue is warranted.[19]

This finding deserves further discussion, specifically as it relates to language continuity and dialect contact in Roswell. As similar age effects have been observed in monolingual Mexican varieties (Lastra & Martín Butragueño 2015; Orozco 2016), it could be the case that this pattern in the Roswell data simply reflects a continuation of trends that began in Mexico. Contact with other Mexican Spanish-speakers may be reinforcing the age effect, resulting in the maintenance of this trend for Roswell speakers and an inhibiting of the type of pronoun increases among younger speakers that might be expected in more established U.S. Spanish varieties. After all, there are several Mexican Spanish dialects in contact in Roswell, which are represented in the

[18] It should also be noted that the current rate of 27% differs from the rate found in Limerick (2017), which also examines Mexican Spanish in Roswell (21% overt SPE). This rate discrepancy could be attributed to sample size and coding differences between the two studies. For instance, 20 speakers were employed in the current study while only 10 were included in Limerick (2017). Furthermore, there are differences in circumscribing the variable context between the studies (See Otheguy and Zentella 2007 for a similar perspective on discrepancies in findings).

[19] I thank an anonymous reviewer for suggesting this potential explanation of the age effect.

current sample. Further, the Roswell speakers reported having more interaction with Spanish-speakers than English-speakers overall in the sociolinguistic interviews. Thus, the speakers' SP patterns for age are likely being reinforced.

6.3 The Role of English/bilingualism

In terms of overall pronoun rates, we observed a marginally higher rate for Roswell Mexicans (27%) than what was reported for Mexico City Spanish (22%, Lastra & Martín Butragueño 2015). Is this sufficient evidence to conclude that Mexicans in Roswell have been influenced by English in terms of pronoun rate? From the perspective of this investigation, an increase in overall pronoun rate alone is not sufficient to make a determination regarding contact-induced change. As previously noted, using overall frequencies to assess language contact has been called into question (e.g. Poplack & Levey 2010; Torres Cacoullos & Travis 2010), and a similar perspective is adopted here. Specifically, it has been widely observed that frequencies of overt SPs differ substantially across dialects (e.g. Poplack & Levey 2010; Torres Cacoullos & Travis 2010; Otheguy & Zentella 2012; Alfaraz 2015; Orozco 2015), and it cannot be ruled out that the increased usage observed in the current data is not simply a reflection of regional variation.[20] In fact, it was shown earlier that within Mexico SP rates vary. Could it be then that the differing regional origins of the current speakers explain this increase when compared to Mexico City? Subsequent calculations of the pronoun rate for exclusively Mexico City speakers (N=8) vs. speakers of other Mexican dialects (N=12) in the current data revealed the following: Mexico City speakers alone show a rate of 24% (closer to the 22% benchmark), and the remaining speakers, 28%, falling closer to the overall average (27%). Thus, the data suggest that dialect contact is playing a role. Since MC speakers are in contact with other Mexican speakers, these trends for frequency that have a source in Mexico appear to be reinforced due to the dialect contact situation. For example, the source dialects of the Roswell speakers may also exhibit SP rates at the level of 27%. The evidence from the current data suggests, then, that the higher rate observed for the non-Mexico City dialects is what is driving up the overall pronoun rate. The different regional Mexican varieties in contact with each other in Roswell would be reflected in the overall rate of 27%.

Furthermore, upon observing trends in the current data for pronoun rates in relation to social factors such as *LOR, AOA,* and *English proficiency*, we do not see direct correlations. Longer established immigrants or those with higher English proficiency do not have higher pronoun rates. In fact, the trends observed pointed in the opposite direction of what an English transfer hypothesis would predict (Silva-Corvalán 1994a; Otheguy & Zentella 2012). The highest pronoun users were the least established immigrants and the oldest arrivals to the U.S. In addition, speakers with both the lowest and highest self-reported English levels had virtually the same overt frequencies.

