

SPANISH SUBJECT EXPRESSION IN ROSWELL, GEORGIA: DIALECT FORMATION IN  
AN EMERGING BILINGUAL COMMUNITY

by

PHILIP P. LIMERICK

(Under the Direction of Chad Howe)

ABSTRACT

This study examines the use of Spanish subject pronouns in Roswell, Georgia. Specifically, I investigate the influence of linguistic and social variables on pronoun use in this variety of Spanish, which has not received the same degree of attention as well-studied varieties (e.g. New York). Sociolinguistic interviews were conducted in Roswell (Wilson 2013) and transcribed to allow for analysis of pronouns and factors that may influence subject expression (e.g. person/number) as well as social variables (e.g. length of residency). Results indicate an overall pronoun rate of 21%, similar to that of Mainland newcomers in New York (Otheguy et al. 2007). However, results from the multivariate analysis suggest that pronoun usage in Roswell diverges from these communities, with differential effects observed for factors such as Person/number and Coreferentiality Index. This analysis of subject expression in the Roswell data reveals an intermediate stage of language shift in this particular community.

INDEX WORDS: subject pronouns, Spanish in the United States, Georgia, language shift dialect formation, Roswell, Mexican Spanish

SPANISH SUBJECT EXPRESSION IN ROSWELL, GEORGIA: DIALECT FORMATION IN  
AN EMERGING BILINGUAL COMMUNITY

by

PHILIP P. LIMERICK

B.A., Indiana University of Pennsylvania, 2009

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment  
of the Requirements for the Degree

MASTER OF ARTS

ATHENS, GEORGIA

2014

© 2014

Philip P. Limerick

All Rights Reserved

SPANISH SUBJECT EXPRESSION IN ROSWELL, GEORGIA: DIALECT FORMATION IN  
AN EMERGING BILINGUAL COMMUNITY

by

PHILIP P. LIMERICK

Major Professor: Chad Howe  
Committee: Sarah Blackwell  
Margaret Quesada

Electronic Version Approved:  
Maureen Grasso  
Dean of the Graduate School  
The University of Georgia  
May 2014

## DEDICATION

This thesis is dedicated to my mother and my father for all their love, support, guidance, and encouragement during my first two years of graduate school.

## ACKNOWLEDGEMENTS

There are numerous people to whom I would like to express my gratitude for helping to make this thesis possible. First, I would like to thank Dr. Chad Howe, my major professor, for all his help and guidance throughout this project. He has devoted a significant amount of time to helping me throughout all the different stages of analysis and writing and has provided me with invaluable insight and suggestions that have greatly improved this thesis, making for a final product I am truly happy with. I would also like to thank my committee members, Drs. Blackwell and Quesada, for their encouragement and helpful suggestions during the final stages of the thesis and also for the knowledge base that they provided me in the courses I have had the opportunity of taking with them. Second, I would like to thank my colleague, Anna Wilson, for her efforts in collecting data in Roswell and for making the interviews available to me for analysis.

Finally, I would like to express my sincere gratitude to all my family and friends and especially to my wonderful girlfriend and best friend, Chantell Smith, for her encouragement, interest, support, and love throughout this endeavor.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	v
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
CHAPTER	
1 INTRODUCTION .....	1
2 LITERATURE REVIEW .....	6
2.1 PRO-DROP IN SPANISH .....	6
2.2 PRAGMATIC CONSTRAINTS .....	7
2.3 SPANISH IN THE UNITED STATES .....	9
2.4 PRONOUN RATES AND COMMON LINGUISTIC AND SOCIAL FACTORS INFLUENCING PRONOUN USAGE .....	10
3 METHODOLOGY .....	17
3.1 DATA AND PARTICIPANTS .....	17
3.2 THE ENVELOPE OF VARIATION .....	19
3.3 CODING OF DATA .....	21
4 RESULTS .....	37
4.1 INDIVIDUAL AND OVERALL PRONOUN RATES .....	37

4.2 FACTORS CONDITIONING THE EXPRESSION OF SUBJECT	
PRONOUNS .....	40
4.3 LANGUAGE SHIFT IN ROSWELL .....	47
5 CONCLUSION AND DISCUSSION .....	51
REFERENCES .....	54



## LIST OF TABLES

	Page
Table 1: Demographic changes in Roswell population between 2000 and 2010 .....	4
Table 2: Overall PN rates of monolingual varieties.....	11
Table 3: Overall PN rates of contact varieties .....	11
Table 4: Variable hierarchies for Caribbean and Mainland newcomers in NYC .....	13
Table 5: Individual PN rates and sociodemographic information .....	38
Table 6: Overall rates of subject pronoun expression in Roswell .....	39
Table 7: Comparison of PN rates between Roswell Spanish and other varieties .....	40
Table 8: Hierarchy of variables related to subject expression .....	42
Table 9: Hierarchy of constraints on SP use.....	43
Table 10: Variable hierarchy as compared with Otheguy et al. (2007:789).....	48

LIST OF FIGURES

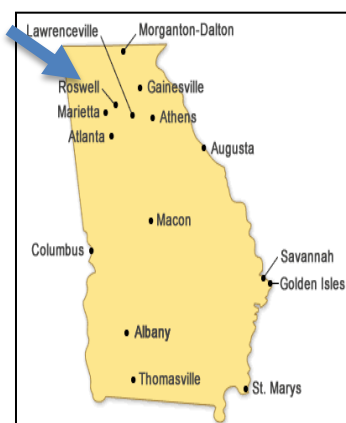
	Page
Figure 1: Map of Georgia .....	1
Figure 2: Population growth of Roswell, GA since 1950.....	3

## CHAPTER 1

### INTRODUCTION

The expression of subject pronouns in Spanish is one of the most extensively studied features both in monolingual and contact varieties of the language (Silva-Corvalán 1982; Bentivoglio 1987; Cameron 1993; Flores-Ferrán 2004; Otheguy, Zentella & Livert 2007; Travis 2007; Orozco & Guy 2008; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Abreu 2012; Erker & Guy 2012; Otheguy & Zentella 2012, among others). There are various linguistic and social factors that have been shown to influence pronoun usage, and while many of these conditioning factors are consistent among the many dialects of Spanish, the differing social situations and underlying grammars associated with each dialect have an effect on and reflect unique speech patterns concerning the explicit or implicit expression of subject pronouns in certain dialects. The general goal of this thesis is to examine the use of Spanish pronouns among speakers in Roswell, Georgia (see Figure 1), an exurb of Atlanta that does not fit into the typical urban/rural dichotomy.

Figure 1: Map of Georgia



This specific geographical region allows for linguistic analysis of a unique, emerging speech community comprised of a significantly increasing Latino population, which grew from approximately 10.6% of the city's total population in 2000 to 16.6% by the 2010 U.S. Census (U.S. Bureau of the Census 2010). Specifically, I investigate the influence of both linguistic and social variables on the distribution of subject pronoun use by speakers of this particular variety of US Spanish, which has not received the same degree of attention as well studied varieties (e.g. New York, Los Angeles). Furthermore, given that the language contact situation in Roswell is much less established than that of bilingual communities in the Southwest or Northeast, a study of this type allows for a unique opportunity to examine the process of linguistic change in its initial stages with regard to Spanish in the United States.

This thesis also contributes to the broader goals of a research organization called the Roswell Voices Project, which began in 2002 as a partnership between researchers at the University of Georgia and the Roswell Convention and Visitors Bureau (Kretzschmar et al. 2007). The primary goal of the Roswell Voices Project is to document life and language among the city's different communities, which has mainly involved research on the African-American and Caucasian communities in particular. The Latino community was first explored by Wilson (2013) by means of sociolinguistic interviews, leading to her analysis of narrative structure in Roswell Spanish.

What makes Roswell an interesting and important region for linguistic investigation is the nature of its growth and changing demographics (see Figure 2 and Table 1 below). Its population has grown immensely, from only a few thousand people in 1950 to approximately 100,000 in 2013 (Wilson 2013). In terms of its demographics, the population of the city has been largely Caucasian, with African-Americans being the largest minority group. However, by the year

2000, Latinos began to outnumber the African-American population and had experienced an approximate 75% increase between 2000 and 2010, putting them at 16.6% of Roswell's total population, which continued to increase to approximately 25% of the total population by 2013 (Wilson 2013).

Figure 2: Population growth of Roswell, GA since 1950. Source: (Wilson 2013:2)

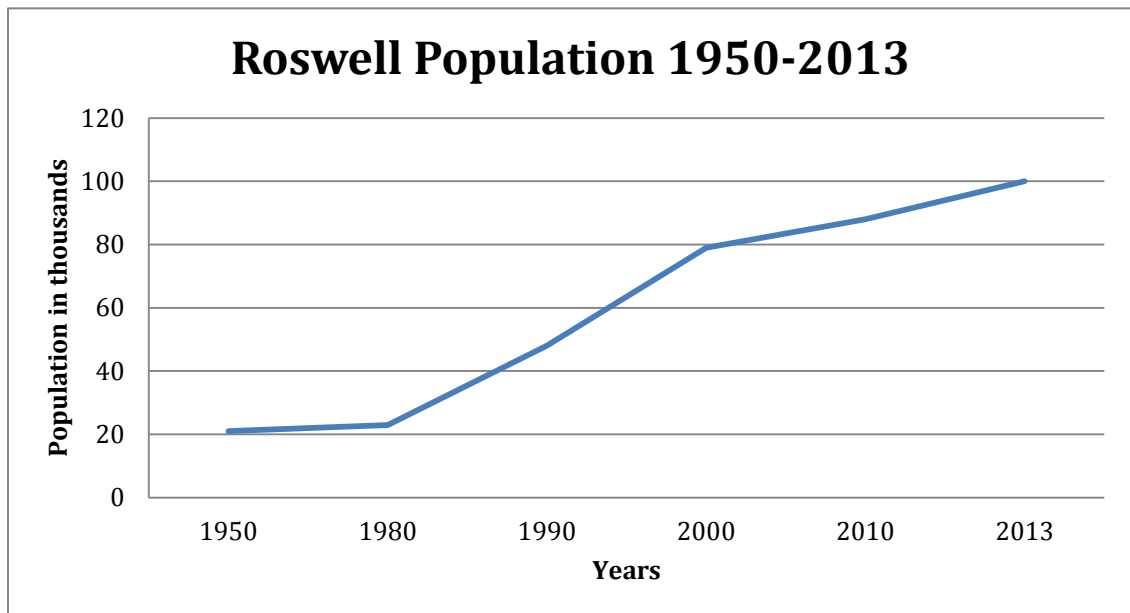


Table 1: Demographic changes in Roswell population between 2000 and 2010

Population	2000	2010
Total Population	79,334	88,346
White	64,666	66,010
Black or African American	6,773	10,373
Asian	2,964	3,565
Some Other Race	3,237	5,846
Two or More Races	1,511	2,241
American Indian or Alaska Native	160	261
Hawaiian Native/other Pacific Islander	23	50
Hispanic or Latino (of any race)	8,421	14,699

Source: U.S. Bureau of the Census, 2010 Census of Population

Returning to the present study, I intend to answer the following research questions:

- 1) How do the overall pronoun rates of Roswell speakers compare to more established contact varieties of Spanish in the US and to Mexican monolingual varieties of Spanish?
- 2) What linguistic variables have the greatest effect on subject pronoun use in Roswell?
- 3) How do social variables, such as age, gender, length of residency, and age of arrival influence pronoun use?
- 4) Is contact with English a significant factor contributing to language shift in Roswell Spanish?

