SUBJECT EXPRESSION IN A SOUTHEASTERN U.S. MEXICAN COMMUNITY

by

PHILIP P. LIMERICK

(Under the Direction of Chad Howe)

ABSTRACT

This dissertation examines language contact between Spanish and English in the Southeastern U.S. by analyzing Spanish spoken in Georgia. Through an analysis of immigrant Spanish in the city of Roswell, an exurb of Atlanta, potential contact-induced language change is investigated through the lens of subject expression. While most research on Spanish has been carried out in regions such as the Southwest and Northeast, the Southeastern U.S. has not received as much scholarly attention. Therefore, the present investigation seeks to examine understudied varieties of U.S. Spanish, specifically regarding the linguistic processes at work in recent language contact situations. Latin American immigration to the Southeast has led to recent demographic shift in this region and substantial Spanish-speaking populations are emerging that historically were not part of the Southeast. The city of Roswell in particular represents this demographic shift in the Southeast, making it an ideal test site for emerging bilingual speech communities.

The current study examines subject expression among 20 Mexican immigrants using sociolinguistic interview data. The speakers' average length of residency (LOR) in the U.S. is 12 years, and their average age of arrival (AOA) is 27. Tokens of subject pronouns from the interviews were coded for language-internal (linguistic) variables previously shown to constrain subject expression (e.g. person/number, switch reference, tense-mood-aspect [TMA],

morphological ambiguity, polarity, specificity) as well as language-external (social) variables (e.g. English proficiency, age, gender, LOR, AOA), and then analyzed using mixed-effects multivariate analysis in Rbrul (Johnson 2009). Results indicate an overall overt pronoun rate of 27% for Mexicans in Roswell, which is higher than what has been reported for monolingual Mexican Spanish. The multivariate analysis showed that several linguistic variables (e.g. person/number, switch reference, morphological ambiguity, polarity) and one social variable (age) played a significant role in pronoun variation. Moreover, differential effects were revealed when compared to monolingual Mexican Spanish for variables such as TMA and verb class, suggesting an influence of bilingualism. Such divergent linguistic configurations in Roswell Spanish suggest that we are seeing an emergent variety of Mexican Spanish in the U.S. with regard to subject pronoun expression.

INDEX WORDS: subject expression, pronouns, Spanish in the United States, Georgia, syntactic variation, Roswell, Mexican Spanish, language contact

SUBJECT EXPRESSION IN A SOUTHEASTERN U.S. MEXICAN COMMUNITY

by

PHILIP P. LIMERICK

B.A., Indiana University of Pennsylvania, 2009

M.A., The University of Georgia, 2014

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

SUBJECT EXPRESSION IN A SOUTHEASTERN U.S. MEXICAN COMMUNITY

by

PHILIP P. LIMERICK

Major Professor:	Chad Howe
Committee:	Sarah Blackwell
	Margaret Quesada

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2018

DEDICATION

This project is dedicated to my wife, Chantell Limerick, whose continual encouragement, love, and support have kept me moving forward. I also dedicate this work to my son, Josiah Philip Limerick, who gives me inexplicable joy and so much to look forward to.

ACKNOWLEDGEMENTS

This dissertation project, as well as my graduate studies in general, would not have been possible without the support of several people. First, I would like to express gratitude to Dr. Chad Howe, my major professor, for the incredible guidance, insight, enthusiasm, encouragement, and support he has shown to me not only while carrying out this dissertation, but also during my other research endeavors and coursework. Dr. Howe has taught me to think about linguistic variation and the social forces that shape it in ways that have greatly inspired me to do the work that I now love to do. I would also like to thank my committee members, Dr. Sarah Blackwell and Dr. Margaret Quesada, for their support, helpful suggestions, and insight in the development of my dissertation and other projects as well as their fascinating and thought-provoking courses that served as a solid foundation for my research. Additionally, I am grateful to Dr. Diana Ranson for her charisma and enthusiasm in the classroom, for her research guidance, and for teaching me to *pensar lingüísticamente* (think linguistically). I am also grateful for my friend and colleague, Kate Bove, who was simultaneously dissertating and is someone with whom I could commiserate as well as share accomplishments and milestones during this long writing process.

I would also like to acknowledge the main sources of funding and support that facilitated both the data collection and data transcription involved in this project. The data collection was funded by the University of Georgia Graduate School Dean's Award, and the transcription of interviews was funded by the Willson Center for Humanities and Arts Graduate Research Award

at the University of Georgia. Furthermore, I am indebted to both Lilian Zhu and Trevor Talmadge for their incredible assistance with data transcription.

I could not have survived this journey without the immense love and support from my family; I express much gratitude to my wife Chantell for her patience and understanding and to my son Josiah for his laughs and smiles; to my mother Linda Mumau and father Patrick Limerick; to my sister Jade and brother Correy; my stepmother Teresa and stepfather Mick; my mother-in-law Christine and father-in law Rendell; my brothers-in-law Trey, Maurice, and Aaron; my sisters-in-law Rachael and Atenea; and to my nieces and nephews Lila, Abel, Judah, RosaLeah, Renzi, and Sophianna. Spending time with all of you during my breaks from writing has given me the joy, laughter, and encouragement I needed to persevere.