Turning now to the linguistic constraints on SP variation, does bilingualism seem to be playing a role here? When comparing variable hierarchies between Mexican Spanish in Roswell and Mexico City Spanish, speakers in Roswell demonstrated some distinctions in relative position in the hierarchy and range values. For instance, the

[20] Additionally, there have been cases documented of Spanish in contact with pro-drop languages in which there was also an increase in overt SP usage (e.g. Prada Pérez 2015; Michnowicz 2015). For example, this was observed by Michnowicz (2015) in his analysis of Spanish in contact with Yucatec Maya and was interpreted as a simplification strategy for bilinguals whereby they lighten the cognitive load associated by simplifying pragmatic constraints. However, since no correlation was found in the current study between English proficiency and pronoun rates, the possiblity of a bilingual simplification strategy is not substantiated.

variables of *morphological ambiguity* and *person/number*, were stronger in Roswell, the range for *morphological ambiguity* was 20 (ranking third overall) compared to 13 (ranking sixth overall) in the results reported by Lastra and Martín Butragueño (2015). For *person/number*, a range comparison relative to *switch-reference* in both varieties revealed that the gap was much wider in Roswell (range difference of 28) than in Mexico City (range difference of 1), showing that *person/number* is a much more important factor than *switch-reference* for Roswell speakers whereas the predictive power of the two variables is about the same for speakers in Mexico City. In comparing this finding with other monolingual dialects, however, we find that the Mexico City data is an exception. Other studies of monolingual Mexican Spanish mirror the current findings in that person/number is undoubtedly the strongest and the gap in strength between person/number and switch-reference is notably wide (Solomon 1999:250; Michnowicz (2015:112); Orozco (2016:5,7).[21] Contrastively, the gap in the Mexico City data is almost non-existent, indicating a rather atypical pattern for Mexico City speakers. Thus, Roswell speakers are actually showing patterns more like most other monolingual dialects. Another constraint divergence was for *polarity*. This was shown by the range value of 10 in the current results, which is more than twice as weak as the strength observed in Mexico City (23). Finally, *TMA* was not significant in the Roswell data, but was significant for Mexico City monolinguals.

What do such divergences suggest in terms of the potential linguistic processes at work for speakers in the current data? As discussed in previous literature, processes such as *simplification*--the weakening of discourse or morphosyntactic constraints to lighten the cognitive load (Silva-Corvalán 1994b; Sorace 2004)--as well as *complexification* (the emergence of constraints) (Shin 2014) can sometimes play a role as an influence of bilingualism in other contact varieties of Spanish. However, for the present analysis, there is little evidence of bilingualism having such effects. Rather, as in the case for SP frequencies, the evidence here suggests that any differences in constraints are simply due to dialectal variation in the data. Crucially, since English proficiency (level of bilingualism) was not shown to be a significant factor, there is no evidence that bilingualism is playing a role. Despite the divergences in the variable hierarchies and the apparent weakened sensitivity to certain constraints for Roswell speakers, we would only be able to point in the direction of a bilingualism effect if we had seen correlations between speakers with the highest English proficiency and higher pronoun rates, a piece of evidence absent from the current investigation.[22] .

In addition, we see ample evidence of linguistic continuity in the variable pronoun system between Mexico City and Roswell speakers. It was found that the first two variables in the hierarchy were the same and in the same order according to importance: *Person/number* ranked 1st and *switch-reference* ranked 2nd in both studies; *style* was ranked last, showing the least amount of importance for both varieties; and the two varieties shared the lack of influence of *specificity* on SP variation. More importantly, upon comparing the *constraint hierarchies* (ordering of constraints within each variable), it was demonstrated that Roswell and Mexico City speakers share the same

[21] Michnowicz (2015): 57 vs. 31; Solomon (1999): 43 vs. 11; Orozco (2016): 50 vs. 28 for *verb class* [switch- reference was not included in Orozco's study; verb class was the second strongest predictor].

[22] It may also be the case that differences in data distribution and configuration between the current analysis and that of Lastra and Martín Butragueño (2015) are motivating these constraint divergences in terms of strength of effect. For example, regarding person/number, more than half of the tokens in the Roswell data are 1sg (strongly favouring overt SPs); 1pl tokens (the second most frequent), which strongly disfavour overt SPs, account for 20% of the data. In Lastra and Martín Butragueño's (2015:43) study, however, these two extremes for person/number make up smaller proportions of their data (46% vs. 9%), which may make the Person/Number effect appear weaker when compared to the current study.