In order to answer these questions, I carry out a quantitative analysis using sociolinguistic interviews that were conducted with Spanish speakers of Roswell, Georgia. I examine factors

such as person/number of the verb, same/switch reference, verb tense, and age of arrival to the US, among others, and their influence on subject expression in this particular Spanish variety.

In the next chapter, I describe the pro-drop phenomenon in Spanish and also provide a summary of previous research on Spanish in the United States in general. I then discuss overall pronoun rates and specific factors that have been shown to influence pronoun usage among different Spanish varieties. I conclude the chapter with a summary of previous studies of Spanish subject expression in the United States. In chapter three, I discuss the methodology for the present study, including a description of the data, participants, method of coding, and the linguistic and social variables included in the study. I then discuss the results of the study in chapter four, including the overall pronoun rate and the specific influencing factors of subject expression in Roswell as compared to other varieties of Spanish. The final chapter discusses general conclusions and considerations for future research.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 PRO-DROP IN SPANISH

Whereas the use of overt subject pronouns is almost always obligatory in English, Spanish is known as a *pro-drop* or null-subject language, in which it is possible to omit the subject pronoun, as seen in example (1), but serves a specific pragmatic purpose when expressed, as seen in (2).

(1)  $\emptyset$  Tengo que salir.

‘(I) have to leave.’

(2) **Yo** tengo que salir.

‘I have to leave.’

According to Bosque & Demonte (1999), the primary function of using overt pronouns in Spanish is that of emphasis or contrast. Explicit pronominal forms that replace null pronominal forms are meant to emphasize the referent or to express contrast among different referents. On the other hand, the absence of the overt pronoun (the null pronoun) is neutral, in other words, not emphatic. Therefore, in example (2) above, the presence of the first person singular pronoun *yo* ‘I’ could serve to emphasize the person speaking and perhaps to show a contrast with other speakers in the conversation (i.e. *I* as opposed to *you/she/he*). However, when there is no particular emphasis desired, the speaker can omit the *yo*, while still leaving the sentence perfectly



grammatical, as seen in example (1). Additionally, overt subject pronouns can be used in contexts of potential ambiguity, for example, when the verb is in the third person:

(3) Ya decía **yo** (Bosque & Demonte 1999:1282).

‘I was already saying’

In (3), the subject pronoun could have also been *él* ‘he’, *ella* ‘she’, or *usted* ‘you (formal)’ according to the verbal inflection. The overt pronoun *yo*, however, is necessary in this sentence to disambiguate the referent, unless there is previous context from which the referent could be identified.

These are just a few of the reasons why a speaker might choose to use an overt instead of a null pronoun. As I demonstrate throughout this paper, however, the choice to use or not to use an overt pronoun is much more complex and is in fact systematically conditioned by semantic, pragmatic, and morphosyntactic factors. In addition, there are contexts in which either the overt pronoun, or more frequently, the null pronoun is obligatory in Spanish, which I discuss in the next section.

## 2.2 PRAGMATIC CONSTRAINTS

Numerous studies have shown that subject pronoun usage in Spanish, as mentioned above, is influenced by discourse-pragmatic factors (Enríquez 1984; Luján 1986; Davidson 1996; Quesada & Blackwell 2009; Blackwell & Quesada 2012). Specifically, speakers tend to use overt SPs in conversation to carry out certain communicative functions such as to express emphasis or contrast (*Él tiene dieciocho años y yo tengo veinticinco años* ‘He is eighteen years old and I am twenty-five years old’), to clarify a potentially ambiguous referent (*Decía yo que...* ‘I was saying that...’), or to add pragmatic weight to an utterance (Davidson 1996). The notion of pragmatic weight is described by Davidson as a way for speakers to “increase their ‘stake’ in

whatever they are saying, either in an argument or in a statement of belief” (1996:551).

Therefore, it is shown that speakers use overt SPs to achieve a specific communicative goal, thereby demonstrating that their use is not random or simply optional. For example, Quesada & Blackwell (2009) examined first person singular SPs in relation to five pragmatic rules or constraints that restrict their use among both Spanish L2 learners and native speakers: *salient referent*, *epistemic parenthetical* (which require null SPs), *switch focus*, *contrastive focus*, and *pragmatic weight* (which require overt SPs). The authors found that while native speakers and advanced L2 learners generally followed such patterns of pragmatic constraints, beginner and intermediate learners diverged from these patterns, exhibiting a superfluous use of both overt and null SPs. Quesada & Blackwell's study, among many others, demonstrates that SPs are not simply used in free variation, but are systematically conditioned in Spanish.

While it is essential to consider such discourse-pragmatic functions with regard to subject expression, the primary focus of the present study, given its quantitative nature, is on the specific language internal (grammatical) and language external (social) factors that condition SP use. For instance, it would not be feasible to include the factor of pragmatic weight mentioned above since it is difficult to quantify and operationalize in a study of this type. However, I do include factors such as person/number, same/switch reference, and clause type, which have been quantified in variationist studies and shown to significantly influence subject pronoun usage. These factors are discussed more specifically below. In the next section, I discuss some general features of Spanish in the United States, and then talk about the use of subject pronouns in particular in both monolingual and US contact varieties of Spanish.

## 2.3 SPANISH IN THE UNITED STATES

There are numerous varieties of Spanish in the United States, which often exhibit unique linguistic features which differ from the typical patterns of their respective monolingual varieties. Such differential features are often claimed to be the result of influence from English. For instance, regarding Spanish in New York, Otheguy, García, & Fernández observed distinct characteristics among Cuban speakers such as the use of loan words (*La casa tiene un **basement** [bEyhmEn]* ‘The house has a basement’), code-switching (*Y si queremos **be**, to be good, we would have to agree on one thing, y no es así* ‘And if we want to be, to be good, we would have to agree on one thing, and it’s not like that’) and word calques (*A mí me gustan las **cartas de béisbol para coleccionarlas*** ‘I like baseball cards because I can collect them’) (1989:43-44). The authors also noted an increase in such usage from first generation to second generation Cubans. Moreover, Silva-Corvalán (1994) observed unique syntactic and pragmatic features among Mexicans in Los Angeles. For example, she noted that Mexican speakers often omit the complementizer *que* ‘that’ as in *Yo creo  $\emptyset$  inventaron el nombre* ‘I think (that) they invented the name’ (p. 136). Specifically, such non-expression of *que* is found only in complement clauses and not in relative clauses (*\*El nombre  $\emptyset$  (ellos) inventaron era extraño* ‘The name (that) they invented was strange’ (Silva-Corvalán 1994:137). Another feature observed in this variety is the use of SV word order when introducing new subject referents into discourse, a context in which the more common structure in monolingual Spanish is VS (Silva-Corvalán 1994). For example, the first mention of a teacher might be expressed as *Llegó la profesora* ‘The teacher arrived’ by a monolingual Spanish speaker while a bilingual might say *La profesora llegó* ‘The teacher arrived’ in the same discursive context. The abovementioned examples are just a few of the

specific features found in Spanish-English contact varieties in the US, and several others have been observed in various regions.

With regard to research conducted specifically on dialect formation in the Mid-Atlantic U.S. South, Wolfram, Kohn, and Callahan-Price (2011) examined the contact between English and Spanish which has resulted in the formation of an emerging ethnic variety of English. Specifically, the authors analyzed the /ai/ diphthong in the speech of both rural and urban speakers in North Carolina, focusing on both the duration and tone of the glide. Their findings indicate that the /ai/ sound of the speakers is somewhere between the Spanish and English targets, making it an “interdialectal” feature that is not originally found in either language. The existence of such vowel sounds, as well as other observed interdialectal features, is a defining characteristic of this emerging socioethnic variety of Hispanic English in North Carolina.

#### 2.4 PRONOUN RATES AND COMMON LINGUISTIC AND SOCIAL FACTORS INFLUENCING PRONOUN USAGE

For the purposes of this study, and following common practice in the analysis of subject pronoun expression, I use the term *overt pronoun* to indicate the presence of a subject pronoun, as seen below in example (4) and the term *null pronoun* when it is absent, as seen in (5).

- 4) Yo empecé a trabajar.  
'I began working.'
- 5) Ø Empecé a trabajar.  
'I began working.'

As a point of reference and for comparative purposes, below I present two tables containing the reported overall overt pronoun rates for numerous monolingual varieties (Table 2) as well as

Spanish varieties in contact with English (Table 3). It is important to note the cross-dialectal variability observed in pronoun rates, both among monolingual Spanish as well as between monolingual and contact varieties.

Table 2: Overall PN rates of monolingual varieties

<b>Variety</b>	<b>PN Rate</b>
Valladolid, Mexico	21%
Puente Genil, Spain	24%
Buenos Aires, Argentina	36%
Santiago, Chile	38%
San Juan, Puerto Rico	45%
Santiago, Dominican Republic	39%
New York Newcomers (Mexican, S. American, Caribbean)	30%

Sources: Valladolid (Solomon 1999) Puente Genil (Ranson 1991) Buenos Aires (Barrenechea & Alonso 1977) Santiago, Chile (Cifuentes 1980), San Juan (Cameron 1994), Santiago, Dominican Republic (Olloqui de Montenegro 1987), New York Newcomers (Otheguy et al. 2007)

Table 3: Overall PN rates of contact varieties

<b>Variety</b>	<b>PN Rate</b>
Los Angeles Mexicans	43%
New York Dominicans	48%
New York born/raised (Mexican, S. American, Caribbean)	38%

Sources: Los Angeles (Silva-Corvalan 1982) New York Dominicans (Flores & Toro 2000), New York Born/Raised (Otheguy et al. 2007) North Carolina (McKnight 2013)

With regard to monolingual Spanish, Mexico and Spain tend to have the lowest pronoun rates, South American varieties such as Argentina and Chile exhibit mid-range pronoun rates, and Caribbean Spanish speakers generally have the highest rates out of all Spanish dialects.

However, as one can see in the above table, varieties of Spanish in contact with English generally show much higher rates when compared to monolingual varieties. Such presence of higher overt pronoun frequencies in contact varieties is commonly attributed to English contact, as overt subject pronouns are nearly obligatory in English but can be frequently omitted in Spanish. Therefore, increased exposure to English has been argued to engender a higher usage of overt subject pronouns in Spanish (Lapidus & Otheguy 2005). The language-contact hypothesis in relation to Roswell Spanish is an issue that will be addressed in the final sections of this paper.

Regarding the specific linguistic and social factors that tend to influence pronoun usage, the following variables have been shown to have a significant effect on subject expression across most dialects of Spanish, all of which will be discussed further in subsequent sections of the paper: person and number of the verb (Silva-Corvalán 1994), same vs. switch reference (Bayley and Pease-Álvarez 1996), clause type where the verb appears (Morales 1997), tense-mood-aspect of the verb (TMA, Cameron 1994), lexical content of the verb (Travis 2007), specificity (specific or non-specific reference, Cameron 1992), reflexivity (reflexive or non-reflexive use of the verb, Otheguy et al. 2007), previous expression (whether the pronoun of the previous verb is null or overt, Travis 2007), as well as social factors including age, gender, social class, and level of education (Carvalho & Child 2011). However, in addition to the above factors there is a different set of variables that play a role in contact varieties of Spanish including length of residency in the US, age of arrival to the US, as well as the degree of use, proficiency, and exposure to English (Otheguy et al. 2007). While many dialects share these same conditioning factors, there are also notable dialectal differences in relation to the underlying grammar of particular dialects, which has been claimed to account for the variability in how subject expression is manifested. For example, Otheguy et al. (2007:789) found dialectal differences among Caribbean and

Mainland newcomers in New York. The following variable hierarchies from their study illustrate differences in both the ranking of conditioning variables as well as which variables affect pronoun use in each dialect.

