

Regional comparisons of subject pronoun expression among Mexican-origin immigrants in Georgia

Philip P. Limerick

Centre College, Danville, KY, USA <philip.limerick@centre.edu>

This paper offers a comparative analysis of variable subject pronoun expression among first-generation Mexican immigrants in the U.S. with differing regional origins. Using sociolinguistic interview data from Georgia, rates and usage patterns of first-person singular SPs were examined among two groups of speakers: those hailing from Mexico City and those from other regions, such as Guerrero, Zacatecas, and Colima, among others. From a variationist perspective, rates and constraints on SPs were examined across the two groups by means of logistic regression analyses in *Rbrul*, with results indicating wide variation in pronoun rates both across and within groups. This suggests that Mexican Spanish may exhibit a wider range of variation in SP rates than was previously assumed. Regarding usage patterns, the linguistic constraints on subject pronoun expression (switch reference, TMA, verb class, polarity) show remarkable similarities between Mexico City speakers and non-Mexico City speakers, suggesting cross-regional uniformity for first-person singular subject pronoun expression in Mexican dialects. Additionally, the social predictors of age and gender were examined, and, while no significant main effects were observed, interaction effects between the linguistic and social predictors were differentially operative between Mexico City and non-Mexico City speakers, reflecting a more nuanced view on the sociolinguistic conditioning of variable subject pronoun expression.

KEYWORDS: Mexican Spanish, subject pronoun expression, Spanish in the U.S., immigrant Spanish.

1. Introduction

The present study seeks to shed light on regional variation among Mexican Spanish-speakers through the lens of subject pronoun expression (SPE). In an immigrant context such as the U.S., where members of a speech community may share a national origin, but differ in regional origin, a unique opportunity is presented to make regional comparisons among speakers. This also can aid in our understanding of varieties of Spanish whose speakers are not homogeneous in terms of regional origin, an opportunity not as easily available outside an immigrant community. While most studies carry out the important and necessary work of focusing on speakers from one particular city or state (e.g. Lastra & Martín Butragueño 2015 [Mexico City]; Solomon 1999 and Michnowicz

2015 [Yucatan]; Orozco 2016 [Xalapa]), the current analysis emphasizes language variation as reflected by speakers across various cities/states who have immigrated to the U.S. In addition, this paper contributes to the field of sociolinguistics by offering a perspective on the interaction between internal (linguistic) predictors and external (social) predictors regarding their conditioning of language variation, specifically as it pertains to Mexican Spanish.¹ It is well-known that variation is impacted by both linguistic and social predictors. Likewise, extensive work demonstrates the interaction of, on the one hand, linguistic predictors (e.g. tense + reference), and on the other hand, the intersection of social predictors (e.g. gender + age). What is less studied, however, is the intersection of the two, that is, how the linguistic and social interact with each other to constrain language variation. The current study aims, among other issues, to explore the potential effects of such interactions on variable SPE.

The variable phenomenon under analysis, SPE (*YO hablo* ‘I speak’ vs *hablo* ‘[I] speak’), has been widely studied across Spanish-speaking countries, including the U.S., and among L1 speakers, heritage speakers, and L2 speakers of Spanish (e.g. Cameron 1994; Silva-Corvalán 1994; Flores-Ferrán 2004; Travis 2005, 2007; Orozco & Guy 2008; Shin & Otheguy 2009; Carvalho & Child 2011; Abreu 2012; Otheguy & Zentella 2012; Orozco 2015; Alfaraz 2015; Lastra & Martín Butragueño 2015; Michnowicz 2015; Torres Cacoullós & Travis 2015; Geeslin *et al.* 2015; Geeslin & Gudmestad 2016; Padilla 2021).² Fewer studies, however, emphasize SPE among first-generation immigrants with varying regional origins. This study will focus on immigrant Spanish in the Southeastern U.S., examining SPE for first-person singular subjects. This method of analyzing first-person singular forms exclusively follows that of other researchers who emphasize the importance of considering one grammatical person alone (e.g. Travis 2005, 2007; Travis & Torres Cacoullós 2012; Bessett 2018). In fact, most research on SPE examines all grammatical persons/numbers in the same analysis, which can obfuscate patterns for the individual person/number. As different grammatical persons of SPs (e.g. first- vs third-person) respond to different contextual and interactional factors (Travis 2005), it is important to continue the study of individual persons/numbers, thus far carried out most extensively for 1SG subjects (e.g. Travis 2005, 2007; Travis & Torres Cacoullós 2012; Prada Pérez 2015; Bessett 2018), and to a lesser extent for third-person singular (Shin 2014), third-person plural (Lapidus & Otheguy 2005), and first-person plural subjects (Posio 2012; Limerick 2021).

Sociolinguistic interview data from a Mexican immigrant community in Georgia (*Corpus of Spanish in Georgia* [Limerick 2022]; see Section 3 for details) will be employed in the current analysis. Pronoun usage will be examined in terms of overall frequency and constraints (see Section 2 for a discussion of frequencies and particular constraints on SPE in previous studies) governing the speakers' overt/null SP variation. The particular speech community under study consists of first-generation Mexican immigrants residing in the Atlanta metropolitan area (Roswell, Georgia). Latin American immigration to the U.S. Southeast has led to recent and significant demographic shift in this region. For instance, between 1990 and 2010 more than a million Latin Americans immigrated to Virginia, North Carolina, South Carolina, and Georgia, with the Latinx populations of North Carolina and Georgia having the highest percentage of growth (Wolfram *et al.* 2011). Thus, substantial Spanish-speaking populations have emerged that historically were not part of the Southeast in the way that they were in the Southwestern and Northeastern U.S.³

The city of Roswell, GA (see Figure 1 below) represents the larger demographic shift in the Southeast and is experiencing recent emergence of Spanish-speaking communities. For example, between 2000 and 2020, the Latinx population in Roswell grew from 11% to 16% (United States Census Bureau 2020). Therefore, Roswell is an excellent test site for studying emerging bilingual and bidialectal speech communities in the U.S. (Limerick 2017).



Figure 1. Roswell, Georgia. Map by Bert Sperling.

The notion of ‘speech community’ that will be adopted in this paper is that of Gumperz (1972 [1968]: 219), defined as “any human aggregate characterized by regular and frequent interaction by means of a shared body of verbal signs and set off from similar aggregates

by significant differences in language usage”. Thus, the current data comprise Mexican immigrants hailing from various regions in Mexico. These data represent one of the major Latinx speech communities in the Atlanta area, with Mexicans as the largest Latinx demographic and with first-generation immigrants who are Spanish-dominant (and who use mostly Spanish in daily work and home life) being the majority given the recently established nature of the Latin American community in Georgia. The speakers in the current data also share common linguistic attitudes toward Spanish and English, having reported their positive view of the use of both languages during the interviews. What delineates the Mexican speech community in Roswell from other communities is the use of Mexican Spanish, while other surrounding communities may use English exclusively, as well as other varieties of Spanish in the area (Colombian, Honduran, Venezuelan, etc.).

The following section will discuss some background on the study of variable SPE in Spanish in terms of pronoun rates and conditioning predictors involved. Section 3 will address the methodology utilized for the present study, and Section 4 presents the results of the analysis. Finally, Section 5 provides a discussion of the results and Section 6 concludes the article.

2. Variable SPE in Spanish

Previous research on variable SPE has examined both occurrence rates and constraints on variation between null and overt SPs. In terms of frequency, scholars have observed a wide range of overt SP rates across dialects of Spanish, ranging from relatively low rates in Mexico and Spain (~20%) to much higher rates in the Caribbean (~50%) (see Otheguy & Zentella 2012). SP rates have also been shown to vary within a country; for instance, with regard to Mexico, speakers in Mexico City, Yucatan, and Xalapa show rates of 21%, 16%, and 25%, respectively (Lastra & Martín Butragueño 2015; Michnowicz 2015; Orozco 2016).⁴ With regard to the predictors that influence SP variation, several morphosyntactic and semantic/pragmatic variables have been shown to exert a significant influence. Table 1 shows some examples of these variables, the particular contexts in which overt SPs are favored, and example studies of the findings.

Regional comparisons of subject pronoun expression

VARIABLE	OVERT SPS FAVORED	EXAMPLE STUDIES
Person/number	Singular verbs	Abreu 2012 (Florida/Puerto Rico) Bentivoglio 1987 (Caracas) Carvalho & Child 2011 (Rivera, Uruguay) Flores-Ferrán 2004 (NYC) Otheguy & Zentella 2012 (NYC)
Switch reference	Disjoint reference	Bayley & Pease-Alvarez 1997 (California) Cameron 1994, 1995 (San Juan/Madrid) Travis 2005 (Cali, Colombia)
Tense-Mood-Aspect (TMA)	Imperfect, conditional	Carvalho & Bessett 2015 (Rivera, Uruguay) Silva-Corvalán 1982 (Los Angeles) Travis 2007 (Cali, Colombia/New Mexico)
Morphological ambiguity	Ambiguous forms	Erker & Guy 2012 (NYC) Prada Pérez 2009 (Minorca) Lastra & Martín Butragueño 2015 (Mexico City) Michnowicz 2015 (Yucatan)
Verb class	Psychological verbs	Bentivoglio 1987 (Caracas) Silva-Corvalán 1994 (Los Angeles) Travis 2007 (Cali, Colombia/New Mexico)
Verbal mood	Indicative	Lastra & Martín Butragueño 2015 (Mexico City)
Specificity	<i>Mexico, Spain</i> : specific reference	Michnowicz 2015 (Yucatan), Cameron 1992 (Madrid)
	<i>Puerto Rico</i> : nonspecific reference	Cameron 1992 (San Juan)
Polarity	Affirmative	Geeslin & Gudmestad 2016 (U.S.) Lastra & Martín Butragueño 2015 (Mexico City)
Speech Style	Casual	Ávila-Jiménez 1996 (Puerto Rico) Lastra & Martín Butragueño 2015 (Mexico City)
Clause type	Main clauses	Orozco & Guy 2008 (Colombia) Otheguy & Zentella 2012 (NYC) Shin & Montes-Alcalá 2014 (NYC)
Reflexivity	Non-reflexive verbs	Carvalho & Child 2011 (Rivera, Uruguay) Otheguy <i>et al.</i> 2007 (NYC)
Priming	Previous overt SP	Cameron 1994 (San Juan/Madrid) Flores-Ferrán 2002 (NYC) Travis 2005 (Cali, Colombia)

Table 1. Overview of general findings for linguistic constraints on SPE.