Moreover, I could not have completed this project without all the wonderful people in the Spanish-speaking community of Roswell, Georgia, whose willingness to be interviewed and share their lives with me has not only made my project possible but extremely enjoyable and enriching. ¡Muchísimas gracias a tod@s!

Finally, and most importantly, I thank my living hope, my firm foundation, the one who guards my heart and mind--Jesus Christ.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1 Goals and justification for the study

The purpose of the present study is to shed light on the linguistic behavior of newly-forming bilingual communities in the United States, particularly at the initial stages of language contact between Spanish and English. Looking to the Southeastern U.S. as a relatively new contact situation, this dissertation also aims to understand the processes of language shift that are at work in situations of language contact and bilingualism.

Latin American immigration to the U.S. Southeast has led to recent and significant demographic shift in this region. For instance, between 1990 and 2010 more than a million Latin Americans immigrated to Virginia, North Carolina, South Carolina, and Georgia, with the Latino populations of North Carolina and Georgia having the highest percentage of growth (Wolfram, Kohn, and Callahan-Price 2011). Thus, substantial Spanish-speaking populations have emerged that historically were not part of the Southeast in the way that they were in the Southwestern and Northeastern U.S. Due to the lack of an established presence of Latinos, Spanish in the Southeastern U.S. has not received the same degree of attention in sociolinguistic research as varieties in other areas such as California, New Mexico, and New York. There are, however, notable exceptions to this dearth of research that demonstrate a growing interest in Southeast Spanish (e.g. Smith 2006; Carter 2014; Lamanna 2012; Wilson 2013, 2014; Montes-Alcalá & Sweetnich 2014; Limerick 2017).

To further explore Spanish in the Southeast, this dissertation will focus on Spanish spoken in Georgia, observing one Latino community in particular: Roswell, GA, a city that is part of the Atlanta metropolitan area (See Figure 1.1).

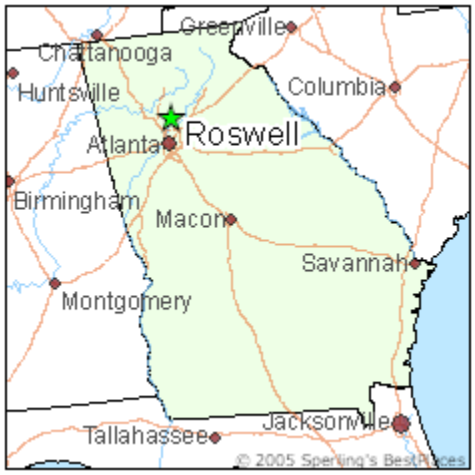

Figure 1.1: Roswell, Georgia. Map by Bert Sperling.

The city of Roswell represents the larger demographic shift in the Southeast and is experiencing recent emergence of Spanish-speaking communities. For example, between 2000 and 2010, the Latino population is Roswell grew (see Table 1.1) from 10.6% to 16.6%, experiencing a 75% increase (U.S. Bureau of the Census 2010). Therefore, Roswell is an excellent test site for studying emerging bilingual speech communities in the Southeast. In addition, given that the language contact situation between Spanish and English is much less established than that of bilingual communities in other U.S. regions, linguistic investigation in this area affords an ideal opportunity to examine the processes of language shift in its initial stages.

Table 1.1: Roswell demographics

Population	**2000**	**2010**
Total Population	79,334	88,346
White	64,666 (82%)	66,010 (75%)
Black or African American	6,773 (9%)	10,373 (12%)
Asian	2,964 (3.7%)	3,565 (4%)
Some Other Race	3,237 (4%)	5,846 (6.6%)
Two or More Races	1,511 (2%)	2,241 (2.5%)
American Indian or Alaska Native	160 (0.2%)	261 (0.3%)
Hawaiian Native/other Pacific Islander	23 (0.03%)	50 (0.05%)
Hispanic or Latino (of any race)	**8,421 (10.6%)**	**14,699 (16.6%)**

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

Furthermore, this dissertation contributes to the broader goals of a research collaboration 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 main objective of the Roswell Voices Project is to document life and language among different communities in Roswell, which has thus far mainly involved research on the African-American and Caucasian communities. The Latino community was first explored by Wilson (2013) by means of sociolinguistic interviews, leading to her analysis of narrative structure in Roswell Spanish. Wilson's fieldwork also led me to conduct my own preliminary research on Spanish in Roswell (Limerick 2017), which I have expanded on for the purposes of this dissertation project.