patterns for each significant variable in terms of which contexts favour the use of overt SPs, and regarding the hierarchical ordering of these contexts. Therefore, considering the above comparisons, we have primarily a case of linguistic continuity in the Mexican Spanish of Roswell, Georgia. Since Roswell speakers use Spanish most of the time and are primarily in contact with Spanish-speakers who also have limited exposure to English, their SP rates and usage patterns are generally maintained. This is consistent with what has been found for other U.S. varieties with tight-knit social networks and limited exposure to English (Shin & Otheguy 2013; Shin & Van Buren 2016). Moreover, the differences in SP usage observed in the current data, rather than having their root in bilingualism and English contact, point toward a reflection of dialectal variation and dialect contact. Further research with more robust dialect groups is needed, however, to confirm this (see, e.g., Otheguy & Zentella 2012 for a substantive treatment of dialect contact effects).

7. Conclusion

Since no evidence of English contact or bilingualism was found, it is more likely the situation of dialect contact emerging in Georgia that promotes the formation of this unique variety. That is, it is not bilingualism that is driving this emergence, but rather the heterogeneous interaction and combination of different Mexican varieties. Being only in the initial stages of language contact (first generation immigrants) and not yet being a well-established variety (i.e. second or third generation immigrant Spanish) with extensive contact with English, we are not (yet) observing evidence of English contact influence. Rather, we are seeing structural continuity on one hand, while simultaneously seeing a reflection of signs of dialects in contact in Georgia regarding SPE, especially since the speakers overall interact more with other Spanish-speakers than English speakers. The present findings reveal insights into the linguistic mechanisms at work in the initial stages of contact (i.e. Spanish in Georgia as a site of recent emergence of Spanish-speaking populations), which is something obscured in the sociolinguistic SP research on Spanish in the U.S. that primarily examines more established varieties.

Future studies should investigate several issues. It would be useful to incorporate a wider range of linguistic variables and to also consider interaction effects between variables. Additionally, it is crucial to explore issues that will shed more light on the role of language/dialect contact. For instance, cross-generational data would be helpful in revealing how the Spanish of second- and third-generation immigrants within this community compares with that of first-generation speakers. Concerning dialect contact, additional data of other dialect groups are essential to fully explore this issue. Aside from the opportunity to study contact between dialects, employing speakers of diverse regional and national origins would contribute to a more representative analysis of Spanish-speakers in the Southeastern U.S.

Philip P. Limerick
Eastern Kentucky University
Department of Languages, Cultures, & Humanities
Richmond, KY. 40475
philiplimerick@eku.edu

References

Abreu, L. (2012). Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish. In K. Geeslin & M. Díaz-Campos (Eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 1-8.

Alfaraz, G. G. (2015). Variation of overt and null subject pronouns in the Spanish of Santo Domingo. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown: Georgetown University Press, pp. 3-16.

Ávila-Jiménez, B. (1996). *Subject pronoun expression in Puerto Rican Spanish: A sociolinguistic, morphological, and discourse analysis*. PhD dissertation, Cornell University.

Barrenechea, A.M., & A. Alonso. (1977). Los pronombres personales sujetos en el español hablado en Buenos Aires. In J.M. Lope Blanch (Ed.), *Estudios sobre el español hablado en las principales ciudades de América*. Mexico City: Universidad Nacional Autónoma de México, pp. 333–49.

Bayley, R., K. Greer & C.C. Holland. (2013). Lexical frequency and syntactic variation: A test of a linguistic hypothesis. *University of Pennsylvania Working Papers in Linguistics* 19(2), pp. 21-30. https://doi.org/10.1017/S0954394500001964

Bayley, R., & L. Pease-Alvarez. (1996). Null and expressed pronoun variation in Mexican-descent children's Spanish. In J. Arnold, R. Blake & B. Davidson (Eds.), *Sociolinguistic variation*: *Data, theory, and analysis* Stanford: Center for the Study of Language and Information.

Bayley, R. & L. Pease-Alvarez. (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9, pp. 349–71.

Bentivoglio, P. (1987). *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Central University of Venezuela.