Social predictors, such as gender and age, while less studied, have also been shown to impact variable SPE. The most consistent findings have been that women favor overt SPs (e.g. Bayley & Pease-Alvarez 1996; Solomon 1999; Carvalho & Child 2011; Otheguy & Zentella 2012; Shin & Otheguy 2013; Alfaraz 2015) and that younger speakers favor null SPs (e.g. Orozco & Guy 2008; Carvalho & Child 2011; Lastra & Martín Butragueño 2015; Limerick 2019). Furthermore, predictors such as length of residency, age of arrival, immigrant generation, and level of bilingualism influence SP use in contact varieties of Spanish (see, e.g., Otheguy & Zentella 2012; Carvalho *et al.* 2015). The current study will primarily focus on four of these predictors: TMA, switch reference, verb class, and polarity.⁵ The following sections will discuss their influence in greater detail according to the findings in previous research.

2.1. Switch-reference

Switch-reference, which considers continuity *vs* shift from one subject to another, shows a strong influence on SPE cross-dialectically (e.g. Bentivoglio 1987; Cameron 1994; Silva-Corvalán 1994; Bayley & Pease-Alvarez 1997; Travis 2005; Prada Pérez 2009; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015). Specifically, when there is a switch in the subject referent, the SP is more likely to be overt, as seen with *yo* in (1); when there is no switch, null SPs are preferred, as in (2).

- (1) *ahora ya la comunidad hispana pues hemos crecido mucho y este pues YO PIENSO que debemos...*
[F39]⁶
'now the Latino community well we have grown a lot and umm well I think that we should...'
- (2) *yo me relaciono mu- muy mucho con mi hermanito y y y s- ...(1.5) Ø PASO mucho tiempo allí*
[M27]
'I relate ve- very much with my little brother and and and s- ...(1.5) I spend a lot of time there'

This pattern is thought to have a functional influence that has to do with referential tracking (Shin & Otheguy 2009). As Cameron (1994: 40-41) explains, “expressed pronominal subjects compensate for the change of information state which occurs with a switch in subject reference.” In other words, overt SPs tend to be used in such contexts in order to facilitate interpretation of the antecedent.

Furthermore, some researchers include an intermediate category, ‘partial switch’, which analyzes cases of switch in subject where the subject is coreferent with the immediately preceding object, as in (3).

- (3) ... en México *me* inculcó mi mamá que *TENÍA* que ir a la iglesia... [M52]
'... in Mexico my mom instilled in me that I had to go to church...'

In this example, there is a switch in subject from *mi mamá* to *yo* (*tenía*), and *yo* is coreferential with the previous object *me*. In general, previous studies have reported either a slight favoring of overt SPs, or a neutral effect for such contexts, that is, neither a favoring nor a disfavoring effect (e.g. Orozco & Guy 2008; Otheguy & Zentella 2012; Lastra & Martín Butragueño 2015; Orozco 2015). Essentially, since there is still a switch in subject reference, promoting a higher use of overt SPs (relative to same reference contexts), but, at the same time, since the coreferential object was just mentioned and is salient in the discourse, overt SPs are less necessary for referential tracking and thus a weaker effect is found.

2.2. Tense-mood-aspect (TMA)

The tense-mood-aspect (TMA) of a verb is also an important predictor that conditions variable SPE. Certain TMAs favor overt SPs while others favor nulls. For instance, Silva-Corvalán (1982) found that imperfects and conditionals favor overt SPs while presents and preterits are more likely to appear with nulls, as was also found by other researchers (e.g. Cameron 1994; Travis 2007; Carvalho & Bessett 2015). To explain such correlations, it has been proposed that imperfects and conditionals favor 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, an explanation that forms part of the Functional Hypothesis (Hochberg 1986). Other studies, however, have found no such correlation (e.g. Enríquez 1984; Bentivoglio 1987; Ranson 1991).

An alternative explanation for the TMA effect, one that is not related to ambiguity, has been proposed by Silva-Corvalán (2001). She discusses imperfects and preterits in relation to discourse functions, suggesting 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). However, this hypothesis has been questioned in more recent work. For instance, Shin (2014) argues rather for the aforementioned ambiguity explanation, finding that imperfects favored overt SPs particularly in contexts of switch reference and with competing referents, both contexts in which referential tracking is more difficult. Given her findings, Shin argues that her analysis lends no support to Silva-Corvalán's (2001) proposal.

2.3. Verb class

Numerous researchers have found that verb class can also determine how an SP is manifested (e.g. Bentivoglio 1987; Silva-Corvalán 1994; Travis 2007; Otheguy & Zentella 2012; Orozco 2015). In general, it has been demonstrated that verbs of psychological/mental activity (e.g. *creer* ‘believe’, *pensar* ‘think’), verbs of communication (e.g. *decir* ‘say’, *hablar* ‘speak’), and copulas (e.g. *ser*, *estar* ‘be’) tend to be expressed with overt pronouns, with psychological verbs showing the highest probability. On the contrary, motion verbs tend to disfavor overts (Bentivoglio 1987; Silva-Corvalán 1994; 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 1994). The speaker “asserts their role in the utterance” by using an overt SP (Travis 2007: 117). With regard to the preference for overt SPs with communication verbs, particularly *decir*, Travis (2007: 117) has postulated that this is perhaps related to the epistemic function of *decir* to express an opinion (e.g. *yo digo que* ‘I say that’), similar to the aforementioned effect for psychological verbs. 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 favored overt SPs while activity verbs were more likely to appear with nulls.

More recently, however, the above categorizations have been called into question. It has been argued that the analysis of verb class effects using these traditional lexical categories exhibits deficiencies as they oversimplify the influence of verb class on SPE (e.g. Posio 2011; Orozco 2018, Orozco & Hurtado 2021). In particular, Orozco (2018) has observed opposing tendencies for verbs within a single category and, therefore, promote a more nuanced method of operationalizing verb class effects (see also Posio 2011). Two new ways of analyzing verb class, particularly in the variationist tradition, are proposed by Orozco & Hurtado (2021) and include (a) verb transitivity and (b) lexical effects as examined by infinitive verb forms as well as pronominal subject + verb collocations. What has been found overall as a result of these methods is the following: Unergative verbs favored overt SPs while reflexive verbs favored null SPs; there were opposing tendencies for verbs within the same semantic class (e.g. *creer* ‘believe/think’ favored overt subjects while *imaginarse* ‘imagine’ favored null subjects) and within different collocations of the same verb (e.g. *es* promoted overt subjects while *son* promoted null subjects) (Orozco & Hurtado 2021: 19). Thus, the stud-

ies cited above clearly demonstrate the need to better understand verb effects on SPE and that we are left with an oversimplified picture if we employ the traditional methods.

2.4. Polarity

Polarity considers affirmative vs non-affirmative clauses and their conditioning on SP variation. While still relatively understudied, this variable has occasionally been analyzed in previous research, with some researchers finding a significant effect (Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016; Limerick 2019) and others reporting a lack of effect (Travis & Torres Cacoullos 2012; Torres Cacoullos & Travis 2015).⁷ For example, Lastra & Martín Butragueño (2015) found that non-affirmatives (negatives and interrogatives) disfavored 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. They hypothesize that “if negated clauses cluster together, it is possible that co-reference across these negated clauses contributes to their disfavoring effect on overt SPPs” (Lastra & Martín Butragueño 2015: 46). In other words, there could be an interaction between the switch reference and polarity predictors such that negative clauses disfavor overt SPs more frequently in coreferential contexts than in switch reference contexts.

Likewise, Geeslin & Gudmestad (2016) found that negation disfavored overt SPs while affirmative clauses favored them. This finding confirmed their prediction based on previous research that the presence of pre-verbal elements makes null SPs more likely. Interestingly, this variable was only significant for 1SG forms and not for second-person singular forms in Geeslin & Gudmestad’s analysis, which calls for further research on the interaction of negation and person.

Furthermore, Travis & Torres Cacoullos (2012), while not finding a significant effect for polarity overall, did find effects upon considering particular verb classes/lexemes. Their study revealed interesting patterns in the distinction between cognitive and non-cognitive verbs and their influence on polarity, namely that negation favored overt SPs, but only for non-cognitive verbs. Moreover, although a significant effect for cognitive verbs was not found overall, the researchers did find an effect when considering only the particular verbal lexeme *creo* ‘(I) think/believe’, specifically that negation highly disfavored overt SPs (e.g. \emptyset NO *creo*) (Travis & Torres Cacoullos 2012: 741, note 23).