1.1.1 The linguistic phenomenon under study

The present investigation focuses on a particular phenomenon in Roswell Spanish, variable subject pronoun (SP) expression, which is one of the most extensively studied structures of the Spanish language. There has been a wide range of research dedicated to SP usage from various theoretical perspectives; in addition, it has been examined among native monolingual speakers (e.g. Enríquez 1984; Bentivoglio 1987; Ranson 1991; Cameron 1993, 1994; Davidson 1996; Solomon 1999; Blackwell 1998; 2003; Travis 2005; Orozco & Guy 2008; Orozco 2015, 2016; Alfaraz 2015; Lastra & Martín Butragueño 2015), bilingual speakers (e.g. Silva-Corvalán 1982, 1994a,b; Flores-Ferrán 2002, 2004; Lapidus & Otheguy 2005; Travis 2007; Shin & Otheguy 2009; Torres Cacoullos & Travis 2010, 2015; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015), and L2 learners of Spanish (e.g. Montrul & Rodríguez Louro 2006; Lozano 2009; Rothman 2009; Quesada & Blackwell 2009; Blackwell & Quesada 2012; Geeslin, Linford, & Fafulas 2015; Quesada 2015; Geeslin & Gudmestad 2016). Whereas explicit SPs are nearly obligatory in English, Spanish is a null-subject language, in which speakers can omit the SP, as in (1), or express it, as in (2).

(1) *Sí, ∅* ***estudié*** *para abogada y… en la ciudad de México…* [F39Mex][1]

'Yes, (I) studied to be a lawyer and…in Mexico City…'

(2) *Ah, no,* ***yo*** *lo* ***estudié*** *cuando recién llegué aquí* [M34Mex]

'Uh, no, I studied it when I first arrived here'

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

There are various linguistic and social factors that have been shown to influence such variation (e.g. person/number of the verb, age of the speaker) and while many of these factors are consistent among numerous dialects of Spanish, the distinct social contexts and grammars associated with each dialect reflect unique patterns concerning the explicit or implicit expression of SPs.

In my previous work (Limerick 2017), I explored subject expression through a preliminary analysis of ten Mexican Spanish-speakers in Roswell, which suggested that bilingualism had an influence on their SP usage in relation to processes of grammatical simplification when compared to monolingual Mexican Spanish.[2] Building on this research, my dissertation seeks to further explore the potential effects of language contact in Roswell through the lens of subject expression. Specifically, I examine variables previously unexplored in this community and rarely studied in previous SP research (e.g. English proficiency, preferred media language, polarity, and verbal mood, among others). Through a study of 20 first-generation Mexican immigrants living and/or working in Roswell, I investigate the frequencies of and constraints on subject expression and further explore the possibility of contact-induced language change in this variety of Spanish. While the central focus of the study is on pronominal subjects, I also examine lexical subjects in the current data (Chapter 5).

1.2 Research questions

This project will investigate the influence of language internal (linguistic) and language external (social) factors on SP variation and will compare pronoun use in Roswell with non-contact varieties of Spanish as well as with other U.S. contact varieties. From a usage-based and

[2] See Chapter 2 for a more detailed summary of this study as well as other studies on Spanish in the Southeastern U.S.

variationist sociolinguistic perspective, this dissertation attempts to answer the following research questions:

1. What is the overall distribution of subject expression in terms of overt and null pronoun usage? How does the rate of overt SP usage compare to that of non-contact varieties of Mexican Spanish?

2. How do contextual/language internal factors (e.g. grammatical person/number, tense-mood-aspect, switch reference) condition SP variation and how do such patterns compare to previous research?

    2a. What do the interactions among the linguistic variables reveal about the Roswell data? In other words, what can the non-independent effects on SP variation show us concerning variable subject expression in the current data?

3. Do language external factors such as English proficiency, age, sex, length of residency in the U.S., and age of arrival to the U.S. play a role in how the speakers use SPs?

4. What role, if any, does English contact play in the conditioning of SP rates in Roswell Spanish? That is, do the overall rate of overt SPs or the patterning of constraints on subject expression in comparison with non-contact Spanish reflect externally-motivated change due to English or do they reflect a lack of such change?

5. What do the rates and/or constraints of subject expression tell us about the emergent nature of Roswell Spanish? How does subject expression in this variety compare to that of more stable contact varieties?

The data employed for this analysis come from sociolinguistic interviews I conducted with the Latino community in Roswell during the summer of 2015. To answer these questions, I carry out a quantitative analysis using variationist sociolinguistic methods (Labov 1972; Tagliamonte 2012). Variationist methodology is grounded in the assumption that language is inherently variable and that variation in language is systematically governed by both language-internal and language-external (i.e. social) factors (Labov 1966).[3]

1.3 Outline of this dissertation

The rest of this dissertation is structured as follows: Chapter 2 begins with an overview of Spanish-speaking populations in the U.S., providing some sociohistorical background as well as a description of some of the linguistic characteristics of U.S. Spanish attested in the literature. The second part of the chapter addresses Spanish in the Southeastern U.S. more specifically. First, a sociocultural background of Latinos in the Southeast is given, and then I summarize previous research conducted on Spanish in the Southeast, particularly varieties of Spanish in Georgia and North Carolina.