Table 4: Variable hierarchies for Caribbean and Mainland newcomers in NYC (Otheguy et al. 2007:789)

	CARIBBEAN NEWCOMERS	MAINLAND NEWCOMERS
<b>Rank</b>	<b>Variable</b>	<b>Variable</b>
1st	Person	Person
2nd	Connect	Connect
3rd	Tense	Set Phrase
4th	Clause	Tense
5th	Lexical	Clause
6th	Genre	Lexical
7th	Picture	Reflexive

For instance, the variables ‘Reflexive’ and ‘Set Phrase’ are not significant in relation to pronoun usage among Caribbeans, whereas they are significant among Mainlanders.<sup>1</sup> Further, the variables ‘Tense,’ ‘Clause,’ ‘Lexical,’ and ‘Genre’ all have a stronger influence on subject expression in Caribbean Spanish than they do in Mainland Spanish, as seen in their respective hierarchical rankings. In fact, the only variables that are consistent in their degree of influence among the two dialects are ‘Person’ and ‘Connect,’ which are ranked first and second, respectively.

One of the first studies carried out in the US regarding Spanish subject expression was Silva-Corvalán’s (1982) analysis of Mexican-American Spanish in Los Angeles. The objective

---

<sup>1</sup> Otheguy et al. use this term to refer to the non-Caribbean speakers in their study, namely Mexicans, Ecuadorians, and Colombians.

of the study was to determine how linguistic variables such as same/switch reference and morphological ambiguity of the verb influence the occurrence of overt versus null subjects in Spanish. The conversational speech of 27 Mexican-Americans who had been living in Los Angeles for at least ten years was analyzed, revealing that switch reference subjects, particularly in situations of contrast, favored overt pronouns while same reference subjects were rarely expressed. Furthermore, ambiguous verb forms favored overt subjects which served to clarify the subject referent. Silva-Corvalán concludes, through a comparison with other dialects, that same/switch reference and morphological ambiguity are constraints that reflect a general characteristic of the Spanish language.

More recently, Otheguy, Zentella & Livert (2007) investigated subject pronoun usage in New York City (NYC) Spanish in order to examine the effects of contact with English on the pronoun rates of two main groups of speakers: Mainlanders (Mexicans, Colombians, Ecuadorians) and Caribbeans (Cubans, Dominicans, Puerto Ricans). The researchers conducted sociolinguistic interviews with the speakers and analyzed, using a variationist approach, the influence of numerous linguistic and social variables on subject pronoun expression. With regard to independent linguistic variables, person, tense and clause type were shown to have the greatest effect on pronoun usage. In general, they found that speakers who have lived in NYC for longer periods of time, whether they were born there or are well established residents, have higher overt pronoun rates than those who have recently arrived, suggesting a significant influence of English on Spanish in New York. These findings present evidence of contact-induced language change and the formation of a distinctive variety of US Spanish.

Utilizing the same corpus as Otheguy et al. (2007), Shin & Otheguy (2013) employ a broader analysis of the influence of social factors, namely social class and gender, on the



increase of pronoun rates in the speech of New York Latinos. Specifically, the authors find a correlation of more affluent communities—Colombians and Cubans—with a greater increase in overt pronoun rates while communities of a lower socioeconomic status, namely Mexicans, Ecuadorians and Dominicans, are more resistant to such language change and do not show a significant increase in overt pronoun rates. Furthermore, the study reveals that women are the leaders of said linguistic change, which is consistent with previous research on gender and language change (Eckert & McConnell-Ginet 2003; Romaine 2003). The authors found that women showed higher overt pronoun rates than men, a finding they attribute to women's more extensive contact with second generation Latinos, including their own children, who exhibit higher overt pronoun rates.

While Otheguy, Shin, and other researchers (Otheguy et al. 2007; Shin & Otheguy 2013) are in favor of a language contact hypothesis based on increases in SP frequencies (among other factors), the evidence that is considered to be necessary to conclude that language change is in fact contact-induced or is occurring at all is highly debated among scholars. For example, Poplack & Levey (2010) posit that changes in SP overall frequencies do not constitute evidence for change at all, whether it be contact-induced or otherwise. These authors affirm that there must be observable differences in constraint hierarchies, that is, the ranked order of factors that are probabilistically shown to favor or disfavor the occurrence of SPs. Torres Cacoullos & Travis (2010) agree with Poplack & Levey 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 also suggest that divergences in grammatical patterning across dialects, which no studies have reported, are necessary to substantiate contact-induced change (p. 4).

The first and, to the best of my knowledge, only study of Roswell Spanish in particular is Wilson's (2013) analysis of narrative structure. Wilson conducted sociolinguistic interviews in Roswell and utilized Labov & Waletzky's (1967) framework of narrative structure to analyze variation in discourse structure between Spanish and English as well as among individual speakers. The author concluded that, among Roswell speakers, there is more variability in narrative structure than was previously supposed, and that Labov & Waletzky's narrative paradigm, which involves a specific set of structures that occur in a specific order, functions differently in Spanish than in English.

The use of Spanish in the Mid-Atlantic South in general has received very little attention, and there are no studies to my knowledge that have specifically addressed subject pronoun usage of Spanish in Roswell, Georgia. Therefore, a primary goal of the present research is to contribute to the sociolinguistic literature and to compare subject expression among Spanish speakers living in Roswell to that of speakers in other areas of the US.

Through a quantitative analysis of pronoun usage using conversational data from Mexican speakers in Roswell, I show that subject expression in Roswell Spanish is greatly influenced by language internal factors such as person/number, lexical content, reflexivity, and coreferentiality index, among others. I also show that Spanish in Roswell exhibits a substantially lower pronoun rate than more established varieties of US Spanish and is more similar to monolingual Mexican varieties and more incipient US varieties. However, divergences are shown in the ranking of constraints on pronoun use, specifically with regard to person/number and coreferentiality index. Finally, regarding English contact, I argue that the similarity in pronoun rate with monolingual varieties and the insignificance of social factors provide no evidence for English influence on pronominal use in Roswell.

## CHAPTER 3

### METHODOLOGY

#### 3.1 DATA AND PARTICIPANTS

Using a variationist sociolinguistic approach, I examined subject expression among 11 native Spanish speakers, the majority of whom are Mexican immigrants (one speaker is Cuban) born in various regions of Mexico (i.e. Mexico City, Veracruz, Michoacán), who are currently residents of Roswell, GA and whose average length of residency in the US is 14 years. The sample includes six females and five males ranging in age from 20 to 54 years old and whose age of arrival to the US ranges from 18 to 38 years old. In addition, their levels of education range from primary school to university, and the majority of the participants work in hotel housekeeping (see Table 5 for sociodemographic information categorized by speaker).

Sociolinguistic interviews, which typically involve “a loosely structured set of topics preselected by the interviewer to mirror current, local and/or individual interests, minimally including childhood games, customs, folklore, recipes and narratives of personal experience” (Poplack 1993:261), were conducted in Roswell (Wilson 2013) and were later transcribed to allow for coding and analysis of pronoun usage and the independent linguistic variables that tend to influence such usage. The nine independent variables included in this study are person and number of the verb, coreferentiality index (same vs. switch reference), clause type, specificity (specific or non-specific reference), tense-mood-aspect of the verb (TMA), reflexivity (reflexive or nonreflexive use of the verb), previous mention (whether the pronoun of the previous verb is null or overt), lexical content of the verb (verbal semantics), and lexical frequency. In addition,

four independent social variables are examined including gender, age, length of residency in the US, and age of arrival to the US. These variables will be described and illustrated below in greater detail.

With regard to the participants' degree of exposure to English, the speakers indicate in the interviews that they generally do not have much interaction with Americans and that they spend most of their time with other Latinos, including their family and friends. Below are some excerpts from the interviews, indicating their answers to a question regarding who they typically interact with:

- (6) “...bueno, más que nada con, más, más con mi familia y a veces, las fiestas como amigos, parientes...”  
 ‘...well, more than anything with, more, more with my family and sometimes, the parties like, friends, relatives...’ [M43]
- (7) “...pues normalmente con mi familia.”  
 ‘...well normally with my family.’ [F37]
- (8) “...pues con latinos.”  
 ‘well with Latinos.’ [F51]
- (9) “...está muy limitado nos, nuestra vida social. Por lo mismo de que, este, estamos tan metidos en el trabajo, conocemos poco anglosajones. Lamentablemente aquí en la ciudad, conocemos menos americanos, gente americana.”  
 ‘...it’s very limited, our social life. Therefore that, umm, we are so involved with work, we know few English speaking people. Regrettably, here in the city, we know less Americans, American people.’ [F50]

### 3.2 THE ENVELOPE OF VARIATION

A total of approximately 2,421 finite clauses were identified in the transcripts, of which 1,307 were excluded from the present study, leaving a total of 1,114 tokens for analysis. These exclusions were made because they were considered to be outside of what is called the *envelope of variation* (Otheguy et al. 2007). With regard to the present study, the envelope of variation includes environments of finite clauses in which variation between overt and null subject pronouns is possible. In other words, any context of null subject pronoun expression in which an overt subject pronoun could also occur belongs inside the envelope of variation and, therefore, is included in the present analysis. Likewise, the appearance of an overt subject pronoun in a context in which a null pronoun could alternatively occur would fit inside the envelope of variation and is included in the study. However, clauses in which either the null or overt pronoun is obligatory or nearly obligatory have been eliminated since they are generally not considered to exhibit variation and consist of the following contexts: verbs within subject headed relative clauses (*La mujer que Ø vino ayer* ‘The woman that came yesterday,’ clauses with full noun phrases (*Las chicas Ø estudiaron mucho* ‘The girls studied a lot’), verbs, such as *haber* or *ser*, when used existentially (*Ø Hay/había muchas personas* ‘There are/were many people’; *Ø Es que no salgo mucho* ‘It’s just that I don’t go out much’), verbs such as *hacer* when making reference to chronological periods (*Ø Hace cinco años* ‘Five years ago’), verbs with inanimate referents (*Ø es bonita*, making reference to a city, ‘It’s nice’), as well as set phrases where an overt or null subject pronoun is preferred (*¿Qué sé yo?* ‘I don’t know’ [literally ‘What do I know?’]; *¿me entiendes Ø* ‘Do you know what I mean?’ [literally ‘Do you understand me?’]). These examples fall outside the envelope of variation because they are essentially categorical with regard to subject pronoun expression; in other words, Spanish speakers would rarely (if ever) alternate

between an overt and null pronoun in these cases. For example, it would be unexpected to find *La mujer que ella vino ayer* or *Las chicas ellas estudiaron mucho* or *Él hace cinco años* or *Ella es bonita* (referring to a city) or *¿Qué sé Ø?*<sup>2</sup>