3. Methods

3.1. Data collection

In 2015, sociolinguistic interviews were conducted with Spanish-speakers who were either living or working in Roswell, Georgia, a suburb of Atlanta, at the time of data collection (see Limerick 2018 for further details). The sample for the present analysis consists of 20 first-generation Mexican immigrants. These interviews form part of the *Corpus of Spanish in Georgia* (Limerick 2022). The interviews lasted between 30 minutes and one hour. They were informal, conversational, and addressed topics of personal history, local community life, differences between the speakers' home countries and the U.S., and experiences adapting to life in the U.S., among others.

3.2. The speakers

The speakers' sociodemographic backgrounds can be summarized in the following way (see also Tables 2 and 3 below): The speakers were born in various regions of Mexico (see Figure 2 below): Mexico City (8), Acapulco, Guerrero (2), the state of Guerrero (1),⁸ 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 ages of arrival (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.

Regional comparisons of subject pronoun expression



Figure 2. Map of Mexico.

SPEAKER	LOR	AOA	EDUCATION	ENGLISH PROFICIENCY
F39Mex	14	25	Partial law school	2
F34Mex	10	24	University	4
F52Mex	2	50	High school	1
F30Mex	10	20	High school + Cosmetology school	3
M41Mex	13	28	Partial university	3
M33Mex	12	21	Partial University	3
M27Mex	16	11	High school (U.S.)	4
M52Mex	15	37	Partial secondary school (2 years)	2

Table 2. Mexico City Speakers.

SPEAKER	CITY/STATE OF ORIGIN	LOR	AOA	EDUCATION	ENGLISH PROFICIENCY
M51Mex	Cuernavaca, Morelos	10	41	Primary school	1
M34Mex	Morelos (state)	10	24	University	3
M32Mex	Tampico, Tamaulipas	16	16	Secondary school	3
M43Mex	San Juan del Río, Querétaro	25	18	Partial high school	2
F49Mex	Juando, Mexico	7	42	Primary school	3
F56Mex	Acapulco, Guerrero	25	31	Secondary school	4
F32Mex	Acapulco, Guerrero	7	25	University	2
F28Mex	Guerrero (state)	2	26	Master's	3
F26Mex	Monterrey, Nuevo León	12	14	High school (U.S.)	4
F43Mex	San Luis Potosí (state)	24	19	High school + Cosmetology school	3
F60Mex	Colima (state)	4	56	High school	1
F25Mex	Zacatecas (state)	13	12	High school + Cosmetology school	4

Table 3. Non-Mexico City speakers.

3.3. The variable context for subject expression

In order to analyze SPE, the audio data were first transcribed. I then extracted all 1SG finite verbs from the interviews 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 tokens that fell outside the variable context. For 1SG, this included set phrases where an overt or null SP was categorical (e.g. *¿Qué sé yo?* 'I don't know'). Speakers did not alternate between an overt and null SP in these cases; thus, these structures were excluded. For all tokens within the variable context (N = 2,565), I first coded whether each verb

appeared with a null or overt SP. I then coded for the four independent variables using the categorizations discussed below.

3.4. Research questions

Based on previous studies of SPE, the present investigation aims to answer the following research questions:

RQ1: What is the distribution of overt and null first-person singular SPs for immigrants from Mexico City as compared to other Mexican regions (e.g. Guerrero, Zacatecas, Morelos)?

RQ2: For the linguistic predictors influencing such variation, how do variable and constraint hierarchies compare for Mexico City vs other regions?

RQ3: What role do language-external (social) predictors play in the SP variation of either of the two speaker groups?

RQ1 is guided by the immense regional variation in SP rates observed both between and within countries in previous research on SPE (e.g. Travis 2007; Torres Cacoullos & Travis 2010; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2018). RQ2 is informed by the method of ‘comparative sociolinguistics’ (Tagliamonte 2002; Meyerhoff 2009), which examines the similarities and differences in variable grammars between two or more speech communities or language varieties. Finally, RQ3 is based on previous SPE scholarship in other speech communities that demonstrates the influence of social predictors on SPE, such as age (e.g. Orozco & Guy 2008; Carvalho & Child 2011; Lastra & Martín Butragueño 2015; Orozco 2015; Limerick 2019) and gender (e.g. Bayley & Pease-Alvarez 1996; Solomon 1999; Carvalho & Child 2011; Otheguy & Zentella 2012; Shin & Otheguy 2013; Alfaraz 2015; Orozco 2018).

3.5. Linguistic predictors

The four internal predictors for the current analysis are the following: Switch-reference, TMA, Verb class, and Polarity.⁹ The inclusion of these predictors is based on their significance in previous SPE research, and the coding methods are primarily based on Otheguy & Zentella’s (2012) analysis of NYC Spanish as well as Lastra & Martín Butragueño’s (2015) study of Mexico City Spanish (see Table 4).¹⁰

VARIABLE	CATEGORIES
Switch-reference	Same-reference Switch-reference Partial-switch (subjects that are coreferential with preceding objects)
TMA	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.
Verb class ¹¹	Mental processes: (e.g. <i>creer</i> 'believe', <i>pensar</i> 'think', <i>saber</i> 'know') Stative verbs: (e.g. <i>ser</i> 'be', <i>estar</i> 'be', <i>tener</i> 'have') <i>Verba dicendi</i> (i.e. communication, e.g. <i>decir</i> 'say', <i>hablar</i> 'speak') Activity verbs (e.g. <i>jugar</i> 'play', <i>hacer</i> 'do/make', <i>ir</i> 'go')
Polarity ¹²	Affirmative Non-affirmative (both negative and interrogative)

Table 4. Linguistic variables and categories.

Based on the discussion of findings in previous scholarship above in Section 2, the following hypotheses are set forth regarding the linguistic predictors:

- a. Switch reference: overt SPs will be favored when there is a switch in subject referent and disfavored in contexts of reference continuity (same reference) (e.g. Travis 2005; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015).
- b. TMA: overt SPs will be favored with imperfect and conditional forms while being disfavored with other forms (e.g. Silva Corvalán 1982; Cameron 1994; Travis 2007; Carvalho & Bessett 2015; Orozco 2015, 2016; Shin & Van Buren 2016; Lastra & Martín Butragueño 2015).
- c. Verb class: Verbs denoting mental processes as well as stative verbs will favor overt SPs while verbs in other categories will promote null SPs (e.g. Silva-Corvalán 1994; Travis 2007; Torres Cacoullos & Travis 2010, 2011; Otheguy & Zentella 2012; Orozco 2015).
- d. Polarity: affirmative contexts will favor overt SPs while non-affirmative contexts disfavor them (e.g. Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016; Limerick 2019).

3.6. Social predictors

Due to their significance in previous studies of subject expression, as discussed in Section 2, the social variables of age and gender were also included in the present study. Age was treated as a continuous variable (ranging from 25 to 60), and gender was coded as either female or male. It is predicted, based on prior findings, that younger speakers will disfavor overt SPs (e.g. Orozco & Guy 2008; Carvalho & Child 2011; Lastra & Martín Butragueño 2015; Orozco 2015; Limerick 2019) and that women will favor overt SPs (e.g. Bayley & Pease-Alvarez 1996; Solomon 1999; Carvalho & Child 2011; Otheguy & Zentella 2012; Shin & Otheguy 2013; Alfaraz 2015; Orozco 2018).

3.7. Statistical Methods

To determine the statistical significance of the linguistic and social predictors and the relative weight of each predictor regarding its conditioning on SPE for Mexico City speakers and non-Mexico City speakers, separate mixed-effects multivariate analyses (logistic regressions) were carried out using *Rbrul* (Johnson 2009) with the inclusion of the speaker as a random effect. Incorporating the speaker as a random effect in the statistical models is a way of controlling for the individual speaker and ensuring that the results obtained are generalizable to the speakers as a whole, and that the patterns are not due to particular speakers skewing the results (Bayley *et al.* 2013; Shin 2014; Michnowicz 2015; Shin & Van Buren 2016). Furthermore, the analysis of lexical effects of the verb was employed by incorporating the 20 most frequent verbal infinitives in the data set as random effects (Orozco 2016; Orozco & Hurtado 2021). Additionally, potential interaction effects were tested for both among the social predictors and between linguistic and social predictors using *Rbrul* and *Language Variation Suite* (Scrivner & Díaz-Campos 2016).

The following section presents the results of the current investigation. It is important first, however, to point out some methodological modifications made to the dataset after the initial coding took place. Due to low token counts for some of the categories, these were collapsed. Specifically, the category *partial switch* was moved to the *switch* category, and the six TMA categories of *perfect*, *present subjunctive*, *past subjunctive*, *synthetic future*, *periphrastic future*, and *conditional* were moved to a separate category named *All other TMAs* (see Lastra & Martín Butragueño 2015 for similar methods). For the multivariate analysis, then, Switch Reference comprised *same* and *switch*, and TMA included four levels: *Present*, *Preterit*, *Imperfect*, and *All Other TMAs*.

4. Results

4.1. Pronominal rates

The analysis of overall overt pronominal rates showed that there is a notable difference between those hailing from Mexico City and those from other Mexican regions. The former group had an average rate of 32% *yo* while the latter group produced *yo* at a rate of 39%.¹³ This distinction is unsurprising given the relatively wide regional variation in SP rates observed in previous studies of Mexican Spanish (Mexico City: 25% [Lastra & Martín Butragueño 2015]; Yucatan: 21% for Spanish monolinguals & 35% for Maya-Spanish bilinguals [Michnowicz 2015]; Xalapa: 33% [Orozco 2016]).