Chapter 3 reviews much of the literature on Spanish subject expression, surveying various perspectives. First, I present the perspective of traditional grammar in relation to subject expression. I then discuss some of the discourse-pragmatics research on subjects. Finally, and most comprehensively, I address issues in the literature on variable subject expression from a variationist sociolinguistic approach. This latter part of the chapter begins with an overview of variationist theory and how it applies to research on subject expression. Secondly, a brief introduction to the factors that influence variation in subject expression is presented. Third, the

[3] A more detailed discussion of variationist theory and methodology is provided in Chapter 3 (Section 3.3), and the process of how variationist methods were applied to the current data is explained and illustrated in Chapter 4 (Section 4.3).

prominent issues and research regarding the investigation of contact-induced language change are discussed. Then, the overt pronoun frequencies reported in previous studies for various dialects are presented, and finally, an exhaustive discussion of each of the linguistic and social variables analyzed in previous work on subject expression is presented.

In Chapter 4, I present the methodology employed for the present analysis in detail. I begin with outlining the data collection process, discussing the nature of the current data, and describing the participants who were interviewed in terms of their sociodemographic information. Then, I discuss both the qualitative and quantitative methods used in analyzing the interview data. This involves the locating of subject tokens in the data as well as the linguistic and social variables included in the study, illustrations of how the data were coded, and an explanation of the statistical methods utilized. The chapter ends with predictions regarding the results of the investigation.

Chapter 5 presents the main results of the current study. First, the results for the frequencies of overt and null pronouns used by the 20 speakers overall are presented, with a comparison of the overt pronoun rate with other varieties of Spanish. Next, the results from the statistical analyses regarding the joint influence of several linguistic and social variables are discussed with comparisons to previous research. Then, I explore the interactions observed between the different independent variables and their effect on pronoun variation. The final section addresses issues related to lexical subjects and presents the findings for the main uses of these types of subjects in the data.

Next, in Chapter 6, I present a discussion of the main findings and implications of the current analysis and address each of the research questions outlined above. Finally, Chapter 7 discusses the main conclusions reached as a result of conducting the present investigation,

outlines some of the scholarly contributions offered in this dissertation, and, lastly, suggests several areas for future research.

CHAPTER 2

SPANISH IN THE UNITED STATES

In this chapter, I will give a brief overview of Spanish in the U.S. in general, including a sociohistorical background for Spanish-speaking populations in the U.S. as well as descriptions of linguistic features observed among different U.S. Spanish varieties. I will then discuss Spanish in the Southeastern U.S in particular, including the sociohistorical dynamics of the presence of Latinos in this region as well as previous studies that have analyzed Spanish in the Southeastern U.S., including North Carolina and Georgia varieties.

2.1 Historical and sociocultural background

For more than four centuries, Spanish-speaking populations have existed in the United States. Beginning in the 1520s, Spanish colonists explored the regions of what are now Florida, Georgia, and North Carolina (Sturtevant 1962), with the first small settlements being established in northern Florida in 1565 and South Carolina in 1566 (Beck, Jr., Moore, & Rodning 2006). The first long-standing settlement was established in 1567 in North Carolina and is considered to be the earliest European colony in what is now the United States (Beck, Jr., Moore, & Rodning 2006:66).

Furthermore, in 1598 permanent Spanish-speaking communities were established in the Southwestern U.S. in what is now New Mexico (Rivera-Mills & Villa 2010:11). The Spanish language was initially introduced by Spanish colonists in the sixteenth century and continued to increase in the seventeenth and eighteenth centuries due to its use by Spanish missionaries throughout the Southwest from California to Texas (Waltermire 2014). Spanish-speakers have

continued to populate the U.S. increasingly, both in the Southwest and other regions of the country. For example, the Northeastern U.S., particularly the states of New York and New Jersey have experienced extensive immigration primarily from the Spanish-speaking Caribbean. This includes primarily Puerto Rican immigration beginning in the early twentieth century in New York City as well as Dominicans immigrating to New York and New Jersey. Moreover, beginning in the nineteenth century, Cuban immigration to Florida became widespread (Waltermire 2014) and then increased dramatically after the Cuban revolution in the mid-twentieth century.

As of the year 2015, there were 57 million Latinos in the U.S. which comprised 17.8% of the country's total population (Steinmetz & Moreno-Fernández 2015:3).[4] This number is a remarkable increase from just nine million Latinos in 1970 (Steinmetz & Moreno-Fernández 2015:4). The states with the largest Latino populations as of 2015 were California (15.2 million), Texas (10.7 million), Florida (5 million), New York (3.7 million), and Illinois (2.2 million) (Flores 2017). We continue to see extensive immigration from Latin America in recent decades, not only to the Southwest, Northeast, and Florida, but also to Mid-Atlantic and Southeastern U.S. States, among other areas. In fact, between the years 2000 and 2015, Georgia experienced the largest percentage of growth in its Latino population with an increase of 118.8% (Flores 2017), reaching 1 million Latinos in 2015 compared to 435,227 in the year 2000.