Below is a passage from one of the participants in which I contextually illustrate the verbs that I included in the study, which are marked in **bold**, and also those which I excluded from the study, which are *italicized*:

(10) **Yo creo** que donde (11) **yo soy**, lo más **agradezco** (12) *es* que este país me (13) *hizo* a madurar. Entonces de allí, (14) **empecé** a trabajar, a trabajar en esta compañía de construcción. Posteriormente por lo mismo la, la habilidad que me (15) *dio* Dios de aprender tan rápido que, este, después de tres años que (16) **trabajé** para esta compañía, pues (17) **di** las gracias. (18) **Empecé** a trabajar por mi cuenta, (19) **puse** mi compañía de construcción, me (20) *fue* muy bien como por cinco años. El trabajar me (21) *fue* muy muy bien, y de ahí (22) *fue* adonde (23) **empezamos** a lograr lo que, lo que (24) **tenemos** ahorita. También lo que, con el trabajo (25) *tiene* mi, mi esposa. Y de ahí (26) **yo te puedo** decir que (27) *hace* año y medio (28) **yo dejé** completamente ya la construcción. Este, lo único que allá (29) **yo, le digo** *es, es, esta* (30) *es* broma, ¿no?, pero le (31) **digo** a mi esposa que a veces no me (32) *gusta* estar con ella porque aquí me (33) **mandan**... (Speaker 013M, Mexican, Age 38, Years in the US: 13)

‘**I think** that where **I am, I am** grateful the most *is* that this country *made* me mature. So from there, **I started** to work, to work at this construction company. Later, for that reason, the, the ability that God *gave* me to learn so quickly that, umm, after three years of working for this company, well **I was** thankful. **I started** working on my own, **I established** my construction company, it *went* very well for like five years. Working *went* very very well and from there *was* where **we started** to achieve what, what **we have** right now. Also, what, with the job that my, my wife *has*. And from there **I can** tell you that a year and a half ago **I** completely **quit** construction, umm, the only thing that there, I, **I tell** her *it’s it’s*, this *is* a joke you know, but **I tell** my wife that sometimes I don’t *like* to be with her because there, **they order** me...’

The italicized verbs in the above passage are outside the envelope of variation for the reasons previously mentioned, and the bolded verbs are inside the envelope because they all have personal referents in contexts of potential null and overt variability (i.e. #10 **Yo creo**, #24

<sup>2</sup> Additionally, in contact situations one might expect to see overt pronouns used where they aren’t permitted in Spanish, but are permitted in English, such as with weather expressions (i.e. \**Ello está lloviendo*, ‘it’s raining’). However, I have found no such examples in my data. Moreover, I strongly suspect that this phenomenon does not occur since, to the best of my knowledge, it has not been observed in previous research on Spanish-English contact varieties.

*[nosotros] tenemos, #33 [ellos] mandan*) and have therefore been coded and analyzed in the present study.

### 3.3 CODING OF DATA

Regarding the coding process, I first extracted all finite verbs within the envelope of variation and then coded whether the verb appeared with a null or overt subject pronoun in order to account for the main dependent variable of the study. Next, I coded for the nine independent linguistic variables as well as for the four relevant social variables. In order to determine the significance of each variable as well as the constraints that condition SP use, I carried out a mixed effects multivariate analysis using Rbrul (Johnson 2009).

#### 3.3.1 LINGUISTIC VARIABLES

##### *Person and number of verb*

In most Spanish varieties, both person and number of the verb have been shown to influence the use of subject pronouns. Specifically, first person singular verbs have been shown to favor overt pronoun usage most frequently, with the exception of Spanish in Madrid, where lower frequencies are found (Cifuentes 1980; Enríquez 1984; Cameron 1992; Silva-Corvalán 1994; Flores-Ferrán 2002, among others). In fact, both Barrenechea and Alonso (1977) and Silva-Corvalán (1994) found that all singular forms for all persons have a greater probability to appear with overt pronouns when compared to plural forms, which have shown a general disfavoring of overt pronouns (Enríquez 1984; Cameron 1992; Silva-Corvalán 1994; Flores-Ferrán 2002). Further, Otheguy et al. (2007) reveal regional differences with regard to person, finding that Caribbean Spanish speaking newcomers to NYC more strongly favored overts with second person singular verbs than they did with third person singular verbs. Conversely, South

American newcomers preferred overts more frequently with third person singular verbs. Below are two examples from my data, which include verbs I have coded for person/number:

(34) Siempre **he** trabajado en el hotel.

‘I have always worked in the hotel.’ [M43Mex] (*Person and number* was coded as First Person Singular)

(35) ...y ellos nos **están** apoyando bastantísimo.

‘...and they are supporting us a lot.’ [M54] (*Person and number* was coded as Third Person Plural)

*Coreferentiality Index (switch vs. same reference)*

This variable, which considers subject continuity--that is, switch vs. same reference from one subject to another--has been shown to strongly influence the manifestation of subject expression cross-dialectically for all persons (Silva-Corvalán 1982; Enríquez, 1984; Bentivoglio 1987; Cameron 1994, 1995; Bayley & Pease-Alvarez 1997; Travis 2005). In contexts where there is a switch in subject referent, the subject pronoun is often overt; on the other hand, when there is no switch in subject referent, there tends to be a favoring of null subject pronouns. However, this pattern is by no means categorical. For example, both Flores-Ferrán (2002) and Otheguy et al. (2007) found that NYC-born Spanish speakers exhibit an increased use of overt pronouns in contexts where the subjects are the same, demonstrating their loss of sensitivity to null pronoun usage in such contexts. Similarly, such non-canonical patterns are found in the present study, as illustrated in examples (36) and (37) below:

(36) ...una **amiga** mía se casó con un americano y **ella** está muy feliz

‘...a friend of mine married an American and she is very happy’ [M23Mex]



- (37) ... **yo** te puedo decir que hace un año y medio **yo** dejé ya completamente a la construcción  
 ‘...I can tell you that a year and a half ago I totally quit construction’ [M38Mex]

Furthermore, for the present study, I have coded this variable using a more extensive method, taking into account not only subjects, but also the coreferentiality between subjects and preceding objects as well as the speech of the interviewer according to their subject reference in the immediate preceding discourse, which in most cases was a question from the interviewer. Specifically, in switch reference contexts where a subject pronoun was coreferential with the preceding object (direct object, indirect object, object of preposition, etc.), I considered such coreference a potential influencing factor for the use of an overt or null pronoun. This type of context is illustrated below:

- (38) A mí me **pusieron** a estudiar derecho, derecho, como abogados, adonde yo **conocí** a la que ahora es mi esposa.

‘They required me to study law, law, like lawyers, where I met who is now my wife.’

[M38Mex]

- (39) ...nos **preguntan** cómo está aquí, pues, **decimos** que está bien...

‘...they ask us how it is here, well, we say that it’s good...’ [F51Mex]

In the above examples, even though there is a switch in reference between the subjects (*ellos* switches to *yo* in (38) and *ellos* switches to *nosotros* in (39)), there is coreference between the underlined objects and the subject of each clause that follows. Therefore, I did not code the tokens *conocí* and *decimos* above as simple cases of switch reference, but rather cases of a switch with subject and coreference (no switch) with object. This more general method of

identifying switch reference has been adopted based on previous studies of subject expression (i.e. Silva-Corvalán 1994; Bayley and Pease-Álvarez 1996; Orozco & Guy 2008).

Moreover, following Otheguy & Zentella (2012), I took the interviewer's speech into account when coding for coreferentiality index. Specifically, I considered the previous reference made by the interviewer when looking at the respondent's first token. For example, I coded same reference for any case in which the interviewer asked a question whose final referent was the same as the respondent's first referent. Below I present an example of this context:

(40) Interviewer: ¿Me puedes hablar un poquito sobre cualquier educación que **has** tenido en tu vida?

‘Can you talk to me about any education you have had in your life?’

Respondent: Solamente **fui** a la primaria.

‘I only went to primary school.’ [F37Mex]

Since there is no switch in reference between the respondent's self-reference *fui* and the interviewer's reference *has* (they both refer to the respondent), this is perhaps one of the reasons for the use of a null pronoun by the respondent.

#### *Clause Type*

Another linguistic factor that can condition subject expression is the type of clause in which the pronominal subject appears. For example, Morales (1997) found that overt subject pronouns are more likely to occur in subordinate clauses—particularly object relative clauses—than they are in main clauses. The three clause categories I coded for in the present study are Main, Subordinate, and Coordinate in order to determine which of these types of clauses might have an effect on pronoun usage in the Roswell data. A fourth category, Indeterminate Clause,

was included for cases in which the structure of the surrounding discourse could not be determined, thus making it impossible to identify the clause type in which the verb appeared.

The following examples illustrate each clause type:

- (41) Acá **tengo** como quince años, creo.

‘I’ve been here about fifteen years, I think.’ [M43Mex] (Clause type was coded as Main Clause)

- (42) Yo **tengo** mi trabajo aquí y me acostumbré...

‘I have my job here, and I got used to...’ [F51Mex] (Clause type was coded as Main Clause)

- (43) He visto que **ayudan** bastante.

‘I’ve seen that they help quite a bit.’ [M43Mex] (Clause type was coded as Subordinate Clause).

- (44) Agradezco mucho a ella porque en esa situación que yo **estaba** pasando...

‘I am very grateful for her because in that situation that I was going through...’  
[F32Mex] (Clause type was coded as Subordinate Clause).

- (45) ...suben a los novios en unas sillas y **empiezan** a hacer como...

‘...they put the bride and groom up on some chairs and they start to do like...’ [F50Mex]  
(Clause type was coded as Coordinate Clause).

- (46) Se llama José y él **tiene** 24 años.

‘His name is José, and he is 24 years old.’ [F43Mex] (Clause type was coded as Coordinate Clause).

*Tense-Mood-Aspect (TMA)*

The tense, mood, or aspect (TMA) of a verb has also been shown to have a significant effect on subject expression, that is, certain TMAs favor overt pronoun use while other TMAs disfavor such use. For instance, Silva Corvalán (1982) found that imperfects and conditionals favor overt pronouns while there is a lesser probability of overt pronouns appearing with presents and preterits. Likewise, other researchers have found correlations between such TMAs and overts (Cameron 1994; Bayley & Pease-Alvarez 1997; Travis 2007). To explain such correlations, it has been proposed that imperfects and conditionals favor overt pronouns due to their potential ambiguity, as the first and third persons of both verb forms are the same, and the use of overt pronouns would therefore serve to disambiguate the referents of such forms (Hochberg 1986). However, other studies have found no such correlation (Bentivoglio 1987; Enríquez 1984; Ranson 1991) which raises the question of whether overt pronouns do in fact serve a disambiguating function or if there is perhaps a more prominent function with regard to TMA. Examples 47-52 below demonstrate the use of different TMAs with their respective coding labels:

- (47) Sí, porque **somos** una familia aunque mis hijos ...  
 ‘Yes, because we are a family even though my children...’ [F37Mex] (TMA is coded as Present Indicative)
- (48) **Estuvimos** luego a su iglesia ...  
 ‘We then went to her church...’ [F32Mex] (TMA is coded as Preterite)
- (49) Él **estaba** aquí...  
 ‘He was here...’ [F32Mex] (TMA is coded as Imperfect Indicative)

(50) **He asistido** a iglesias cristianas...