As another point of reference, it is interesting to compare the pronominal rates observed here with other speech communities outside of Mexico (see Table 5). For instance, in Medellín, Colombia and Rivera, Uruguay, we see similar rates to those observed in the current data. On the other hand, speech communities such as Huancayo, Peru and Santo Domingo show drastic differences. These comparisons further evidence the wide rate of variation in SPE frequencies across the Spanish-speaking world.

SPEECH COMMUNITY	% 1SG OVERT SPE
Santo Domingo, Dominican Republic (Alfaraz 2015)	52
Barranquilla, Colombia (Orozco 2015)	45
Medellín, Colombia (Orozco & Hurtado 2021)	32
Rivera, Uruguay (Carvalho & Bessett 2015)	32
Santiago de Chile (Martínez-Lara <i>et al.</i> 2021)	29
Malabo, Equatorial Guinea (Padilla 2021)	25
Valladolid, Spain (Prada Pérez 2015)	20
Huancayo, Peru (Cerrón-Palomino 2018)	19

Table 5. First-person singular overt pronoun rates in speech communities outside Mexico.

It is also interesting to consider individual speaker variation in terms of overt frequency (see Tables 6 and 7). While the average rate among multiple speakers is certainly useful for comparison, this average may hide a vast amount of individual variation, such as what we see below.

Regional comparisons of subject pronoun expression

SPEAKER	OVERT SP RATE	N TOKENS	% DATA
F39	46%	47/103	10%
M27	39%	41/106	10%
F52	35%	67/193	18%
M41	35%	43/124	12%
M33	29%	23/80	8%
F30	27%	21/79	8%
M52	26%	38/146	14%
F34	25%	55/220	21%

Table 6. Individual pronoun rates for Mexico City speakers.

SPEAKER	OVERT SP RATE	REGION	N TOKENS	% DATA
F56	57%	Acapulco, Guerrero	105/184	12%
M51	55%	Cuernavaca, Morelos	17/31	2%
F32	50%	Acapulco, Guerrero	75/151	10%
F43	49%	San Luis Potosí (state)	49/99	7%
F60	47%	Colima (state)	54/116	8%
F26	42%	Monterrey, Nuevo León	60/144	10%
F49	39%	Juando, Mexico	30/78	5%
M32	36%	Tampico, Tamaulipas	53/149	10%
M34	35%	Morelos (state)	41/116	8%
F28	32%	Guerrero (state)	51/160	11%
F25	23%	Zacatecas (state)	37/163	11%
M43	12%	San Juan del Río, Querétaro	15/123	8%

Table 7. Individual pronoun rates for non-Mexico City speakers.

While the overall rate for Mexico City speakers was 32%, notice that half of the speakers show rates far away from this average (e.g. 25%, 26% on the low end; 39%, 46% on the high end). The range in overt frequency is even more dramatic for non-MC speakers, whose average rate was 39%. On the lower end, the data show rates of 12% and 23%, and, interestingly, five of the highest pronoun users show

rates ranging from the high 40s to high 50s, 10%-18% above the average rate (39%). The most striking findings are the unusually low rate of 12% overt SPE and the unusually high rates of 50-57% overt SPE, the latter rates typically associated with Caribbean varieties of Spanish and uncharacteristic of Mexican varieties. This wide range of rates demonstrates the vast amount of variation that can occur within a single country, in this case within Mexico. While this is by no means representative of each of these Mexican regions, the current data raise questions as to how the larger communities in these regions use SPs. Thus, further research is warranted that explores more representative samples of pronoun use in Guerrero, Morelos, Zacatecas, Colima, etc.

4.2. Linguistic predictors

Despite such variation in pronominal rates, when we look at the linguistic predictors governing SPE across Mexico City and non-Mexico City speakers, they are remarkably similar (see Table 8). First, the same four predictors have a statistically significant influence on SPE in each of the two groups: Switch reference, TMA, verb class, and polarity. Moreover, the variable hierarchy (determined by the range) is the same, with the strongest predictor being switch reference and the least powerful predictor being polarity. These data are consistent with previous research demonstrating that while pronominal rates vary widely across Spanish dialects, the conditioning linguistic predictors generally remain stable (e.g. Cameron 1992; Torres Cacoulllos & Travis 2010; Otheguy & Zentella 2012; Carvalho & Bessett 2015; Orozco 2015; Shin & Erker 2015; Martínez-Lara *et al.* 2021).

MEXICO CITY	RANGE	NON-MEXICO CITY	RANGE
Switch reference	22	Switch reference	28
TMA	19	TMA	19
Verb class	18	Verb class	18
Polarity	15	Polarity	8

Table 8. Comparison of variable hierarchies.

Moreover, upon examining the direction of effect within each predictor, we again see striking similarities for the two groups. Tables 9 and 10 below present the constraint hierarchies for each of the four predictors and compare Mexico City to non-Mexico City. The first column shows each predictor 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 favoring of overt SPs. When it is closer to 0, it generally indicates a disfavoring of overt SPs (see Tagliamonte 2006: 145, 156). The two groups exhibit similarities the most for switch reference and polarity, where we see the same direction of effect and very similar FWs. There is slight variation with the predictor of TMA, whereby Mexico City speakers prefer null SPs with the preterit (fw = .43) while non-Mexico City speakers show a neutral effect (fw = .49). There are also some minor differences for verb class, in which Mexico City speakers show a preference for nulls with stative verbs (fw = .43) while non-Mexico City speakers prefer nulls with activity verbs (fw = .41).¹⁴

FACTOR GROUP	FACTOR WEIGHT	% OVERT	N TOKENS
Switch reference			
switch	.61	43%	206/479
same	.39	23%	131/572
RANGE	22		
TMA			
Imperfect	.62	42%	57/135
All Other TMAs	.51	34%	27/80
present	.44	32%	185/578
preterit	.43	27%	70/258
RANGE	19		
Verb Class			
Mental	.61	43%	136/317
Communicative/Activity	.46	28%	123/440
Stative	.43	26%	76/294
RANGE	18		
Polarity			
Affirmative	.58	33%	303/917
Negative	.43	23%	31/134
RANGE	15		

p-values: switch reference (2.52e-10); TMA (0.00267); Verb class (6.92e-05); Polarity (0.00675)

Table 9. Constraint hierarchies: Mexico City (n = 1051).

FACTOR GROUP	FACTOR WEIGHT	% OVERT	N TOKENS
Switch reference			
switch	.64	54%	356/659
same	.36	27%	231/855
RANGE	28		
TMA			
Imperfect	.62	50%	114/228
Preterit	.49	37%	128/345
All Other TMAs	.46	37%	62/168
Present	.43	37%	286/773
RANGE	19		
Verb Class			
Mental	.59	48%	184/384
Communicative/stative	.51	39%	208/533
Activity	.41	32%	191/597
RANGE	18		
Polarity			
Affirmative	.54	40%	525/1313
Negative	.46	34%	68/201
RANGE	8		

p-values: switch reference (3.22e-24); TMA (0.000125); Verb class (1.51e-05); Polarity (0.0423)

Table 10. Constraint hierarchies: non-Mexico City (n = 1514).

Overall, the constraint hierarchies show patterns that are consistent with previous research for the linguistic predictors under study, namely that there is a favoring of overt SPs in the following contexts:

- switch in subject referent (Travis 2005; Torres Cacoullós & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015)
- imperfect aspect (Silva Corvalán 1982; Cameron 1994; Travis 2007; Carvalho & Bessett 2015; Orozco 2015, 2016; Shin & Van Buren 2016; Lastra & Martín Butragueño 2015)

- mental verbs (Silva-Corvalán 1994; Travis 2007; Torres Cacoullós & Travis 2010, 2011; Otheguy & Zentella 2012; Orozco 2015)
- affirmative contexts (Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016)

4.2.1. Lexical effect of the verb

Additionally, to further explore the effect of the verb on SP variation beyond the traditional verb class predictor, lexical effects were analyzed using the most frequent verbs in SPE contexts in the data set by analyzing the infinitive of each verb as a random effect (see Orozco 2016; Orozco & Hurtado 2021).¹⁵ Following Orozco & Hurtado (2021), verbs that comprised at least 0.5% of the dataset (for the current data, those occurring at least 13 times) were included. The results are presented below in Table 11, which includes the 20 most frequent verbs. The tendencies of the verbs regarding their effects on SPE are the following: *creer* ‘believe’ and *pensar* ‘think’ are the verbs that most strongly favor overt SPE, with FWs of .73 and .70, respectively. Other verbs also show a favoring effect, albeit to a lesser degree, such as *ser* ‘be’, *saber* ‘know’, *llegar* ‘arrive’, and *decir* ‘say’, among others. Moreover, a disfavoring of overt SPE is observed with verbs such as *tener* ‘have’, *poner* ‘put’, *venir* ‘come’, *ir* ‘go’, and *conocer* ‘know’, among others, while other verbs exhibit a neutral effect, such as *querer* ‘want’ and *poder* ‘can’.