2.2 Linguistic features of Spanish in the U.S.

There are numerous varieties of Spanish in the U.S., which often exhibit unique linguistic features and which differ from the typical patterns of their respective monolingual varieties. Such

[4] The term *Latino* is used in this dissertation to refer to people living in the U.S. who are of Latin American descent, whether they be immigrants or U.S. born, and whether they speak Spanish or not.

differential features are often claimed to be the result of influence from English. According to Lipski (2008:223), "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". The three phenomena that Lipski refers to are lexical borrowing, calques, and code-switching. Below I will present some examples of these phenomena attested in the literature on U.S. Spanish as well as some of the morphosyntactic and discourse-pragmatic features observed.

2.2.1 Lexical borrowing, calques, and code-switching

These three phenomena can be observed all throughout U.S. Spanish and demonstrate obvious cases of contact-induced change. For instance, Otheguy, García, and Fernández (1989:43-44) observed all three phenomena among Cuban speakers in New York: lexical borrowing (e.g. *La casa tiene un* ***basement*** *[bEyhmEn]* 'The house has a basement'), code-switching (e.g. *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 (e.g. *A mí me gustan las* ***cartas*** *de béisbol para* ***colectar****las* 'I like baseball cards because I can collect them').[5] The authors also noted an increase in such usage from first generation to second generation Cubans. Additional examples of lexical borrowing from English to Spanish attested in the literature, among countless others, include nouns such as *cash*, *army*, *weekend*, *choice* (Shin 2010), as well as verbs such as *lunchear*, *chequear*, *parquear* and *watchear* (Rothman & Beth Rell 2005). Other calques include *te llamo pa'tras* 'I'll call you

[5] Word calques are distinguished from borrowings in that they involve a transfer only of word meanings and not of word forms (Otheguy, García, and Fernández 1989:44). That is, calques involve words that already exist in Spanish but whose meanings are extended to other contexts (e.g. *cartas* 'letters' is semantically extended to also mean 'cards').

back' (Sánchez 2001), *tener un buen tiempo* 'to have a good time' (Llombart 2003), and ¿*Cómo te gusta*? 'how do you like it?' (Silva-Corvalán 2008). Contact features are not only manifested lexically and semantically as seen in the abovementioned cases of borrowing and calques, but also at the morphosyntactic and discourse-pragmatic levels.

2.2.2 Morphosyntactic features

Silva-Corvalán (1994a) observed certain syntactic features among Mexicans in Los Angeles. For instance, she observed that Mexican speakers often omit the complementizer *que* 'that' as in *Yo creo Ø inventaron el nombre* 'I think (that) they invented the name' (p. 136). Specifically, such non-expression of *que* was attested only in complement clauses and not in relative clauses (**El nombre Ø (ellos) inventaron era extraño* 'The name (that) they invented was strange' (Silva-Corvalán 1994a:137). This example demonstrates the use of a syntactic structure absent from general Spanish, but manifests under contact with English. In other cases, some structures already in use in monolingual Spanish are simply elevated in situations of contact, such as a higher use of *estar* in contexts generally reserved for *ser* (Silva-Corvalán 1994a), an increased use of the indicative over the subjunctive (Silva-Corvalán 1994b; Zentella 1997; Lynch 1999), a greater use of progressive constructions (*estar* + gerund) (Solé 1977, cited in Klee 2009:238), and, directly related to this dissertation, a heightened frequency of overt subject pronouns (Otheguy & Zentella 2012).

2.2.3 Discourse-pragmatic features

Another contact feature observed in Los Angeles Spanish, one that relates to discourse constraints, 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 1994a). For example, the first mention of a teacher would typically 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. Moreover, numerous researchers have observed the insertion of English discourse markers in bilingual speech (e.g. Torres 2002; Lipski 2005; Said-Mohand 2008). For instance, Torres (2002) studied the use of markers such as *so*, *y'know*, *and*, *because*, and *I mean* among Puerto Ricans in New York. Furthermore, although much less studied, some researchers have observed influence of English on the production of certain speech acts, such as compliments used by Chicana women (Yáñez 1990) and requests made by Cubans in Miami (Gutiérrez-Rivas 2007). The abovementioned examples illustrate just a few of the specific features found in Spanish-English contact varieties in the U.S., and several others have been documented in various regions.[6]

2.3 Spanish in the Southeastern United States[7]

As stated in the Chapter 1, Spanish in the U.S. Southeast is relatively understudied in comparison to other regions such as the Southwest and Northeast. Nevertheless, there is growing interest in Spanish speaking populations in the Southeastern US. In this section, I will first discuss some of the sociohistorical background regarding Spanish-speaking populations in the Southeast and then I will summarize some of the work that has been done in this region.[8]

[6] For extensive discussions on Spanish in the U.S., see, for example, Silva-Corvalán (1994a); Lipski (2008); Klee & Lynch (2009, Ch. 6); Escobar & Potowski (2015)

[7] In this dissertation, the Southeast includes both Mid-Atlantic Southern states, such as North Carolina, as well as those further South (e.g. Georgia), with the exception of Florida.