‘I’ve attended Christian churches...’ [F37Mex] (TMA is coded as Present Perfect)

(51) ...no te **sabría** decir bien ...

‘...I wouldn’t know how to tell you...’ [F32Mex] (TMA is coded as Conditional)

(52) Lo más importante es que las **involucren** en, en algo positivo.

‘The most important thing is that they get them involved in, in something positive.’

[M43Mex] (TMA is coded as Present Subjunctive).

### *Lexical Content*

Several researchers have found that the lexical content of verbs can also determine how a subject pronoun is manifested (Bentivoglio 1987; Enríquez 1984; Silva-Corvalán 1994; Otheguy et al. 2007; Travis 2007). In general, it has been noted that verbs of psychological/mental activity, verbs of communication, and copulas tend to be expressed with overt pronouns, with psychological verbs showing the highest probability (e.g. *creer, pensar, decir, hablar, ser, estar*). On the contrary, motion verbs tend to disfavor overts (Bentivoglio 1987; Enríquez 1984; 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). I have employed each of the four semantic categories stated above in the present study, specifically adopting the categorization used by Travis (2007). Furthermore, I propose an additional category that includes verbs of perception such as *ver, escuchar, and notar* as these types of verbs appear often in the data, and they frequently appear with overt pronouns, which I suspect to be significant. For the remaining verbs that do not fit into these specific categories, I have coded as “other,” also in accordance with Travis (2007).

- (53) ...yo **pienso** que, para mí, yo he escuchado mucho, que les gusta Roswell, que a mí me gusta, está bonito aquí.  
 ‘I think that, for me, I’ve heard a lot, that they like Roswell, that, I like it, it’s nice here.’  
 [F51Mex] (Coded as a psychological verb)
- (54) Después de Chalco, me **vine** para acá...  
 ‘After Chalco, I came here...’ [M23Mex] (Coded as a motion verb)
- (55) **Soy** el más chico...  
 ‘I’m the youngest...’ [M38] (Coded as a copula)
- (56) ... no ha cambiado mucho, ni ha progresado mucho porque, **hablan** de problemas económicos.  
 ‘...it hasn’t changed much, neither has it progressed much because they talk about Economic problems.’ [F50Mex] (Coded as a verb of communication)
- (57) Yo lo **veo** bien, o sea, que están bien orientados más que nada.  
 ‘I see it as something good, that is, that they are well oriented more than anything.’  
 [M43Mex] (Coded as a verb of perception).
- (58) No **puedo** contar mucho, pero....  
 ‘I can’t tell much, but...’ [M23Mex] (Coded as Other)

### *Specificity*

Cameron’s (1992) analysis of the Spanish varieties of San Juan and Madrid indicates a significant influence of specificity of subject referent on pronoun usage. In other words, the use of a specific referent, as opposed to a more general or nonspecific referent tends to condition whether or not a pronoun is overt or null. The following two examples illustrate this distinction between specific and non-specific with regard to the subject *tú*:

(59) ¿Tú **trabajas** en esa compañía?

‘Do you work in that company?’ (Flores-Ferrán 2007b:638)

(60) Si tú **reclamas** el dinero a tiempo, no tienes que pagar penalidades.

‘If you (one) claim the money on time, (you (one)) don’t have to pay penalties.’ (Flores-Ferrán 2007b:638)

In (59), *tú* is used by the speaker to refer specifically to the interlocutor or listener; however, in (60) the subject referent *tú*, depending on the context, could potentially be referring to a nonspecific entity (‘one’).

In Cameron’s (1992) study, dialectal differences were found between speakers in Madrid and speakers in San Juan with regard to the interaction of specificity and subject expression. Specifically, it was the non-specific subject referents that favored overt pronouns in San Juan while the specific referents favored overts in Madrid. However, specificity is not found to be significant in all dialects; for instance, Otheguy et al. (2007) found that specificity was only a conditioning factor for pronoun use among NYC born mainlanders (Mexicans, Colombians, Ecuadorians) and had no influence on the Spanish of Caribbeans (both NYC born and newcomers) nor for Mainland newcomers.

Another common subject referent that can be used in both a specific and non-specific manner is *ellos*. Below are two examples from my data that illustrate these two forms:

(61) A veces sí **pueden** ir conmigo...

‘Sometimes they can go with me...’ [M43Mex] (Specificity is coded as Specific)

(62) Ellos lo **hacen** sentir uno comfortable allí.

‘They make one feel comfortable there.’ [F37Mex] (Coded as Non-specific)

In (61) the speaker is referring specifically to his children and not simply to “them” in a general sense. In contrast, this general “they” is indeed referred to in (62), where the speaker is talking about the church personnel in general. Another example, this time using the non-specific subject *tú* mentioned above, is presented below in which the speaker is not referring specifically to the interlocutor, but to the “general you” or “one.”

(63) Entonces, ya no **te sientes** con la seguridad...

‘So, you no longer feel the security...’ [M43Mex] (Coded as Non-specific)

### *Reflexivity*

Whether a verb is used reflexively or non-reflexively is also a conditioning factor of subject expression that has been shown in the literature (Otheguy et al. 2007; Carvalho & Child 2011). For example, Carvalho & Child (2011) found that verbs that contain a reflexive pronoun disfavor overt pronouns whereas verbs without a reflexive pronoun neither favor nor disfavor overts. According to Carvalho & Child, “this is probably due to the added referential information contained within the reflexive pronoun that might make the use of the subject personal pronoun seem redundant” (2011:20). The following segments from my data illustrate the factor of reflexivity in the present study:

(64) ...no **están** permitidos que **se casen**...

‘...They aren’t permitted to get married...’ [M43Mex] (Coded as Non-reflexive for *están* and Reflexive for *se casen*).

(65) ...él no supo que yo **me fui**...

‘...he didn’t know that I went...’ [M54] (Coded as Reflexive)



*Previous mention*

This variable, which serves to determine a possible “priming” effect, takes into account whether the previous verb’s subject pronoun is overt or null. Priming is defined by Travis as “the process whereby the use of a certain structure in one utterance functions as a prime on a subsequent utterance, such that that same structure is repeated” (2007:101). In terms of subject expression, an effect of this type has been observed in various studies in which the use of an overt pronoun in a given clause has been found to be conditioned by the use of an overt subject pronoun in a preceding clause. Likewise, it has been shown that null subject pronouns display the same effect, whereby the use of a null pronoun primes the use of a subsequent null pronoun (Cameron 1994; Flores-Ferrán 2002; Travis 2007). In order to determine if a priming effect occurs in the data of the present study, I have coded each token in relation to the subject expression of the previous token as is seen in example (66) below.

(66) **Yo** he visto que **ayudan** bastante...

‘I’ve seen that they help quite a bit...’ [M43Mex] (Coded as Overt since the verb *he*, which precedes the verb *ayudan*, appears with the overt pronoun *yo*).

However, following Travis (2007), I have excluded the first token of each interview since it has no preceding clause, which makes it impossible to determine whether a priming effect is involved. I have also excluded first mentions in discourse (i.e. after a question from the interviewer) since the interactional nature of this context often causes an interruption in the cohesiveness of the interviewee’s speech, making it difficult to determine a priming effect. Below I provide two example passages from my data that illustrate potential priming effect contexts for both overt and null subject pronouns.

(67) **Tengo** tres niños y **me dedico** a mis niños en la tarde cuando no **trabajo**, **me dedico** a ellos.

‘I have three children and I devote myself to my children in the afternoon when I don’t work, I devote myself to them.’ [F37Mex] (Coded as null for all verbs)

(68) ...**yo quería** estudiar algo muy diferente. **Yo quería** ser piloto, piloto aviador. Entonces **yo hice** exámenes para La Marina...

‘I wanted to study something very different. I wanted to be a pilot, airplane pilot. So I took exams for the Marines...’ [M38Mex] (Coded as overt for all verbs)

In (67), it is possible that the null subject pronoun with *tengo* primes the subsequent nulls with *me dedico*, *trabajo*, and *me dedico*. In (68), a series of overts appear, whereby the subject pronoun *yo* produced with *queria* potentially conditions the subsequent overts in the clauses that follow. In fact, this is precisely what Travis found for first-person singular subjects, which she describes as a “yo-yo effect,” explaining that “one *yo* leads to another *yo*, and one implicit first-person singular subject leads to another implicit subject” (2005:330).

### *Lexical Frequency*

A very understudied but intriguing variable to consider in relation to subject expression is lexical frequency. Taking into account the frequency of verbs used in speech, Erker and Guy (2012) measured this variable discretely, considering frequent versus infrequent forms used by speakers and their effect on SP use. Rather than observing an independent and direct influence on SP occurrence, the researchers found that lexical frequency more directly interacts with other independent variables<sup>3</sup> (e.g. TMA, person/number), revealing that high-frequency forms tend to strengthen certain constraints on SP use whereas low-frequency forms tend to weaken or negate

<sup>3</sup> Erker and Guy refer to this effect as “non-orthogonal”. The typical assumption in statistical analyses is the independent variables are in fact independent from, or orthogonal to, one another.

their effects. For instance, they find that first and third person verb forms of higher frequency are associated with a decrease of overt SPs whereas second person high frequency forms are associated with an increase of overt usage. Following Erker & Guy (2012), I first calculated the raw frequency of each verb's non-finite (infinitive) form, in other words, the number of times any given verb form appeared in the data. This raw frequency was then log transformed. Each token was then coded using the log frequency of its respective infinitive. The log frequencies ranged from 0 (i.e. tokens with a raw frequency of only 1), which included verbs such as *acabar*, *cambiar*, *deber*, *limpiar*, *pedir*, *terminar*, to 4.852, represented by the verb *tener* which occurred a total of 127 times in the data.

#### *Predictions concerning linguistic variables*

Based on the findings of previous studies regarding correlations between independent linguistic variables and Spanish subject expression, I predict the following concerning SP use in Roswell:

- (i) Singular verb forms, especially those in the first person, will favor overt SPs while plural forms will generally show a disfavoring of overts.
- (ii) Contexts in which there is a switch in subject referent will favor overt SPs while same reference contexts will favor null SPs. However, these contextual effects may be weaker than those observed in monolingual and less established contact varieties (for example, Newcomers in the Otheguy et al.'s data) since I have noted an increase of overts in same reference contexts in the data. Additionally, contexts in which there is a switch in subject but no switch with the preceding object referent will disfavor overt SPs.
- (iii) Verbs within subordinate clauses will most highly favor overt SPs when compared to verbs in main and coordinate clauses.

- (iv) Imperfect and conditional verb forms will favor overts while present and preterit forms disfavor them.
- (v) Psychological and perception verbs, copulas, and verbs of communication will show a general favoring of overts, with psychological verbs showing the highest probability, while motion verbs will generally disfavor overts.
- (vi) Since effects for specificity of the referent have been somewhat inconsistent in previous studies, it is difficult to predict its influence on SP usage in Roswell. However, from a general impressionistic analysis of this variable in the data, I predict that non-specific referents will disfavor overt SPs while specific referents favor overts.
- (vii) Verbs that are used reflexively will disfavor overts while non-reflexive verbs will generally favor them.
- (viii) Preceding overt SPs will favor overts while preceding null SPs will favor nulls.
- (ix) The frequency of a verb in the data will have no significant influence on the probability of occurrence of a null or overt SP.