Bentivoglio, P., L.A. Ortiz, & C. Silva-Corvalán. (2011) La variable expresión del sujeto pronominal. Guía de codificación. In *PRESEEA – Grupo de análisis de expresión de sujetos*. Retrieved from *http://preseea.linguas.net/portals/0/metodologia /guia_codificacion _ sujetos_julio_2011.pdf*

Blackwell, S.E. & M.L. Quesada. (2012). Third-person subjects in native speakers' and L2 learners' narratives: Testing (and revising) the givenness hierarchy for Spanish. In K. Geeslin & M. Díaz-Campos (eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, pp. 142-64.

Cameron, R. (1992). *Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation*. Philadelphia, PA: PhD dissertation, University of Pennsylvania.

Cameron, R. (1993). Ambiguous agreement, functional compensation, and nonspecific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5, pp. 305-34. https://doi.org/10.1017/S0954394500001526

Cameron, R. (1994). Switch reference, verb class and priming in a variable syntax. In *Papers from the regional meeting of the Chicago Linguistic Society*: Parasession on variation in linguistic theory 30, pp. 27–45.

Carter, P.M. (2014). National narratives, institutional ideologies, and local talk: The discursive production of Spanish in a 'new' US Latino community. *Language in Society* 43(2), pp. 209-40. https://doi.org/10.1017/S0047404514000049

Carvalho, A.M., & R.M. Bessett. (2015). Subject pronoun expression in Spanish in contact with Portuguese. In A.M. Carvalho, R. Orozco, & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown: Georgetown University Press, pp. 143-65.

Carvalho, A.M. & M. Child. (2011). Subject pronoun expression in a variety of Spanish in contact with Portuguese. In J. Michnowicz & R. Dodsworth (Eds.), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* Somerville, MA: Cascadilla Proceedings Project, pp. 14-25.

Enríquez, E. (1984). *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: CSIC.

Erker, D. & G.R. Guy. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88, pp. 526-57. https://doi.org/10.1353/lan.2012.0050

Flores, N. & J. Toro. (2000). The persistence of dialect features under conditions of contact and leveling. *The Southwest Journal of Linguistics* 19(2), pp. 31-42.

Flores-Ferrán, N. (2002). *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: A sociolinguistic perspective*. Munich: Lincom.

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, pp. 49-73. https://doi.org/10.1017/S0954394504161048

Flores-Ferrán, N. (2007). Los mexicanos in New Jersey: Pronominal expression and ethnolinguistic aspects. In J. Holmquist, A. Lorenzino & L. Sayahi (Eds.), *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics.* Somerville, MA: Cascadilla Proceedings Project, pp. 85-91.

Geeslin, K., & A. Gudmestad. (2016). Subject expression in Spanish: Contrasts between native and non-native speakers for first and second-person singular referents *Spanish in Context* 13(1), pp. 53-79. https://doi.org/10.1075/sic.13.1.03gee

Geeslin, K., B. Linford, & S. Fafulas. (2015). Variable subject expression in second language Spanish: Uncovering the developmental sequence and predictive linguistic factors. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown: Georgetown University Press, pp. 191-210.

Hochberg, J.G. (1986). Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62(3), pp. 609-21.

Johnson, D.E. (2009). Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3(1), pp. 359–83. https://doi.org/10.1111/j.1749-818X.2008.00108.x

Klee, C.A. & A. Lynch. (2009). *El español en contacto con otras lenguas.* Washington, D.C.: Georgetown University Press.

Labov, W. (1972). *Language in the inner city*. Philadelphia: University of Pennsylvania Press.

Lamanna, S.G. (2012). *Colombian Spanish in North Carolina: The role of language and dialect contact in the formation of a new variety of U.S. Spanish*. PhD dissertation, Indiana University.

Lapidus, N., & R. Otheguy. (2005). Overt nonspecific *ellos* in Spanish in New York. *Spanish in Context* 2(2), pp. 157–74. https://doi.org/10.1075/sic.2.2.03lap

Lastra, Y., & P. Martín Butragueño. (2015). Subject pronoun expression in oral Mexican Spanish. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject*

pronoun expression in Spanish: A cross-dialectal perspective. Georgetown: Georgetown University Press, pp. 39-57.