[8] Aside from Spanish, other studies of language contact in the Southeast have examined Hispanic English, such as that of Wolfram, Kohn, and Callahan-Price (2011), which discusses the formation of an emerging ethnic variety of English. These 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 (p. 5). The existence of such vowel sounds, as well as other observed interdialectal features, is a defining characteristic of this emerging variety of Hispanic English in North Carolina. For related research on Hispanic English, see also Wolfram, Carter, and Moriello (2004) and Carter (2005).

2.3.1 Sociohistorical background

Within the last 30 years, the Southeastern U.S. has seen immense immigration from Latin America, with many Southern cities experiencing growth rates in their Latino populations above 500% since 1990 (Winders 2005) and several states above 100%. For instance, between 2000 and 2011, the states of Georgia, North Carolina, and South Carolina had growth rates of 102%, 119.6%, and 154.5%, respectively (Motel & Patten 2013). Such demographic shifts are not only seen in a few cities or states, but have been evidenced in the Southeast as a whole. Overall, the Southern U.S. had the highest growth rate (57.3%) of all regions of the country in the 2010 U.S. Census. As a point of reference, Table 2.1 below shows the Hispanic or Latino populations of several Southeastern states in both 2000 and 2016. As these numbers demonstrate, the states with the highest Latino populations as of 2016 were Georgia, North Carolina, and Virginia. One can also see the large increases that occurred within these 16 years, with Latino populations in some states doubling.

Table 2.1: Hispanic or Latino populations in the Southeastern U.S. in 2000 and 2016[9]

State	**2000 population**	**2016 population**
Georgia	435,227 (5.3%)	926,990 (9.2%)
North Carolina	378,963 (4.7%)	884,763 (8.9%)
South Carolina	95,076 (2.4%)	258,361 (5.3%)
Virginia	329,540 (4.7%)	725,092 (8.7%)
Alabama	75,830 (1.7%)	193,503 (4%)
Tennessee	123,838 (2.2%)	327,345 (5%)
Kentucky	59,939 (1.5%)	146,945 (3.3%)
Mississippi	39,569 (1.4%)	86,704 (2.9%)

[9] https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml

Some of the factors motivating such widespread immigration in recent decades include economic hardship, particularly the 1980s debt crises in Mexico and other Latin American countries (Pastor 1998), as well as the need in the U.S. South for low-wage workers in the service sector (Winder 2005). This influx of Latin American immigrants has led to a widespread use of Spanish in Southeastern states which, during most of the twentieth century had been almost entirely English-speaking, such as in Georgia and North Carolina (Klee & Lynch 2009).

2.3.2 Previous research on Southeast Spanish

Some notable studies, particularly on North Carolina Spanish, include Lamanna (2012) and Carter (2014). These studies in particular are revealing of the emergent nature of Southeast Spanish and begin to show a picture both of newly forming varieties of U.S. Spanish as well as discourses about the Spanish language in recently established U.S. Latino communities. Lamanna (2012) studied Colombian Spanish in North Carolina, specifically examining the imperfect/imperfect progressive distinction to assess possible English contact influence. Since English only has one past progressive form (an analytic form equivalent to the imperfect progressive in Spanish), the prediction was that bilingual speakers would use the analytic imperfect progressive form more often. Lamanna indeed found a significant increase in this form among NC speakers when compared to monolingual Colombian speakers, which he attributes to English contact. Lamanna also observed evidence of dialect contact influence among Colombian and Mexican Spanish-speakers, particularly in their use of pronominal address forms. He concludes that these cases of contact-induced change constitute the formation of a new variety of U.S. Spanish.

From a more sociocultural and discourse analytic perspective, Carter (2014) analyzed discourse about Spanish and U.S. Latinos by non-Latinos in a North Carolina middle school,

finding that much of the discourse was harmful and dehumanizing in a way that further marginalized Latinos and devalued the Spanish language. Much of the talk that Carter observed involved, among other characteristics, anti-immigrant sentiment and revealed ideologies related to Latinos as dangerous and a threat. Linking this discourse to broader societal and institutional practices, Carter maintains that these negative perceptions toward Spanish and Latinos by non-Latinos are influenced, among other factors, by the employment structure of the school itself, such as the fact that the higher-ranking positions are occupied by English monolinguals while the lower-ranking ones are filled by Spanish monolinguals. Other problematic issues argued to play a role in the production of such harmful discourse concerning Spanish and Latinos include the prohibition of speaking Spanish by certain teachers, the general favoring of English monolingualism, and the stigmatization of Spanish monolingualism in the school. In sum, Carter argues that these types of discourses, institutional practices, and ideologies have a very limiting and detrimental effect on the identities of Latino students in North Carolina.