VERB	FACTOR WEIGHT	N	% OVERT	% DATA
<i>Creer</i> ‘believe’	.73	173	67%	7%
<i>Pensar</i> ‘think’	.70	48	69%	3%
<i>Ser</i> ‘be’	.56	93	47%	4%
<i>Saber</i> ‘know’	.56	52	48%	2%
<i>Llegar</i> ‘arrive’	.56	86	47%	3%
<i>Decir</i> ‘say’	.55	60	47%	2%
<i>Ver</i> ‘see’	.54	66	46%	3%
<i>Vivir</i> ‘live’	.54	56	45%	2%
<i>Querer</i> ‘want’	.51	58	41%	2%
<i>Poder</i> ‘can’	.51	94	40%	4%
<i>Hacer</i> ‘do/make’	.48	30	37%	1%
<i>Estar</i> ‘be’	.46	145	36%	6%

<i>Conocer</i> ‘know’	.45	63	33%	2%
<i>Ir</i> ‘go’	.43	139	32%	5%
<i>Haber</i> ‘have’	.43	167	32%	7%
<i>Trabajar</i> ‘work’	.42	82	31%	3%
<i>Empezar</i> ‘begin’	.42	30	27%	1%
<i>Venir</i> ‘come’	.40	53	26%	2%
<i>Poner</i> ‘put’	.39	14	14%	0.5%
<i>Tener</i> ‘have’	.36	283	26%	11%

Table 11. Lexical effects of the verb (n = 1792).

These findings show partial consistency to those of other speech communities. Regarding Mexican Spanish, Orozco (2016: 8) also found a favoring of overt SPE with *creer* and *ser* in Xalapa. However, the strongest favoring effect in Orozco’s study was observed for *ir* (fw = .67), a verb for which the present study shows a disfavoring effect (fw = .43). Another similarity, for instance, is that both *estar* and *tener* disfavor overt SPs in the two Mexican varieties. To take another speech community, Medellín, Colombia, we also see some comparable patterns. Orozco & Hurtado (2021: 17) found that the verbs *creer* (the strongest effect), *pensar*, *decir*, and *vivir*, among others, favored overt SPE in Medellín, all findings that are consistent with the current data. Similarly, both *poner* and *venir* promote null subjects in both communities. There are, nevertheless, some differences as well, such as the fact that *trabajar* favors overt subjects in Medellín (fw = .723) while favoring null subjects (fw = .42) for Mexicans in Roswell, Georgia in the current study.¹⁶

The fact that *creer* is in the lead for favoring overt SPs suggests that this verb in particular (as well as *pensar*) is driving the overall verb class effect for mental verbs discussed above and found in numerous other investigations (e.g. Silva-Corvalán 1994; Travis 2007; Torres Cacoullós & Travis 2010, 2011; Travis & Torres Cacoullós 2012; Otheguy & Zentella 2012; Orozco 2015).¹⁷ In fact, *conocer*, another mental verb, disfavors overt subjects (fw = .45) and *saber* only slightly favors them (fw = .56) in the current data. These results also lend support to the findings of more recent research that there are opposing tendencies within the same verb class category (e.g. Posio 2011; Orozco 2016, 2018; Orozco & Hurtado 2021). Returning to *creer*, all but 2 cases in the dataset are in the present indicative ([yo] *creo*), with overt *yo creo* constituting 66% of all cases.¹⁸ Both the high relative frequency of *creer* (7%

of the data) and the high overt frequency suggests that the construction *yo + creo* may constitute a prefab or formulaic sequence (Posio 2011, 2015), as has been argued in previous analyses of other speech communities (e.g. Travis & Torres Cacoullos 2012; Posio 2015; Orozco & Hurtado 2021).

4.3. Social predictors

Neither age nor gender exhibited a significant main effect on SPE among either speaker group. However, upon considering the interaction of these social predictors with the linguistic predictors analyzed in the study, some unexpected results emerged. In particular, statistically significant interaction effects were found for *gender:polarity*, *age:verb class*, and *gender:TMA*. Tables 12-13 and Figure 3 below show the cross tabulations and a conditional inference tree for these predictors, demonstrating the particular interaction effects they have on SPE. Interestingly, this exploration of interactions in the current data revealed differential effects for Mexico City vs non-Mexico City speakers. For the former group, gender interacted with polarity (see Table 12) and age interacted with verb class (see Figure 3). For the latter group, however, neither of these interactions were operative. Instead, gender interacted with TMA for non-Mexico City speakers (see Table 13). These differences between the two groups reflect regional differences regarding the intersection of linguistic and social predictors governing SPE for Mexican speakers.

Table 12 below shows that the polarity effect is dependent upon the gender of the speaker in that it is only operative for women (34% affirmative vs 15% negative). The distinction in overt SPE disappears for males, with the same proportion of overt SPE in both affirmative and negative contexts (32%).

GENDER	AFFIRMATIVE	NEGATIVE	TOTAL
Male	32%	32%	32%
Female	34%	15%	32%
Total	33%	23%	32%

p = 6.84e-03

Table 12. Interaction between gender and polarity for SPE (% overt) (Mexico City; n = 1051).

Also for Mexico City speakers, as stated above, age intersects with verb class. The conditional inference tree in Figure 3 demonstrates the specifics of this interaction, namely that the verb class effect (favoring

of overt SPE with mental verbs) is amplified among middle-aged and younger speakers (ages 41 and below). As shown in the tree, speakers 41 and younger exhibit a significantly higher frequency of overt SPs (48%) with mental verbs than speakers older than 41 (28%).

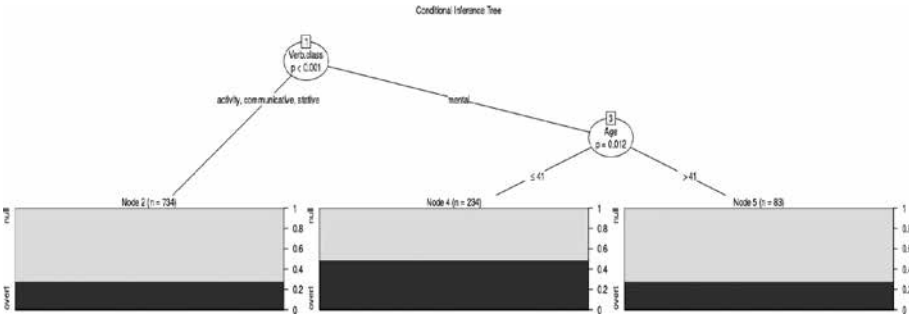


Figure 3. Interaction between age and verb class for SPE (Mexico City; n = 1051).

Moving to the non-Mexico City group, a different type of interaction was observed. Specifically, gender interacted with the linguistic predictor of TMA (see Table 13).

GENDER	IMPERFECT	PRETERIT	PRESENT	ALL OTHER TMAs	TOTAL
Male	47%	19%	29%	38%	30%
Female	51%	45%	39%	36%	42%
Total	50%	37%	37%	37%	39%

$p = 0.0213$

Table 13. Interaction between gender and TMA for SPE (% overt) (non-Mexico City; n = 1514).

In particular, women only marginally distinguished between the imperfect and preterit in their overt SP usage (51% vs 45%) whereas men showed a strong distinction (47% vs 19%). The sharp increase of overt SPE in preterit contexts for women as compared to men is what neutralizes or weakens their imperfect/preterit distinction. Women continued to use a high percentage of overt pronouns with the imperfect (45%) but showed a low proportion of pronouns with the preterit (19%). The following sections of the paper will further discuss the above findings, summarize the study, and provide suggestions for future work on variable SPE.

5. Discussion

5.1. Pronoun rates and internal predictors

Regarding the overall frequencies of overt SPE, it was found that Mexico City speakers produced overt *yo* 32% of the time and non-Mexico City speakers showed an overt rate of 39%, a notable difference according to region. Aside from comparisons between the two groups, a wide amount of variation in rates was observed within each group, especially the non-Mexico City group, with rates ranging from 12% to 57% overt SPE. A particularly interesting finding were the rates on the higher end in this group, uncharacteristic for Mexican Spanish as documented in the literature (e.g. Lastra & Martín Butragueño 2015; Michnowicz 2015; Orozco 2016). This finding shows that Mexican Spanish may exhibit a wider range of variation in SP rates than was previously assumed.

The internal predictors played a very similar role between speaker groups, suggesting that, despite hailing from different regions in Mexico, the speakers' variable grammars are more or less the same. In particular, the production of *yo* was favored with a switch in subject referent, a widely repeated finding in the variationist literature (e.g. Travis 2005; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015). The idea is that overt SPs are produced less for purposes of referential tracking in same reference contexts since the subject was just mentioned and are produced more for reference tracking when both the previous subject and object were noncoreferential (Shin & Otheguy 2009). Such relation between pronoun use and referential tracking is also associated with the notions of 'accessibility' and 'salience' (Givón 1983; Ariel 1994). In this respect, switch reference contexts make referents less accessible/salient while in cases of same reference the referent, just being mentioned in the previous clause, is more accessible. Thus, the former context promotes overt SPs while the latter promotes nulls.

With respect to TMA, overt SPs were favored with the imperfect, another general finding in previous scholarship (Silva Corvalán 1982; Cameron 1994; Travis 2007; Carvalho & Bessett 2015; Orozco 2015, 2016; Shin & Van Buren 2016; Lastra & Martín Butragueño 2015; but see Shin 2014 and Limerick 2019 for exceptions). This could be due to the morphological ambiguity of the first-person singular in the imperfect; however, a qualitative analysis of the discourse context is needed to confirm whether referential ambiguity plays a significant role.