### 3.3.2 SOCIAL VARIABLES

In addition, I coded for the following four independent social variables in order to determine whether such factors influence subject pronoun use in Roswell: gender, age, length of residency in the US, and age of arrival to the US. These social variables, among others such as level of education and social class, have been examined in previous studies in order to determine how they impact subject pronoun use in different dialects (Enríquez 1984; Bentivoglio 1987; Cameron 1992; Otheguy et al. 2007; Carvalho & Child 2011; Shin & Otheguy 2013).

### *Gender*

The speaker's gender is not typically shown to be significant in relation to SP usage (Cameron 1992; Otheguy et al. 2007); however, some authors have in fact found that females and males differ substantially in their SP frequencies. For instance, in their study of Uruguayan Spanish in contact with Portuguese, Carvalho & Child (2011) found that females favored the use of overt SPs whereas males favored null SPs. Similarly, Shin & Otheguy (2013) observed a higher rate of overt SPs among women speakers in New York, a finding they attribute to women's more extensive contact with US born bilinguals, who generally exhibit relatively high pronoun rates.

### *Age*

Age has also been shown to affect the variable use of SPs. In addition to finding gender to be a significant factor, Carvalho & Child (2011) also found that there is a favoring of null SPs among younger speakers (ages 16-29) while second and third generation speakers favor overt SPs. However, due to a small sample size, the authors are unable to generalize this finding to the dialect as a whole. On the contrary, Flores-Ferrán (2002) found that older speakers (50+) 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. Otheguy et al. (2007), however, did not find age to be a significant factor for SP usage.

### *Investigating English Contact: Length of Residency and Age of Arrival*

Two important factors to consider when looking at potential influence of English on Spanish language change are the length of residency of the speakers, which normally suggests greater exposure to English, and the age of the speakers when they first immigrated to the US, which indicates speakers' susceptibility to change. For instance, Otheguy et al. (2007) found that

a longer length of residency in New York correlated with higher frequencies of overt SPs due to the fact that speakers who had lived there longer had much more exposure to English than those who were recent arrivals. Specifically, they found that New York born speakers had an overall pronoun rate of 38%, while recent arrivals showed a rate of only 30% (Otheguy et al. 2007:783), a difference they attribute to English contact. In addition, the authors find that younger arrivals, who are more susceptible to English influence, exhibit higher pronoun rates than older arrivals, who tend to use less English and whose Spanish is more proficient, thereby making them less susceptible to influence from English.

Using the above findings as a base, I predict that the speakers in the present study who have lived in the US for longer periods of time will have higher frequencies of overt SPs than those with shorter lengths of residency and that those who arrived at younger ages will exhibit higher overt frequencies than older arrivals. However, due to the inconsistent findings on age and gender and given the fact that most studies on Spanish subject expression do not find such factors to be significant, I do not expect either variable to be a conditioning factor of SP use in Roswell.

## CHAPTER 4

### RESULTS

#### 4.1 INDIVIDUAL AND OVERALL PRONOUN RATES

As noted above, one Cuban speaker was included among the speakers from Mexico in the sample, which was done strictly for comparative purposes in terms of individual pronoun rates. Table 5 below illustrates the individual rates from highest to lowest for all eleven speakers, along with their respective speaker codes – including their gender, age, and country of origin – as well as their ages of arrival to the US, length of residency in the US, and their cities and/or states of origin.

Table 5: Individual PN rates and sociodemographic information

<b>Ranking</b>	<b>Speaker</b>	<b>City/State of Origin</b>	<b>Length of Residency</b>	<b>Age of Arrival</b>	<b>PN Rate</b>
1	M54Mex	Tampico, Tamaulipas	16	38	40%
2	M38Mex	Mexico City, D.F.	13	25	29%
3	F51Mex	Tamaulipas (state)	19	32	29%
4	F37Mex	Altamira, Tamaulipas	19	18	21%
5	M43Mex	Pachuca, Hidalgo	15	28	19%
6	F43Mex	Veracruz (state)	14	29	17%
7	M44Mex	Veracruz (state)	21	23	16%
8	M23Mex	Texcoco, Mexico	10	13	16%
9	F32Mex	Michoacán (state)	7	25	14%
10	F20Cuba	Havana	2	18	12%
11	F50Mex	Mexico City, D.F.	22	28	7%

Here I show the variation in rates among speakers in Roswell. While individual pronoun frequencies differ and some are closer to the general rates of mid/high range varieties of Spanish, the overall pronoun rate when considering the speakers as a whole is the more relevant factor for comparing Spanish in Roswell to other varieties. Nonetheless, it is interesting to note that the Cuban speaker only used an overt pronoun 12% of the time, which is surprising considering that in general, Caribbeans typically have the highest rates out of all Spanish varieties. While this



seems to be anomalous, it is important to recognize that individual speakers do not always exhibit rates that are similar to overall rates. This idea reflects the danger of using a single speaker to represent an entire speech community. Therefore, when I refer to the pronoun rate in Roswell, this does not mean that every speaker exhibits this rate but that it is the average rate of all speakers together.

In determining the overall pronoun rate (as well as for all subsequent analyses in the study), I excluded the data from the Cuban speaker in order to have a relatively homogenous sample in terms of national origin. As seen in Table 6, out of a total of 1,005 verbs, 214 occurred with an overt SP and 791 occurred with a null SP, indicating an overall overt pronoun rate of 21% for Mexicans in Roswell.

Table 6: Overall rates of subject pronoun expression in Roswell

	<b>N Verbs</b>	<b>%</b>
Verbs with overt subject pronouns	214	21
Verbs with null subject pronouns	791	79
Total	1,005	100
Number of speakers = 10		

Comparing this result to other studies<sup>4</sup> (see Table 7 below), the rate of PN expression in the Roswell data is only marginally higher than that of Mexican newcomers in New York (19%) and Mexicans in North Carolina (17%), marginally lower than that of Mexican newcomers in New

---

<sup>4</sup> Sources: Yucatan (Michnowicz, forthcoming) Mexico City (Lastra & Butragueño, forthcoming) Valladolid (Solomon 1999) Puente Genil (Ranson 1991) Buenos Aires (Barrenechea & Alonso 1977) Santiago, Chile (Cifuentes 1980), San Juan (Cameron 1994), Santiago, Dominican Republic (Olloqui de Montenegro 1987), New York Newcomers (Otheguy et al. 2007) Los Angeles (Silva-Corvalan 1982) New York Dominicans (Flores & Toro 2000), New York Born/Raised (Otheguy et al. 2007) North Carolina (McKnight 2013), New Jersey Mexicans (Flores-Ferrán 2007a)

Jersey, identical to that of monolingual Mexicans in Valladolid and Mexico City (21%), and significantly lower than the rate of Mexicans in Los Angeles (43%). Therefore, in terms of overall pronoun rate, the Spanish of Mexican speakers in Roswell is more like monolingual and incipient contact varieties and less like enduring contact varieties.

Table 7: Comparison of PN rates between Roswell Spanish and other varieties

<b>Variety (monolingual)</b>	<b>PN Rate</b>	<b>Variety (contact)</b>	<b>PN Rate</b>
Yucatan, Mexico	16%	Los Angeles Mexicans	43%
Mexico City	21%	New York Dominicans	48%
Valladolid, Mexico	21%	New York born/raised (Mexican, S. American, Caribbean)	38%
Puente Genil, Spain	24%	North Carolina Mexicans	17%
Buenos Aires, Argentina	36%	<b>Roswell Mexicans</b>	<b>21%</b>
Santiago, Chile	38%	New Jersey Mexicans	24%
San Juan, Puerto Rico	45%		
Santiago, Dominican Republic	39%		
New York Newcomers (Mexican, S. American, Caribbean)	30%		

#### 4.2 FACTORS CONDITIONING THE EXPRESSION OF SUBJECT PRONOUNS

For the multivariate analysis carried out in Rbrul, I included all aforementioned independent linguistic and social variables. Regarding the lexical frequency variable, I included the log frequency as a continuous predictor. Also, with regard to the TMA variable, I excluded

all uses of the synthetic future (e.g. *comeré*), conditional, and imperfect subjunctive due to the very low number of tokens found in the data for these categories and recoded the very low number of periphrastic future tokens (e.g. *voy a comer*), adding them to the present indicative category. I also simplified the coreferentiality index variable, excluding all verbs that were the first tokens in the discourse (e.g. after a question from the interviewer). Additionally, I included the speaker as a random intercept.

The linguistic variables selected as significant in the multivariate analysis include person/number, lexical content, reflexivity, coreferentiality index, previous expression and clause type. The three linguistic factors that did not have a significant effect on SP usage were TMA, specificity, and lexical frequency. Regarding the social variables of gender, age, length of residency and age of arrival, none were selected as significant to subject expression. The resulting variable hierarchy is presented below, showing the ranking of significant factors. Person/number, which is ranked first, has the most powerful effect on subject expression while clause type, which is ranked sixth, has the weakest effect. This ranking was determined by calculating the range of probability weights for the constraints within each variable. Table 9 shows the specific constraints on SP usage, which are also ranked hierarchically. In the next section, I discuss the significant variables and constraints in greater detail.

Table 8: Hierarchy of variables related to subject expression

N verbs = 698

\*p &lt; 0.05

\*\*p &lt; 0.01

<b>Rank</b>	<b>Variable</b>
1 <sup>st</sup>	Person/Number**
2 <sup>nd</sup>	Lexical Content**
3 <sup>rd</sup>	Reflexivity**
4 <sup>th</sup>	Coreferentiality Index**
5 <sup>th</sup>	Previous Expression**
6 <sup>th</sup>	Clause Type*

Table 9: Hierarchy of constraints on SP use

Factor Group	Probability Weight	% Overt	N
<b>Person/Number</b>			
3s	.78	29	65
1s	.66	27	318
2s	.57	18	56
3pl	.49	12	169
1pl	.10	2	90
RANGE 68			
<b>Lexical Content</b>			
Perception	.69	36	14
Psychological	.66	33	105
Copula	.59	23	184
Motion	.44	14	95
Other	.41	14	271
Communication	.23	10	29
RANGE 46			
<b>Reflexivity</b>			
Non-reflexive	.66	21	623
Reflexive	.34	8	75
RANGE 32			
<b>Coreferentiality Index</b>			
Switch with subject, coreferential with object	.67	33	27
Switch reference	.46	23	279
Same reference	.37	16	392
RANGE 30			
<b>Previous Expression</b>			
Overt	.64	31	150
Null	.37	16	548
RANGE 27			
<b>Clause Type</b>			
Subordinate	.62	24	177
Main	.58	23	301
Coordinate	.42	11	201
Indeterminate	.39	10	19
RANGE 23			
Speaker (random) Std. Dev. .47			

#### 4.2.1 PERSON/NUMBER

As predicted, the singular verb forms show a favoring of overt pronouns whereas the plural forms show a disfavoring of overts. Specifically, third person singular forms favor overts the most with a probability weight of .78 and first person plural forms highly disfavor overts with a probability weight of .10. In fact, only 2% of first person plural forms in the data occur with an overt SP. This general finding with regard to singular and plural forms is consistent with previous studies (Cifuentes 1980; Enríquez 1984; Cameron 1992; Silva-Corvalán 1994; Flores-Ferrán 2002). However, while most studies have found that first person singular verbs favor overts most frequently, it is the third person singular forms, as mentioned above, that most frequently appear with overt SPs in the present study. More specifically, this finding shows a divergence from both Mexican newcomers in New York (Shin 2012) and Mexicans in Querétaro, Mexico (Quesada & Blackwell 2009; Blackwell & Quesada 2012), whose first person singular forms outweigh their third person singular forms in the expression of overt SPs. Additionally, third person plural forms, which typically disfavor overt SPs, seem to show neither a favoring or disfavoring with a probability weight of .49, thus demonstrating an increased use of overts with these forms.