The other main region of the Southeast on which research has been carried out is Georgia. For instance, Smith (2006) studied codeswitching in a Northeast Georgia Latino community among Latin American immigrants and observed convergence toward English when their codeswitched utterances reached a certain quantity of English use. Specifically, Smith found that, when less than 70% of speakers' utterances were monolingual Spanish, (as opposed to bilingual Spanish-English utterances [i.e. codeswitching]), their Spanish showed patterns of morphosyntactic convergence toward English. For example, he observed a lack of distinction between person/number with past tense verbs in some cases (e.g. *ellos* ***dijo*** 'they said') (p. 217), which, according to Smith, illustrates as potential English influence since English does not distinguish these verb forms. Smith's analysis demonstrates that (non)convergence can be

predicted by the amount of codeswitching speakers perform. Moreover, Montes-Alcalá and Sweetnich (2014) studied language shift from Spanish to English among Latinos in Atlanta. Arguing against the "shift happens" theory (Hurtado & Vega 2004), they found that speakers can maintain Spanish (and not shift to English) after the second generation. One of the main factors influencing Spanish maintenance in their study was greater contact with the language, particularly Spanish use with family and community. Crucially, such language use favored positive attitudes toward Spanish, which in turn predicted maintenance among third generation speakers. Other influencing factors included Spanish proficiency, length of residency, and age of immigration.

Finally, a few studies have examined Spanish in Roswell, Georgia, the city of focus for the present investigation (Wilson 2013; Limerick 2017). Wilson (2013), for example, examined narrative structure among Roswell speakers. For this analysis, she conducted sociolinguistic interviews in Roswell and utilized Labov and Waletzky's (1967) framework of narrative structure to analyze variation in discourse structure between Spanish and English as well as among individual speakers. Wilson concluded that, among Roswell speakers, there is more variability in narrative structure than was previously supposed, and that Labov and Waletzky's narrative paradigm, which involves a specific set of structures that occur in a specific order, functions differently in Spanish than in English. For instance, she found that Spanish narratives tended to focus on background information and the outcome of the story while narratives in English dedicated more time to the action of the story and gave less importance to the people involved or the outcome.

Regarding subject expression in particular, my previous work (Limerick 2017) examined pronoun use among Mexican immigrants utilizing Wilson's (2013) interview data. For that

analysis, I explored the initial stages of language contact in Roswell by analyzing SP patterns and considering possible contact effects from English. I found that subject expression was constrained by person/number, switch reference, TMA, lexical content, and clause type, the same factors (among many others) reported as significant in numerous other studies. In terms of the possibility of direct English contact effects, little to no evidence was found. Specifically, the overt pronoun rate was very similar to that of monolingual Mexican Spanish reported in other studies, and, crucially, the social factors analyzed, such as length of residency and age of arrival, were not found to be significant. Nevertheless, a more indirect influence of bilingualism was suggested due to a reduced effect observed for the switch reference constraint. This reduced effect was interpreted as a process of bilingual simplification on part of the speakers whereby they were potentially losing sensitivity to switch reference as evidenced by increased overt SP usage in same reference contexts, a context that more strongly favors nulls in monolingual varieties.[10] In order to expand and further develop this research, this dissertation draws from a larger data set of sociolinguistic interviews that I collected in Roswell and more closely examines language contact effects by analyzing additional factors (e.g. English proficiency; preferred media language) and comparing SP patterns with non-contact varieties of Spanish.

2.4 Summary

Although Spanish has been spoken continuously for centuries in many regions of the U.S., the emergence of Spanish-speaking populations in the Southeast is relatively new. Scholarship on Southeast Spanish, therefore, is comparatively sparse, but fortunately there is recent interest in this area, as demonstrated by the previous research summarized in this chapter. Several open questions remain regarding Spanish spoken in the Southeast. For example, what do

[10] See Chapter 3 for a more detailed discussion of investigating English contact through subject expression.

patterns of mood choice, copula choice (*ser*/*estar*), final /s/ weakening, or complementizer omission look like among Spanish-speakers in the Southeast? What role, if any, does English contact play in the production of speech acts in Southeast Spanish? How does the use of codeswitching in other Southeastern states compare to the findings of Smith (2006) in Georgia? How do the imperfect/imperfect progressive distinction or the use of pronominal address forms compare to the patterns observed by Lamanna (2012)? Additionally, do we see Spanish maintenance beyond the second generation in other Southeastern states, and how might geographical context (e.g. urban vs. rural) and other factors influence the shift/maintenance of Spanish, particularly in light of the findings of Montes-Alcalá and Sweetnich (2014) discussed above?

Moreover, several issues remain unexplored related to subject expression in Spanish spoken in the Southeast, such as the following: Does the preferred language for media have an effect on the use of subject pronouns? Does the level of English proficiency of Spanish-speakers influence subject expression patterns? What additional language internal and language external factors are at play in regulating the subject expression behavior of Southeast Spanish-speakers? It is precisely these types of questions, among others, that will be addressed in this dissertation.