The next internal predictor, verb class, exhibited a common pattern for SP variation, namely that overt SPs were favored with mental verbs (see Silva-Corvalán 1994; Travis 2007; Torres Cacoullos & Travis 2010, 2011; Otheguy & Zentella 2012; Orozco 2015). This is perhaps attributable to the implied contrast expressed by the speaker and the fact that speakers convey their point of view with these types of verbs (Silva-Corvalán 1994).

The analysis of verb class was extended to include the lexical effects of the 20 most frequent verbs in the dataset, with both similarities and differences observed between Mexican speakers in Roswell with speakers of other speech communities. Overall, *creer* ‘believe/think’ and *pensar* ‘think’ mostly strongly favored overt SPs while *tener* ‘have’ and *poner* ‘put’ highly promoted null subjects. The present findings also corroborate previous investigations that found opposing tendencies for different verbs in the same semantic category (e.g. Posio 2011; Orozco 2016, 2018; Orozco & Hurtado 2021), further supporting new methodologies in studying verb effects on SPE rather than utilizing the traditional categories based on mere semantic classifications. For instance, the current data demonstrated that, within the mental verbs category, *creer* favored overt SPs while *conocer* ‘know’ promoted null SPs.

The differences observed regarding which verbs favor or disfavor overt SPs when compared to other speech communities, such as Xalapa, Mexico (Orozco 2016) and Medellín, Colombia (Orozco & Hurtado 2021) also lend support to the idea that lexical effects tend to be idiosyncratic to each variety, that is, tendencies for certain verbs to favor or disfavor subject expression (with the exception of *creer* and *pensar*) are not consistent across different speech communities (Orozco & Hurtado 2021). As only the verbal infinitives were analyzed in the current investigation, future work on this speech community needs to extend the lexical effects analysis to include tendencies among different collocations of pronominal subjects + verbal inflections (see Orozco & Hurtado 2021) to further our knowledge of verb lexeme effects in this variety.

Finally, regarding polarity, affirmative contexts favored the use of *yo*, as shown in other studies (Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016). According to Geeslin & Gudmestad (2016), the presence of pre-verbal elements (e.g. negation) make null SPs more likely.

5.2. External predictors

While neither of the social predictors explored (age and gender) showed significant main effects on SPE in the current data, upon considering their intersection with linguistic predictors, some interesting interac-

tion effects emerged. First, there was a significant interaction between gender and polarity in that the main effect for polarity (overt SPs favored with affirmative polarity) applies only to women; the male speakers in the data did not show a polarity effect, with the exact same rate of overt SPE in both affirmative and negative contexts (32%). It is unclear why women would exhibit this effect, and why men would not; nevertheless, it is worth considering potential implications of this effect in future research. The second significant interaction effect found was that between age and verb class. In particular, there was a divide between younger and middle-aged speakers (41 and younger) on the one hand, and older speakers (older than 41) on the other, in that the overall verb class effect (overt SPE favored with mental verbs) was heightened for the former group. Why would older speakers no longer demonstrate a distinction in overt SP rate for different verb classes in their variable grammar? Put another way, why would younger and middle-aged speakers show an increased verb class effect? It could be that older speakers do not feel the pragmatic need to assert their role or emphasize their responsibility regarding their utterance with the use of *yo* in assertions such as *YO creo* 'I believe' / *YO pienso* 'I think' (Aijón Oliva & Serrano 2010) while younger speakers are more cognizant of such mitigation. Further analysis of the discourse context would be worthwhile to explore these possibilities. The third and final interaction effect observed in the present analysis was that between gender and TMA. Men showed a significant distinction between the imperfect (47% overt) and preterit (19% overt), mirroring what has been found in the variationist scholarship in general for SPE, that is, a high use of overt SPE with the imperfect versus a low use with the preterit (Silva Corvalán 1982; Cameron 1994; Travis 2007; Carvalho & Bessett 2015; Orozco 2015, 2016; Shin & Van Buren 2016; Lastra & Martín Butragueño 2015). Women, however, did not exhibit a significant distinction in these contexts, with 51% overt SPE in imperfect contexts and 45% with the preterit, a difference of only 6%.

These interaction effects reveal new tendencies that have been overlooked in previous research on SPE in Spanish. In fact, examining interactions between linguistic and social predictors is not commonly practiced in variationist sociolinguistics in general (but see, e.g., Poplack 1997 and Villarreal *et al.* 2021 for important exceptions). We see here, for example, that the effects for external predictors (e.g. age, gender) are more complex when interactions with internal predictors are considered. The current data raise new questions and broaden research avenues regarding the sociolinguistic conditioning of variable phenomena in general, and SPE in particular. Future studies are warranted to corroborate these findings for interaction effects on SPE.

Another important finding is that the above interaction effects were region specific. Specifically, the gender:polarity and age:verb class effects were observed among Mexico City speakers while the gender:TMA effect was shown in the non-Mexico City group. This finding suggests that while the variable grammars of speakers remain very similar cross-dialectally, the differing sociolinguistic background and contexts of different regions may promote differential interaction effects between the linguistic and social, in particular with regard to variable SP use.

6. Conclusion

This paper reported on a comparative analysis of variable SPE among first-generation Mexican immigrants in the U.S. with differing regional origins. Using sociolinguistic interview data collected in Georgia, occurrence rates and usage patterns of first-person singular SPs were examined. The primary comparison was made between speakers hailing from Mexico City and speakers from other regions in Mexico (Guerrero, Zacatecas, Morelos, etc.) in addition to the analysis of variation among individual speakers. Regarding overall overt pronoun rates for *yo*, Mexico City speakers produced fewer (32%) than non-Mexico City speakers (39%). A wide range of variation in rates was also seen within each of the two groups, with non-Mexico City speakers showing the largest amount of variation from speaker to speaker as well as displaying unusually high SP rates not typically reported for Mexican Spanish.

Regarding the linguistic predictors conditioning pronoun use, it was found that all four predictors (switch reference, TMA, verb class, polarity) were significant for both groups. Remarkably similar usage patterns were also observed, with Switch reference being the strongest predictor and polarity being the least powerful influence on variable SPE. Further, the constraint hierarchies were strikingly similar. These findings reflect the generalization that while pronoun frequencies vary greatly across varieties, the constraints on pronoun variation remain consistent (Carvalho *et al.* 2015). The additional lexical analysis of verb effects also revealed a more nuanced view of how verbs condition SP variation by showing which particular verbs promoted overt subjects and which did not. Crucially, divergent tendencies were observed for different verbs within the same semantic category, offering additional support to recent scholarship that points out the deficiencies and oversimplifications involved with the more traditional methods (Orozco 2016; Orozco & Hurtado 2021). The more innovative methods provide us with a more complex picture of not only the influence of specific

verbs, but also the unique nature of verb effects between varieties of Spanish as well as the existence of grammaticalized constructions or formulaic sequences of subject pronoun + verbal lexeme across varieties (Travis & Torres Cacoullous 2012; Posio 2015; Orozco & Hurtado 2021).

For the social predictors under study (age and gender), neither was significant as a main or direct influence on SP variation. They did exhibit, however, significant interaction effects, with gender interacting with polarity and age interacting with verb class among speakers from Mexico City, in addition to gender interacting with TMA among the group from other regions in Mexico. Thus, we are seeing non-orthogonality of linguistic and social predictors; in other words, their effects on variable SPE are not independent of each other but rather interact with each other. Interaction effects between internal predictors, on the one hand, and between external predictors, on the other, are well-documented in the variationist literature. However, the (much less documented) interaction effects observed in the current analysis reflect a more nuanced perspective of the interplay of internal and external predictors influencing variable SPE, and also raise new questions for future sociolinguistic research.

What remains to be seen, among other issues, is whether the aforementioned variation in *yo* frequencies is representative of the regional varieties sampled here. Future research that incorporates more speakers from each of the non-Mexico City regions is needed to determine if the wide variation in rates simply reflects individual speaker usage or if other regional dialects can be delimited based on pronoun frequencies.¹⁹ A thorough qualitative analysis of the discourse context could also help us begin to answer questions regarding the interaction of gender:age with internal predictors such as polarity, verb class, and TMA. Additionally, other linguistic as well as social predictors should be explored in future studies that would perhaps provide a more comprehensive analysis of the first-person singular pronoun behavior for this group of speakers. Finally, incorporating speakers of different national origins aside from Mexico would also engender a more representative view of Spanish in Georgia.

Acknowledgements

The author wishes to thank two anonymous reviewers for their invaluable suggestions in improving the paper. The data collection for the *Corpus of Spanish in Georgia* was funded by the University of Georgia Graduate School Dean's Award, and the transcription of interviews was partially funded by the Willson Center for Humanities and Arts at the University of Georgia. The author also wishes to thank Lilian Zhu and Trevor Talmadge for their transcription assistance.

Notes

¹ For further discussion, the reader is referred to studies of Mexican Spanish in Mexico (e.g. Lastra & Martín Butragueño 2015; Michnowicz 2015; Shin & Erker 2015; Orozco 2016) as well as Mexican Spanish in the U.S. (e.g. Silva-Corvalán 1994; Flores-Ferrán 2007; Shin & Erker 2015; Shin & Van Buren 2016).

² See Table 1 (Section 2) for details on the particular speech communities analyzed among different SPE studies.

³ For extensive discussions on Spanish in the U.S., see, for example, Silva-Corvalán (1994); Lipski (2008); Klee & Lynch (2009); Escobar & Potowski (2015); Fuller & Leeman (2020).