Limerick, P. P. (2017). Language contact in the US Southeast: The case of Spanish subject expression in an emerging bilingual community in Georgia. *Spanish in Context* 14(1), pp. 53-77. https://doi.org/10.1075/sic.14.1.03lim

Lozano, C. (2009). Selective deficits at the syntax-pragmatics interface: Evidence from the CEDEL2 corpus. In N. Snape, Y. Ingrid Leung & M. Sharwood Smith (Eds.), *Representational deficits in second language acquisition*. Amsterdam: John Benjamins. pp. 127-66. https://doi.org/10.1075/lald.47.09loz

McKnight, R. (2013). *Pronombres de sujeto en Raleigh*. Raleigh, NC: Master's thesis, North Carolina State University.

Michnowicz, J. (2015). Subject pronoun expression in contact with Maya in Yucatan Spanish. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown: Georgetown University Press. pp. 101-119.

Montes-Alcalá, C., & L. Sweetnich. (2014). Español en el Sureste de EEUU: El papel de las actitudes lingüísticas en el mantenimiento o pérdida de la lengua. *Revista Internacional de Lingüística Iberoamericana* 12(23), pp. 77-92.

Morales, A. (1997). La hipótesis funcional y la aparición del sujeto no nominal: El español de Puerto Rico. *Hispania* 80, pp. 153–65. https://doi.org/10.2307/345995

Motel, S., & E. Patten. (2013). 2011, Hispanics in the United States statistical portrait' Retrieved from http://www.pewhispanic.org/2013/02/15/2011-statistical-information-on-hispanics-in-united-states/#change-in-the-hispanic-population-by-state-2000-and-2011

Olloqui de Montenegro, L. (1987). Un aspecto de la sintaxis: Los pronombres personales sujeto en el habla estudiantil santiaguera. In H. López Morales & M. Vaquero (Eds.), *Actas del I Congreso Internacional sobre el Español de América* San Juan, Puerto Rico: Academia Puertorriqueña de la Lengua Española. pp. 753–64.

Orozco, R., & G. Guy. (2008). El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana? In M. Westmoreland & J. Antonio Thomas (Eds.), *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics* Somerville, MA: Cascadilla Proceedings Project. pp. 70-80. https://doi.org/10.3765/plsa.v1i0.3703

Orozco, R. (2015). Pronominal variation in Colombian Costeño Spanish. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown: Georgetown University Press. pp. 17-37.

Orozco, R. (2016). Subject pronoun expression in Mexican Spanish: ¿Qué pasa en Xalapa? In *Proceedings of the Linguistic Society of America Vol. 1, Article 7*. pp. 1-15.

Otheguy, R., A.C. Zentella, & D. Livert. (2007). Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 83, pp. 770-802. https://doi.org/10.1353/lan.2008.0019

Otheguy, R. & A.C. Zentella. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199737406.001.0001

Poplack, S., & S. Levey. (2010). Contact induced grammatical change: A cautionary tale. In P. Auer & J. Erich Schmidt (Eds.), *Language and space: An*

international handbook of linguistic variation. Berlin: Mouton de Gruyter. pp. 391-419. https://doi.org/10.1515/9783110220278.391

Prada Pérez, A. de. (2009). *Subject expression in Minorcan Spanish: Consequences of contact with Catalan.* PhD dissertation, Pennsylvania State University.

Prada Pérez, A. de. (2015). First person singular subject pronoun expression in Spanish in contact with Catalan. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* Georgetown: Georgetown University Press. pp. 121-142.

Ranson, D.L. (1991). Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change* 3, pp. 133-52. https://doi.org/10.1017/S0954394500000491

Roever, C., & D. E. Powers. (2005). Effects of language of administration on a self-assessment of language skills. Princeton, NJ: Educational Testing Service.

Shin, N.L. (2012). Variable use of Spanish subject pronouns by monolingual children in Mexico. In K. Geeslin & M. Díaz-Campos (Eds.), *Selected proceedings of the 14th Hispanic linguistics symposium* Somerville, MA: Cascadilla Proceedings Project. pp. 130-141. https://doi.org/10.1017/S095439451400012X

Shin, N.L. (2014). Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change* 26, pp. 303-30.

Shin, N.L., & D. Erker. (2015). The emergence of structured variability in morphosyntax: Childhood acquisition of Spanish subject pronouns. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective.* Georgetown: Georgetown University Press. pp. 169-190.