#### 4.2.2 LEXICAL CONTENT

The constraints on subject expression for the present study related to lexical content of the verb include verbs of perception, psychological verbs, and copulas, which all favor overt SPs, as well as motion verbs, and verbs of communication, which both disfavor overts. These findings are generally consistent with previous studies, with the exception of verbs of communication, which have shown a general favoring of overt use in most other studies (Bentivoglio 1987;

Enrriquez 1984; Silva-Corvalán 1994; Travis 2007), but exhibit a strong disfavoring of overts with a probability weight of .23 in the present study, thus indicating a divergence from other Spanish varieties. With regard to psychological verbs, Roswell speakers follow the trend of frequently using overt SPs with these verbs (probability weight = .66), an observation that Silva-Corvalán (1994) attributes to the fact that psychological verbs tend to express the speaker's point of view and have an implied contrastive function. Further, the additional lexical category that I proposed, verbs of perception, which includes verbs such as *ver*, *escuchar*, and *notar*, show the highest favoring of overt SPs with a probability weight of .69. However, the small number of tokens of this type in the data (N=14) makes a conclusive finding problematic.

#### 4.2.3 REFLEXIVITY

The reflexive or non-reflexive use of a verb was also a conditioning factor for SP use in the data. Specifically, it was the non-reflexive forms that favored overt SPs and the reflexive forms that disfavored overts. This finding is in line with previous research that addressed the factor of reflexivity (Otheguy et al. 2007; Carvalho & Child 2011) and is perhaps due to the added referential clarity of reflexive pronouns, which would make the use of an overt SP unnecessary or excessive (Carvalho & Child 2011:20). The non-reflexive verbs in the present study have a probability weight of .66 and occur with an overt SP 21% of the time while the reflexive verbs' weight is .34 and appear with an overt only 8% of the time.

#### 4.2.4 COREFERENTIALITY INDEX

The continuity or switch between subjects also had an effect on SP usage. As reported in all prior studies, same reference contexts favor null SPs and disfavor overts (Silva-Corvalán 1982; Enrriquez, 1984; Bentivoglio 1987; Cameron 1994, 1995; Bayley & Pease-Alvarez 1997; Travis 2005). Likewise, Spanish in Roswell follows this pattern: where there is no change in

subject referent, the use of an overt SP is disfavored with a probability weight of .37. Further, when there is in fact a change in subject, Roswell Spanish also shows a disfavoring of overt SPs (probability weight = .46), which is surprising considering that most other studies find that such contexts favor overts. This suggests that in Roswell, there seems to be an increasing use of null SPs when there is a switch in referent compared to other varieties of Spanish. Finally, in contexts where there is a switch in subject, but no switch with the preceding object, there was a favoring of overt SPs. This finding is the opposite of what I predicted, and it seems that further analysis of this constraint is needed in order to better understand it. Due to the fact that very few studies employ this constraint and that there are methodological differences in the studies that do, it is difficult to compare findings.

#### 4.2.5 PREVIOUS EXPRESSION

An analysis of the influence of previous SP expression has revealed a priming effect on SP usage in Roswell. An overt SP in the previous clause favors another overt with a probability weight of .64 whereas a previous null SP disfavors an overt with a weight of .37. In other words, the occurrence of an overt SP in one clause conditions the occurrence of an overt in the following clause. Likewise, the occurrence of a null SP in one clause conditions the occurrence of a null in the following clause. This finding is consistent with what Cameron (1994), Flores-Ferrán (2002), and Travis (2007) observed with regard to previous expression.

#### 4.2.6 CLAUSE TYPE

Verbs that appear in both subordinate and main clauses were found to favor overt SPs with probability weights of .62 and .58, respectively. On the other hand, coordinate clause contexts showed a disfavoring of overts with a probability weight of .42. Further, verbs that appeared in contexts where it was difficult to determine the clause type also showed a



disfavoring of overts. The observation that subordinate clauses favor overt SPs is in line with what Morales (1997) found. Additionally, the finding that main clauses favor overt SPs while coordinate clauses disfavor overts is consistent with what Otheguy & Zentella (2012) observed overall for Spanish in New York.

### 4.3 LANGUAGE SHIFT IN ROSWELL

Although the overall pronoun rate of Roswell Mexicans did not indicate any significant divergence from monolingual or incipient contact varieties, a comparison of the effects of specific variables and constraints suggests that Spanish in Roswell may be undergoing change. As illustrated above, the linguistic variables that condition the presence or absence of a subject personal pronoun in Roswell Spanish are, from the strongest to the weakest effect, Person/Number, Lexical Content, Reflexivity, Coreferentiality Index, Previous Expression, and Clause Type. On the other hand, the variables TMA, Specificity, and Lexical Frequency do not have a significant influence on SP use. When comparing this ranking with that of Mainland Newcomers in New York (see Table 10 below), I noted the following differences: Lexical Content and Reflexivity are both ranked much higher in Roswell than in New York. On the other hand, Coreferentiality Index is ranked lower in Roswell. Further, while TMA was found to be a significant factor in New York, it is not significant in Roswell. A further observation of divergence is noted when comparing the ranking of constraints within Person/number between Mexicans in New York (Shin 2012) and Querétaro (Quesada & Blackwell 2009; Blackwell & Quesada 2012) with Roswell Mexicans: while the Spanish of the latter groups favors overt SPs most with first person singular verbs, Spanish in Roswell favors overts most with third person singular verbs.

Table 10: Variable hierarchy as compared with Otheguy et al. (2007:789)

<b>ROSWELL</b>		<b>NEW YORK</b>	
<b>Rank</b>	<b>Variable</b>	<b>Rank</b>	<b>Variable</b>
1 <sup>st</sup>	<b>Person/Number</b>	1 <sup>st</sup>	<b>Person/Number</b>
2 <sup>nd</sup>	<b>Lexical Content</b>	2 <sup>nd</sup>	<b>Coreferentiality Index</b>
3 <sup>rd</sup>	<b>Reflexivity</b>	3 <sup>rd</sup>	Set Phrase
4 <sup>th</sup>	<b>Coreferentiality Index</b>	4 <sup>th</sup>	TMA
5 <sup>th</sup>	Previous Expression	5 <sup>th</sup>	<b>Clause Type</b>
6 <sup>th</sup>	<b>Clause Type</b>	6 <sup>th</sup>	<b>Lexical Content</b>
		7 <sup>th</sup>	<b>Reflexivity</b>
		8 <sup>th</sup>	Genre

Could language shift in Roswell Spanish be due to English contact? Considering the lack of increase in the overall pronoun rate (21%) and the fact that neither length of residency nor age of arrival had a significant impact on overt pronoun frequencies, there seems to be a lack of evidence for the influence of English on Roswell Spanish. However, a language contact hypothesis should not be completely discarded. Following the above mentioned assertions of Poplack & Levey (2010) as well as Torres Cacoulios & Travis (2010) regarding evidence for contact-induced change, the divergences concerning variable and constraint hierarchies noted above could potentially reflect language contact.

Furthermore, there are in fact certain lexical and morphosyntactic features apart from subject pronouns in the data that demonstrate clear evidence of English influence on Spanish in Roswell. According to Lipski (2008), “when two languages come together for sustained periods

of time—in various parts of the world and in a wide range of circumstances—fluent bilinguals inevitably engage in three contact-induced speech phenomena” (p. 223). The three phenomena that Lipski refers to are lexical borrowing, calques, and code-switching, all of which are found in the Roswell data. Examples 69-70 below illustrate the use of lexical borrowing from English while examples 71-72 show the use of word calques. Finally, code-switching from Spanish to English is shown in examples 73-74.

(69) Ellos lo hacen sentir uno *comfortable* allí.

‘They make one feel comfortable there.’ [F37Mex]

(70) ...inculco a mis hijos que sean sociables, que saluden, que, que sean *nice* como personas...”

‘I instill it in my children to be sociable, to greet, to, to be nice as people...’ [F51Mex]

(71) ...y bueno ya no hace nada, [unintelligible] universidad y con altos *grados*...

‘...and well, he doesn’t do anything now, university and with high grades...’ [M23Mex]

(72) ... tienen que *mover* del estado por eso.

‘...they have to move from the state for that reason.’ [F32Mex]

(73) Yo estaba en *high school*, I mean, you are grownups supposedly in *high school*.

‘I was in high school, I mean, you are grownups supposedly in high school.’ [M23Mex]

(74) Estuve trabajando con un señor, ¿cómo se llama? private investigator y... y... *oh my god!* poniendo, poniendo...

‘I was working with a man, what’s his name? private investigator and...and...oh my god! putting, putting...’ [M23Mex]

Additionally, there is evidence of morphosyntactic influences from English. In (75) below, the use of *tienen pena* instead of *les da pena* demonstrates that the speaker is using a structure that is not typical of native Spanish. In (76) the speaker omits the definite article *los* before the word *exámenes*, which is typical in English, but Spanish requires the definite article in this context.

(75) ...no saben y ellos *tienen pena*.

‘...they don’t know and they are embarrassed.’ [M44Mex]

(76) Entonces yo hice exámenes para La Marina...

‘So I took exams for the Marines...’ [M38Mex]

Finally, the use of the present progressive is used in Spanish in contexts which require other forms. In the following two examples, the present progressive would be used in English, but it would be more common in native Spanish to simply use the present indicative (*trabajo*) and the infinitive forms (*limpiar*):

(77) Ella a veces va con la niña porque yo *estoy trabajando*.

‘She sometimes goes with the little girl because I’m working.’ [M43]

(78) Empecé *limpiando*, casas, cuidando niños.

‘I started cleaning, houses, taking care of children. [F32]

## CHAPTER 5

### CONCLUSION AND DISCUSSION

In this thesis, I have attempted to describe the behavior of subject pronouns in the Spanish of Roswell, Georgia, a region that, despite its large and increasing Latino population, lacks the linguistic attention that has been given to regions with more established Spanish-English contact situations in the United States such as those of the West, Southwest, and Northeast. Through a variationist sociolinguistic analysis of the factors contributing to the expression or non-expression of a subject pronoun in this particular Spanish variety as well as a comparison of pronoun behavior with other varieties of Spanish, I have been able to make some conclusions. First, I was able to determine the general pronoun rate among the 10 speakers in the sample and to locate, to a certain degree, the Roswell data with respect to the rates of other US contact varieties as well as monolingual Spanish varieties. I have come to the conclusion that Mexican Spanish in Roswell, at least in terms of overall pronoun rate, is much more similar to monolingual and incipient contact varieties of Mexican Spanish than it is to more established contact varieties of Mexican Spanish. With an overall rate of 21% overt pronouns in Roswell, this is very similar to the rates observed in New York (Otheguy et al. 2007), New Jersey (Flores-Ferrán 2007a), and North Carolina (McKnight 2013), and is identical to the rate observed in Valladolid, Mexico and Mexico City (Solomon 1999; Butragueño, forthcoming). However, this rate is substantially lower than that observed in Los Angeles.