⁴ The studies referenced here included all persons/numbers, not solely first-person singular.

⁵ See Limerick (2019) for a language contact/bilingualism perspective on immigrant Spanish in Georgia.

⁶ Speaker codes indicate sex and age.

⁷ Some differences in terminology and methods among these studies should be noted: Lastra & Martín Butragueño (2015: 46) use the term ‘enunciative type’ to refer to polarity (positive vs negative), but with the addition of interrogative tokens being included in a ‘non-affirmative’ category. Other researchers cited in this section use the terms ‘polarity’ (Travis & Torres Cacoullos 2012) and ‘verbal negation’ (Geeslin & Gudmestad 2016) and simply use the classification of affirmative vs negative sentences. See Section 3 for further methodological details.

⁸ Some speakers did not specify their city of origin.

⁹ Although many other predictors condition SPE in Spanish, as shown in Section 2, only four are examined here to avoid overfitting of the statistical models. Given the relatively modest number of speakers and tokens for each of the two groups, the number of predictor variables included in the analysis was constrained in order to maximize the reliability of the results. Furthermore, employing these four predictors is a reflection of their importance and consistent use in the SPE literature. See Limerick (2019) for an analysis that includes numerous other predictors not examined here and their impact on SPE in Mexican Spanish spoken in Georgia more broadly.

¹⁰ A question may arise concerning the potential influence of language contact/bilingualism on the overt SP rates observed in the current data compared to monolingual Mexican varieties. Nonetheless, previous research of this same group of speakers has ruled out this possibility (Limerick 2019). In particular, no significant correlations were found between overt pronoun rates and English proficiency, length of residency in the U.S., or age of arrival to the U.S.

¹¹ These categories are based on Bentivoglio *et al.* (2011). The author continues to employ the traditional verb class categories used in the variationist literature in order to ensure comparability with previous studies. However, the more innovative methods proposed by Orozco (2018) and Orozco & Hurtado (2021) will also be employed in the current analysis (see Section 4).

¹² I adopt Lastra & Martín Butragueño’s (2015) categories for what they call ‘enunciative type’.

¹³ It should be noted that this difference in rates of *yo* is not statistically significant ($p = 0.13$), as observed by a separate regression model that included Region (Mexico City vs Non-Mexico City).

¹⁴ Regarding verb class, some of the levels were combined after the initial regressions due to issues of overlap/collinearity for both regional models. Specifically, the % overt did not descend in order with the FWs. For Mexico City, *communicative*

showed a FW of .45 with 26% overt while *activity* had a factor weight of .42 with 27% overt. After combining the levels, the new model was not statistically significantly different from the previous one as revealed by a chi-square test in *Rbrul* ($p=0.378$). Regarding the Non-Mexico City data, *communicative* had a factor weight of .51 with 39% overt while *stative* had a FW of .50 with 40% overt. As with the Mexico City data, a comparison of the two models did not show a statistically significant difference ($p=0.913$). See Tagliamonte (2012) for further discussion of such statistical issues.

¹⁵ This analysis employs the entire data set as a whole, that is, Mexico City and Non-Mexico City speakers combined.

¹⁶ Aside from the analysis of different speech communities, the inconsistencies in the results between studies could also be due to the fact that the current study solely examines 1SG subjects while both Orozco (2016) and Orozco & Hurtado (2021) analyzed all persons/numbers.

¹⁷ Orozco & Hurtado (2021) also argued that *creer* and *pensar* were responsible for the overall cognitive verb effect in Medellín.

¹⁸ The tokens for these 2 exceptional cases were *yo creía* 'I thought' and *crea* '(I) believe'

¹⁹ Another potential source for such individual variation could have to do with the mere distribution of the data. If we compare, for example, the highest pronoun user (F56, 57% overt SPE) with the lowest pronoun user (M43, 12% overt SPE), the mere proportions of contexts that favor overt SPs are quite different. For instance, for cases of imperfect aspect, switch reference, and mental verbs, F56 simply produced more tokens than M43 (Imperfect: 53/184 (29%) vs 23/123 (19%); Switch reference: 80/184 (43%) vs 43/123 (35%); Mental verbs: 36/184 (20%) vs 16/123 (13%), respectively). Future research is needed to tease apart distribution from regional differences to determine to what extent each of these predictors play a role in SP variation.

Bibliographical References

- Abreu, Laurel 2012. Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish. In Geeslin, Kimberly & Díaz-Campos, Manuel (eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project. 1-8.
- Aijón Olivia, Miguel Ángel & Serrano, María José 2010. El hablante en su discurso: expresión y omisión del sujeto de creó. *Oralia* 13. 7-38.
- Alfaraz, Gabriela G. 2015. Variation of overt and null subject pronouns in the Spanish of Santo Domingo. In Carvalho; Orozco & Shin (eds.) 2015. 3-16.
- Ariel, Mira 1994. Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. *Journal of Linguistics* 30. 3-42.
- Ávila-Jiménez, Bárbara 1996. *Subject pronoun expression in Puerto Rican Spanish: A sociolinguistic morphological, and discourse analysis*. PhD dissertation. Cornell University.
- Bayley, Robert; Greer, Kristen & Holland, Cory C. 2013. Lexical frequency and syntactic variation: A test of a linguistic hypothesis. *University of Pennsylvania Working Papers in Linguistics* 19,2. 21-30.
- Bayley, Robert & Pease-Alvarez, Lucinda 1996. Null and expressed pronoun variation in Mexican-descent children's Spanish. In Arnold, Jennifer; Blake, Renee & Davidson, Brad (eds.), *Sociolinguistic Variation: Data, Theory, and*

- Analysis*. Stanford: Center for the Study of Language and Information. 85-99. <doi.org/10.1017/S0954394500001964>.
- Bayley, Robert & Pease-Alvarez, Lucinda 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9. 349-371. <doi.org/10.1017/S0954394500001964>.
- Bentivoglio, Paola; Ortiz, Luis A. & Silva-Corvalán, Carmen 2011. La variable expresión del sujeto pronominal. Guía de codificación. In *PRESEEA – Grupo de análisis de expresión de sujetos*. Retrieved from <preseea.linguas.net/portals/0/metodologia/guia_codificacion_sujetos_julio_2011.pdf>.
- Bentivoglio, Paola 1987. *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Central University of Venezuela.
- Bessett, Ryan M. 2018. Testing English influence on first person singular “yo” subject pronoun expression in Sonoran Spanish. In J MacDonald, Jonathan E. (ed.), *Contemporary Trends in Hispanic and Lusophone Linguistics: Selected papers from the Hispanic Linguistic Symposium 2015*. Amsterdam: John Benjamins. 355-372. <doi.org/10.1075/ihll.15.17bes>.
- Cameron, Richard 1992. Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation. PhD dissertation. University of Pennsylvania.
- Cameron, Richard 1994. Switch reference, verb class and priming in a variable syntax. Papers from the regional meeting of the Chicago Linguistic Society: Parasession on variation in linguistic theory 30. 27-45.
- Cameron, Richard 1995. The scope and limits of switch reference as a constraint on pronominal subject expression. *Hispanic Linguistics* 6,7. 1-27.
- Carvalho, Ana M. & Bessett, Ryan M. 2015. Subject pronoun expression in Spanish in contact with Portuguese. In Carvalho; Orozco & Shin (eds.) 2015. 143-165.
- Carvalho, Ana M. & Child, Michael 2011. Subject pronoun expression in a variety of Spanish in contact with Portuguese. In Michnowicz, Jim & Dodsworth, Robin (eds.), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 14-25.
- Carvalho, Ana M.; Orozco, Rafael & Shin, Naomi. L. (eds.) 2015. *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown: Georgetown University Press.
- Cerrón-Palomino, Álvaro 2018. Variable subject pronoun expression in Andean Spanish: A drift from the acrolect. *Onomázein* 42. 53-73.
- Enríquez, Emilia V. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
- Erker, Daniel & Guy, Gregory R. 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88. 526-557. <doi.org/10.1353/lan.2012.0050>.
- Escobar, Anna María & Potowski, Kim 2015. *El español de los Estados Unidos*. Cambridge: Cambridge University Press.
- Flores-Ferrán, Nydia 2002. *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: A sociolinguistic perspective*. Munich: Lincom Europa.
- Flores-Ferrán, Nydia 2004. Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English con-