Shin, N.L., & C. Montes-Alcalá. (2014). El uso contextual del pronombre sujeto como factor predictivo de la influencia del inglés en el español en Nueva York. *Sociolinguistic Studies* 8(1), pp. 85-110. https://doi.org/10.1558/sols.v8i1.85

Shin, N.L., & R. Otheguy. (2009). Shifting sensitivity to continuity of reference: Subject pronoun use in Spanish in New York City. In M. LaCorte & J. Leeman (Eds.), *Español en Estados Unidos y otros contextos: Sociolingüística, ideología y pedagogía* Madrid: Iberoamericana. pp. 111-36.

Shin, N.L., & R. Otheguy. (2013). Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. *Language in Society* 42(4), pp. 429-52. https://doi.org/10.1017/S0047404513000468

Shin, N.L., & J. Van Buren. (2016). Maintenance of Spanish subject pronoun expression patterns among bilingual children of farmworkers in Washington/Montana. *Spanish in Context* 13(2), pp. 173-94. https://doi.org/10.1075/sic.13.2.01shi

Silva-Corvalán, C. (1982). Subject expression and placement in Mexican-American Spanish. In J. Amastae & L. Elías-Olivares (Eds.), *Spanish in the United States: Sociolinguistic aspects* Cambridge: Cambridge University Press. pp. 93–120.

Silva-Corvalán, C. (1994a). *Language contact and change: Spanish in Los Angeles* Oxford: Clarendon.

Silva-Corvalán, C. (1994b). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change* 6(3), pp. 255-72. https://doi.org/10.1017/S095439450000168X

Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español* Washington, D.C.: Georgetown University Press.

Smith, D.J. (2006). Thresholds leading to shift: Spanish/English codeswitching and convergence in Georgia, U.S.A. *International Journal of Bilingualism* 10(2), pp. 207–40. https://doi.org/10.1177/13670069060100020501

Solomon, J. (1999). *Phonological and syntactic variation in the Spanish of Valladolid, Yucatán* Stanford, CA: PhD dissertation, Stanford University.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition* 7(2), pp. 143-45. https://doi.org/10.1017/S1366728904001543

Tagliamonte, S.A. (2006). *Analysing sociolinguistic variation* Cambridge, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511801624

Tagliamonte, S.A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Malden, MA: Wiley-Blackwell.

Torres Cacoullos, R., & C.E. Travis. (2010). Variable yo expression in New Mexico: English influence? In S. Rivera-Mills & D. Villa (Eds.), *Spanish of the U.S. Southwest: A language in transition* Madrid: Iberoamericana/Vervuert. pp. 185-206. https://doi.org/10.31819/9783865278692-012

Torres Cacoullos. R., & C.E. Travis. (2015). Foundations for the study of subject pronoun expression in Spanish in contact with English: Assessing interlinguistic (dis)similarity via intralinguistic variability. In A.M. Carvalho, R. Orozco & N.L. Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* Georgetown: Georgetown University Press. pp. 81-100.

Travis, C.E. (2005). The yo-yo effect: Priming in subject expression in Colombian Spanish. In R. Gess & E.J. Rubin (Eds.), *Selected Papers from the 34th Linguistic Symposium on Romance Languages (LSRL)* Amsterdam: John Benjamins. pp. 329–49.

Travis, C.E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19, pp. 101-35. https://doi.org/10.1017/S0954394507070081

Travis, C.E., & R. Torres Cacoullos. (2012). What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics* 23(4), pp. 711-48. https://doi.org/10.1515/cog-2012-0022

Travis, C.E., & R. Torres Cacoullos. (2013). Making voices count: Corpus compilation in bilingual communities. *Australian Journal of Linguistics* 33(2), pp. 170-194. https://doi.org/10.1080/07268602.2013.814529

Veltman, C. (2000). The American linguistic mosaic: Understanding language shift in the United States. In S. McKay & S.L. Cynthia Wong (Eds.), *New immigrants in the United States* Cambridge: Cambridge University Press. pp. 58-95.

Wilson, A. (2013). *Stories of Roswell, Georgia: A sociolinguistic study of narrative structure* Athens, GA: Undergraduate honors thesis, University of Georgia.

Winders, J. (2005). Changing politics of race and region: Latino migration to the US South. *Progress in Human Geography* 29(6), pp. 683–99. https://doi.org/10.1191/0309132505ph577oa