Second, I have found that the language internal factors having the most significant effect on pronoun usage in Roswell are the person and number of the verb, the lexical content or semantics of the verb, and whether or not the verb is used reflexively. These factors are followed by coreferentiality index (same vs. switch reference), pronoun expression of the previous verb, and finally, the clause type in which the verb appears, which have weaker effects on subject expression but are nevertheless statistically significant. On the other hand, the tense-mood-aspect of the verb, the verb's lexical frequency, and whether the verb's subject refers to a specific or non-specific referent were not found to be significant, which was also observed for all language external factors under analysis (gender, age, length of residency in the US, age of arrival to the US).

Third, there are significant differences with regard to the grammar of Roswell Mexicans when compared to monolingual and incipient varieties. The abovementioned hierarchical order of variables differs from that of other US varieties, namely that of mainland newcomers in New York, who consist of Mexican, Colombian, and Ecuadorian speakers. More specifically, as observed in the constraint hierarchies, the person/number variable differs from that of New York Mexicans (Shin 2012) as well as Mexicans in Querétaro (Quesada & Blackwell 2009; Blackwell & Quesada 2012), with the third person singular form superseding the first person singular in the favoring of overt pronoun expression. Also, while most studies find that switch reference contexts favor overt pronouns, such contexts disfavor overts in Roswell.

Finally, as part of my broader research goal, I have considered potential influence from English on the Spanish of Roswell. While it is certainly possible that English is a force shaping Roswell Spanish, I have found no conclusive evidence, at least with regard to pronoun use, of such contact-induced language change given the similarity in overall pronoun rate to

monolingual and incipient Mexican varieties as well as the insignificance of social factors, namely length of residency and age of arrival. Nonetheless, I do in fact conclude that Spanish in Roswell is in a stage of language shift considering its notable divergences from other varieties. I have also observed certain lexical and morphosyntactic features apart from pronouns in the Roswell data that—in a very transparent way—provide evidence for contact-induced change.

While the results of this study certainly reveal, to some degree, a number of important findings about Spanish in Roswell, the small sample size of only 10 speakers makes it difficult to generalize such findings as being representative of the speech community as a whole. Therefore, future research should employ a larger sample of speakers. In addition, it would be useful for future studies to consider a wider range of social factors, for example to take into account the speakers' social class and level of education as well as specific information regarding the speakers' English proficiency as well as amount of exposure to and use of English. Such a wider scope would be particularly beneficial for gaining more insight into the factor of English contact, and more generally, would lead to a more complete description of Spanish in Roswell with regard to subject pronoun expression.

## REFERENCES

- Abreu, Laurel. 2012. Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish. *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, ed. by Kimberly Geeslin and Manuel Díaz-Campos, 1-8. Somerville, MA: Cascadilla Proceedings Project.
- Barrenechea, Ana María, and Alicia Alonso. 1977. Los pronombres personales sujetos en el español hablado en Buenos Aires. *Estudios sobre el español hablado en las principales ciudades de América*, ed. by Juan M. Lope Blanch, 333–49. Mexico City: Universidad Nacional Autónoma de México.
- Bayley, Robert and Pease-Alvarez, Lucinda. 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9.349–71.
- Bentivoglio, Paola. 1987. *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Central University of Venezuela.
- Blackwell, Sarah E. and Margaret Lubbers Quesada. 2012. Third-person subjects in native speakers' and L2 learners' narratives: Testing (and revising) the givenness hierarchy for Spanish. *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, ed. by Kimberly Geeslin and Manuel Díaz-Campos, 142-64. Somerville, MA: Cascadilla Proceedings Project.
- Bosque, Ignacio, and Demonte, Violeta. 1999. *Gramática descriptiva de la lengua española*. Madrid: Espasa.



- Cameron, Richard. 1992. *Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation*. Philadelphia, PA: University of Pennsylvania Doctoral dissertation.
- Cameron, Richard. 1993. Ambiguous agreement, functional compensation, and nonspecific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5.305-34.
- 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. and Child, Michael. 2011. Subject pronoun expression in a variety of Spanish in contact with Portuguese. *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, ed. by Jim Michnowicz and Robin Dodsworth, 14-25. Somerville, MA: Cascadilla Proceedings Project.
- Cifuentes, Hugo. 1980. Presencia y ausencia del pronombre sujeto en el habla culta de Chile. *Boletín del Instituto de la Universidad de Chile* 31.743-52.
- Davidson, Brad. 1996. ‘Pragmatic weight’ and Spanish subject pronouns: The pragmatic and discourse uses of *tú* and *yo* in spoken Madrid Spanish. *Journal of Pragmatics* 26.543-65.
- Eckert, Penelope, and McConnell-Ginet, Sally. 2003. *Language and gender*. Cambridge: Cambridge University Press.

- Enríquez, Emilia V. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas, Instituto Miguel de Cervantes.
- Erker, Daniel and Guy, Gregory R. 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88.526-57.
- 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 contact? *Language Variation and Change* 16.49-73.
- Flores-Ferrán, Nydia. 2007a. Los mexicanos in New Jersey: Pronominal expression and ethnolinguistic aspects. *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*, ed. by Jonathan Holmquist, Augusto Lorenzino, and Lotfi Sayahi, 85-91. Somerville, MA: Cascadilla Proceedings Project.
- Flores-Ferran, Nydia. 2007b. A bend in the road: Subject personal pronoun expression in Spanish after 30 years of sociolinguistic research. *Language and Linguistics Compass* 1.624-52.
- Johnson, Daniel E. 2009. Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3/1.359–83.
- Kretschmar, William A., Jr., Sonja Lanehart, Bridget Anderson, and Becky Childs. 2007. The relevance of community language studies to HEL: The view from Roswell. In *SHEL III - Managing chaos: Strategies for identifying change in English*, ed. by Christopher Cain and Geoffrey Russom, 173-86. Berlin: Mouton de Gruyter.

- Lapidus, Naomi, and Ricardo Otheguy. 2005. Contact induced change? Overt nonspecific ellos in Spanish in New York. *Selected proceedings of the 2nd Workshop on Spanish Sociolinguistics*, ed. by Lotfi Sayahi and Maurice Westmoreland, 67–75. Somerville, MA: Cascadilla.
- Lastra, Yolanda and Pedro Martín Butragueño. Forthcoming. Subject pronoun expression in oral Mexican Spanish. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin. Georgetown: Georgetown University Press.
- Lipski, John M. 2008. *Varieties of Spanish in the United States*. Washington, D.C.: Georgetown University Press.
- Luján, Marta. 1986. Stress and binding of pronouns. *Papers from the Regional Meeting of the Chicago Linguistic Society* 22.248-62.
- McKnight, Rebecca. 2013. *Pronombres de sujeto en Raleigh*. Raleigh, NC: North Carolina State University Master's thesis.
- Michnowicz, Jim. Forthcoming. Subject pronoun expression in Yucatan Spanish. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin. Georgetown: Georgetown University Press.
- Morales, Amparo. 1997. La hipótesis funcional y la aparición del sujeto no nominal: El español de Puerto Rico. *Hispania* 80.153–65.
- Ollolqui de Montenegro, Liliana de. 1987. Un aspecto de la sintaxis: Los pronombres personales sujeto en el habla estudiantil santiaguera. *Actas del I Congreso Internacional sobre el Español de América*, ed. by Humberto López Morales and María Vaquero, 753–64. San Juan, Puerto Rico: Academia Puertorriqueña de la Lengua Española.

- Orozco, Rafael and Guy, Gregory. 2008. El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana? *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*, ed. by Maurice Westmoreland and Juan Antonio Thomas, 70-80. Somerville, MA: Cascadilla Proceedings Project.
- Otheguy, Ricardo, García, Ofelia and Fernández, Mariela. 1989. Transferring, switching, and modeling in West New York Spanish: An intergenerational study. *International Journal of the Sociology of Language* 79.41-52.
- Otheguy, Ricardo, Zentella, Ana Celia and Livert, David. 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 83.770-802.
- Otheguy, Ricardo and Zentella, Ana Celia. 2012. *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford University Press: New York.
- Poplack, Shana. 1993. Variation theory and language contact. In *American Dialect Research*, ed. by Dennis R. Preston with the assistance of the American Dialect Society Centennial Research Committee, 251-86. Amsterdam, The Netherlands: John Benjamins Publishing Company.
- Poplack, Shana, and Levey, Steven. 2010. Contact induced grammatical change: A cautionary tale. In *Language and space: An international handbook of linguistic variation*. Berlin: Mouton de Gruyter.
- Quesada, Margaret Lubbers and Sarah Blackwell. 2009. The L2 acquisition of null and overt Spanish subject pronouns: A pragmatic approach. *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, ed. by Joseph Collentine, Maryellen García, Barbara Lafford & Francisco Marcos Marín, 117-130. Somerville, MA: Cascadilla Proceedings Project.

- Ranson, Diana L. 1991. Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change* 3.133-52.
- Romaine, Suzanne. 2003. Variation in language and gender. *The handbook of language and gender*, ed. by Janet Holmes and Miriam Meyerhoff, 98-118. Oxford: Blackwell.
- Shin, Naomi Lapidus. 2012. Variable use of Spanish subject pronouns by monolingual children in Mexico. *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, ed. by Kimberly Geeslin and Manuel Díaz-Campos, 130-41. Somerville, MA: Cascadia Proceedings Project.
- Shin, Naomi Lapidus and Otheguy, Ricardo. 2013. Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. *Language in Society* 42(4), 429-52.
- Silva-Corvalán, Carmen. 1982. Subject expression and placement in Mexican-American Spanish. *Spanish in the United States: Sociolinguistic aspects*, ed. by Jon Amastae and Lucía Elías-Olivares, 93–120. Cambridge: Cambridge University Press.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon.
- Solomon, Julie. 1999. *Phonological and syntactic variation in the Spanish of Valladolid, Yucatán*. Stanford, CA: Stanford University Doctoral dissertation.
- Torres Cacoullos, Rena and Travis, Catherine E. 2010. Variable yo expression in New Mexico: English influence? *Spanish of the U.S. Southwest: A language in transition*, ed. by Susana Rivera-Mills and Daniel Villa, 185-206. Madrid: Iberoamericana/Vervuert.

- Travis, Catherine E. 2005. The yo-yo effect: Priming in subject expression in Colombian Spanish. *Selected Papers from the 34th Linguistic Symposium on Romance Languages (LSRL)*, ed. by R. Gess and E. J. Rubin. Salt Lake City, 2004. Amsterdam: Benjamins. 329–49.
- Travis, Catherine E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19.101-35.
- U.S. Bureau of the Census. 2010. Census of Population. *Demographics & Population, 2010*. Retrieved from <http://www.roswellgov.com/index.aspx?nid=933>
- Wilson, Anna. 2013. *Stories of Roswell, Georgia: A sociolinguistic study of narrative structure*. Athens, GA: University of Georgia Undergraduate honors thesis.
- Wolfram, Walt, Mary E. Kohn, and Erin Callahan-Price. 2011. Southern-Bred Hispanic English: An emerging socioethnic variety. *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, ed. by Jim Michnowicz and Robin Dodsworth, 1-13. Somerville, MA: Cascadilla Proceedings Project.