- tact? *Language Variation and Change* 16. 49-73. <doi.org/10.1017/S0954394504161048>.
- Flores-Ferrán, Nydia 2007. Los mexicanos in New Jersey: Pronominal expression and ethnolinguistic aspects. In Holmquist, Jonathan; Lorenzino, Augusto & Sayahi, Lotfi (eds.), *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 85-91.
- Fuller, Janet M. & Leeman, Jennifer 2020. *Speaking Spanish in the US: The Sociopolitics of Language*. Bristol, Blue Ridge Summit: Multilingual Matters.
- Geeslin, Kimberly & Gudmestad, Aarnes 2016. Subject expression in Spanish: Contrasts between native and non-native speakers for first and second-person singular referents. *Spanish in Context* 13,1. 53-79. <doi.org/10.1075/sic.13.1.03gee>.
- Geeslin, Kimberly; Linford, Bret & Fafulas, Stephen 2015. Variable subject expression in second language Spanish: Uncovering the developmental sequence and predictive linguistic factors. In Carvalho; Orozco & Shin (eds.) 2015. 191-210.
- Givón, Talmy 1983. Topic continuity in discourse: An introduction. In Givón, Talmy (ed.), *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins. 1-41.
- Gumperz, John J. 1972 [1968]. The Speech Community. In Giglioli, Pier Paolo (ed.), *Language and Social Context*. London: Penguin Books. 219-231.
- Hochberg, Judith G. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62,3. 609-621.
- Johnson, Daniel E. 2009. Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3. 359-383. <doi.org/10.1111/j.1749-818X.2008.00108.x>.
- Klee, Carol A. & Lynch, Andrew 2009. *El español en contacto con otras lenguas*. Washington, D.C.: Georgetown University Press.
- Lastra, Yolanda & Martín Butragueño, Pedro 2015. Subject pronoun expression in oral Mexican Spanish. In Carvalho; Orozco & Shin (eds.) 2015. 39-57.
- Limerick, Philip 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. 53-77.
- Limerick, Philip P. 2018. Subject expression in a Southeastern U.S. Mexican community. PhD dissertation. Athens, GA.
- Limerick, Philip P. 2019. Subject expression in a Southeastern U.S. Mexican community. *Borealis – An International Journal of Hispanic Linguistics* 8,2. 243-273.
- Limerick, Philip P. 2021. First-person plural subject pronoun expression in Mexican Spanish spoken in Georgia. *Studies in Hispanic and Lusophone Linguistics* 14,2. 411-432.
- Limerick, Philip P. 2022. *Corpus of Spanish in Georgia*. Available at <corpusofspanishingeorgia.omeka.net>
- Lipski, John M. 2008. *Varieties of Spanish in the United States*. Washington, D.C.: Georgetown University Press.
- Martínez-Lara, José A.; Guerrero González, Silvana & González Vergara, Carlos 2021. (Yo) opino: la variación del sujeto pronominal en el habla de Santiago de Chile. In Rogers, Brandon M.A. & Figueroa Candia, Mauricio A.

- (eds.), *Lingüística del castellano chileno: estudios sobre variación, innovación, contacto e identidad*. Vernon Press. 283-308.
- Meyerhoff, Miriam 2009. Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change* 21. 297-317.
- Michnowicz, Jim. 2015. Subject pronoun expression in contact with Maya in Yucatan Spanish. In Carvalho; Orozco & Shin (eds.) 2015. 101-119.
- Orozco, Rafael & Guy, Gregory R. 2008. El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana? In Westmoreland, Maurice & Thomas, Juan Antonio (eds.), *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 70-80.
- Orozco, Rafael. 2015. Pronominal variation in Colombian Costeño Spanish. In Carvalho; Orozco & Shin (eds.) 2015. 17-37.
- Orozco, Rafael 2016. Subject pronoun expression in Mexican Spanish: ¿Qué pasa en Xalapa? In *Proceedings of the Linguistic Society of America*, Vol. 1, Article 7. 1-15. <doi.org/10.3765/plsa.v1i0.3703>.
- Orozco, Rafael 2018. El castellano del Caribe colombiano en la ciudad de Nueva York: El uso variable de sujetos pronominales. *Studies in Lusophone and Hispanic Linguistics* 11,1. 89-129. <doi.org/10.1515/shll-2018-0004>.
- Orozco, Rafael & Hurtado, Luz Marcela 2021. A Variationist Study of Subject Pronoun Expression in Medellín, Colombia. *Languages* 6,5. 1-29. <doi.org/10.3390/languages6010005>.
- Otheguy, Ricardo & Zentella, Ana Celia 2012. *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford University Press: New York.
- Otheguy, Ricardo; Zentella, Ana Celia & Livert, David 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 83. 770-802. <doi.org/10.1353/lan.2008.0019>.
- Padilla, Lillie 2021. First person singular subject pronoun expression in Equatoguinean Spanish. *Journal of Monolingual and Bilingual Speech* 3,2. 171-194.
- Poplack, Shana 1997. The sociolinguistic dynamics of apparent convergence. *Amsterdam studies in the theory and history of linguistic science series* 4. 285-310.
- Posio, Pekka 2011. Spanish subject pronoun usage and verb semantics revisited: First and second person singular subject pronouns and focusing of attention in spoken Peninsular Spanish. *Journal of Pragmatics* 43,3. 777-798.
- Posio, Pekka 2015. Subject pronoun usage in formulaic sequences. In Carvalho; Orozco & Shin (eds.) 2015. 59-78.
- Prada Pérez, Ana de 2009. *Subject expression in Minorcan Spanish: Consequences of contact with Catalan*. PhD dissertation. Pennsylvania State University, University Park, PA.
- Prada Perez, Ana de 2015. First person singular subject pronoun expression in Spanish in contact with Catalan. In Carvalho; Orozco & Shin (eds.) 2015. 121-142.
- Ranson, Diana L. 1991. Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change* 3. 133-152. <doi.

- org/10.1017/S0954394500000491 > .
- Scrivner, Olga & Díaz-Campos, Manuel 2016. Language Variation Suite: A Theoretical and Methodological Contribution for Linguistic Data Analysis. In *Proceedings of the Linguistic Society of America*, Vol. 1, Article 29. 1-15.
- Shin, Naomi Lapidus 2014. Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change* 26. 303-30. <doi.org/10.1017/S095439451400012X> .
- Shin, Naomi Lapidus & Erker, Daniel 2015. The emergence of structured variability in morphosyntax: Childhood acquisition of Spanish subject pronouns. In Carvalho; Orozco & Shin (eds.) 2015. 169-190.
- Shin, Naomi Lapidus & Montes-Alcalá, Cecilia 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. 85-110. <doi.org/10.1558/sols.v8i1.85> .
- Shin, Naomi Lapidus & Van Buren, Jackelyn 2016. Maintenance of Spanish subject pronoun expression patterns among bilingual children of farmworkers in Washington/Montana. *Spanish in Context* 13,2. 173-194. <doi.org/10.1075/sic.13.2.01shi> .
- Shin, Naomi Lapidus & Otheguy, Ricardo 2009. Shifting sensitivity to continuity of reference: Subject pronoun use in Spanish in New York City. In LaCorte, Manel & Leeman, Jennifer (eds.), *Español en Estados Unidos y otros contextos: Sociolingüística, ideología y pedagogía*. Madrid: Iberoamericana. 111-136.
- Shin, Naomi Lapidus & Otheguy, Ricardo 2013. Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. *Language in Society* 42,4. 429-452. <doi.org/10.1017/S0047404513000468> .
- Silva-Corvalán, Carmen 1982. Subject expression and placement in Mexican-American Spanish. In Amastae, Jon & Elías-Olivares, Lucía (eds.), *Spanish in the United States: Sociolinguistic aspects*. Cambridge: Cambridge University Press. 93-120.
- Silva-Corvalán, Carmen 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon.
- Silva-Corvalán, Carmen 2001. *Sociolingüística y pragmática del español*. Washington, D.C.: Georgetown University Press.
- Solomon, Julie 1999. *Phonological and syntactic variation in the Spanish of Valladolid, Yucatán*. PhD dissertation. Stanford University, Stanford, CA.
- Tagliamonte, Sali A. 2002. Comparative sociolinguistics. In Chambers, J.K.; Trudgill, Peter & Schilling-Estes, Natalie (eds.), *The handbook of language variation and change*. Malden / Oxford: Blackwell. 729-763.
- Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge, NY: Cambridge University Press. <doi.org/10.1017/CBO9780511801624> .
- Tagliamonte, Sali A. 2012. *Variationist sociolinguistics: Change, observation, interpretation*. Malden, MA: Wiley-Blackwell.
- Torres Cacoullous, Rena & Travis, Catherine E. 2010. Variable yo expression in New Mexico: English influence? In Rivera-Mills, Susana & Villa, Daniel (eds.), *Spanish of the U.S. Southwest: A language in transition*. Madrid: Iberoamericana/Vervuert. 185-206. <doi.org/10.31819/9783865278692-012> .

- Torres Cacoulios, Rena & Travis, Catherine E. 2015. Foundations for the study of subject pronoun expression in Spanish in contact with English: Assessing interlinguistic (dis)similarity via intralinguistic variability. In Carvalho; Orozco & Shin (eds.) 2015. 81-100
- Travis, Catherine E. 2005. The yo-yo effect: Priming in subject expression In Colombian Spanish. In Gess, Randall & Rubin, Edward J. (eds.), *Selected Papers from the 34th Linguistic Symposium on Romance Languages (LSRL)*. Salt Lake City, 2004. Amsterdam: Benjamins. 329-349. <doi.org/10.1075/cilt.272.20tra> .
- Travis, Catherine E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19. 101-135. <doi.org/10.1017/S0954394507070081 > .
- Travis, Catherine E. & Torres Cacoulios, Rena 2012. What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics* 23,4. 711-748. <doi.org/10.1515/cog-2012-0022> .
- United States Census Bureau 2020. Quick Facts: Roswell city, Georgia. Retrieved from <www.census.gov/quickfacts/fact/table/roswellcitygeorgia/POP010220> .
- Villarreal, Dan; Clark, Lynn; Hay, Jennifer & Watson, Kevin 2021. Gender separation and the speech community: Rhoticity in early 20th century Southland New Zealand English. *Language Variation and Change* 33,2. 245-266.
- Wolfram, Walt; Kohn, Mary E. & Callahan-Price, Erin 2011. Southern-Bred Hispanic English: An emerging socioethnic variety. In Michnowicz, Jim & Dodsworth, Robin (eds.), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 1-13.