Finally, one issue of interest that will be under discussion in the present investigation is the notion of *emergent* varieties of Spanish (Moreno-Fernández 2009). In previous research, emergent or new varieties of Spanish have been described in terms of linguistic characteristics or patterns that arise in a given bilingual/bidialectal variety that distinguish it from its monolingual counterpart. Moreno-Fernández (2009:217) defines emergent varieties as those in situations of language/dialect contact that adopt characteristics that are not part of the original linguistic configuration of the non-contact variety. For instance, as discussed previously in this chapter,

studies have shown that many varieties of Spanish in the U.S. have adopted distinct lexical, morphosyntactic, and pragmatic characteristics that distinguish them from their monolingual counterparts. For the Southeast in particular, the aforementioned use of codeswitching in Georgia (Smith 2006) and heightened use of the imperfect progressive in North Carolina (Lamanna 2012) demonstrate a few of these characteristics that would constitute an emergent variety.[11] Aside from being distinct from monolingual varieties, emergent varieties would also be different from the more established bilingual varieties of U.S. born speakers in the sense that they appear only recently due to immigration (Lamanna 2012). With emergent varieties, we see substantial changes from monolingual speech, but also processes of language shift that are not as advanced or developed as those observed in established second- or third-generation immigrant communities.[12] The following chapter discusses previous literature on Spanish subject expression from multiple perspectives.

[11] In a similar vein, Lamanna (2012:2) uses the term *new* variety to refer to "a linguistic variety that has come into being only recently as a result of immigration and that exhibits contact-induced change when compared with a corresponding monolingual variety."

[12] See Chapter 6 for a detailed discussion of emergent varieties in relation to the findings of the present investigation.

CHAPTER 3

SPANISH SUBJECT EXPRESSION

In this chapter, I will discuss previous literature on subject expression in Spanish from multiple perspectives, including traditional grammar, discourse-pragmatics, and, most extensively, variationist sociolinguistic approaches. This latter perspective will include the following: (a) a brief overview of variationist theory and how it is applied to subject expression research, (b) a brief outline of the factors shown to influence variable subject expression, (c) the prominent issues in the investigation of contact-induced change from English, (d) the overt pronoun rates for various dialects reported in the literature, and (e) an in-depth discussion of each of the linguistic and social variables examined in previous research that are relevant to this dissertation.

3.1 Views of subject expression from traditional grammar

According to traditional grammarians (e.g. Bosque & Demonte 1999; Butt & Benjamin 2004; King & Suñer 2004), the primary function of using overt pronouns in Spanish is that of emphasis and contrast. Explicit pronominal forms in place of null pronominal forms are meant to emphasize the subject referent or to express contrast among different subject referents. On the other hand, the absence of the overt pronoun (the null pronoun) is neutral, in other words, not emphatic in the sense that there is no intention by the speaker to make a contrast between two or more subjects. Therefore, in an utterance such as ***yo*** *tengo que salir* 'I have to leave', 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, etc.).[13] However, when there is no particular emphasis desired, the speaker can omit the *yo*, while still leaving the sentence perfectly grammatical (**Ø** *tengo que salir* '(I) have to leave'). Additionally, overt SPs can be used in contexts of potential ambiguity, for example, when the verb is in the third person, as in the following example presented by Bosque and Demonte (1999:1282):

(1) *Ya decía* ***yo***

'I was already saying'

In (1), the SP could have also been *él* 'he', *ella* 'she', or *usted* 'you (formal)' according to the verbal inflection. The overt *yo*, however, is necessary to disambiguate the referent, unless there is previous context from which the referent could be identified.

Furthermore, according to Butt and Benjamin (2004) overt SPs should only be used and are only used for the aforementioned functions, and posit that "[i]t is a bad error, common among English-speakers, to use Spanish subject pronouns unnecessarily" (p. 130). They go on to cite the following example and explanation from Bosque and Demonte's (1999) *Gramática descriptiva de la lengua española* in order to illustrate this error:

(2) ***yo*** *me vestí, y después* ***yo*** *fui a recoger a mi hijo, pero* ***yo*** *llegué tarde*

'I got dressed, and afterwards I went to pick up my son, but I arrived late'

(Butt and Benjamin 2004:130)

Butt and Benjamin argue that all of the overt SPs must be omitted in this example, with the possible exception of the first one provided that it is needed for the above-mentioned functions. However, as variationist research has shown (e.g. Travis 2005), it is completely natural for native Spanish speakers to produce such utterances with multiple consecutive overt SPs without

[13] See Section 3.2.3 as well as Chapter 4 for a more in-depth discussion of contrast and its relation to pronoun usage.

necessarily expressions the functions of emphasis, contrast, or clarification of ambiguity in each case. For instance, *priming* effects have been observed whereby speakers tend to repeat a recently used structure; in the case of SPs, it has been found that overt SPs are likely to be used when they were recently used in the immediately preceding discourse and that the same is true for null SPs (Cameron 1994; Flores-Ferrán 2002; Travis 2005, 2007). Thus, the use of consecutive overt (or null) SPs can be explained, at least in part, by this priming effect that is common in native Spanish.

These are just a few of the reasons why a speaker might use an overt instead of a null SP. As I will demonstrate throughout this dissertation, however, the variable use of pronouns is much more complex and is in fact systematically conditioned by semantic, pragmatic, and morphosyntactic factors. In addition, there are many contexts in which either the overt pronoun, or more frequently, the null pronoun is (nearly) obligatory in Spanish, which I discuss in Chapter 4.

3.2 Discourse-pragmatic approaches

Numerous studies have shown that SP usage in Spanish is influenced by discourse-pragmatic factors (e.g. Enríquez 1984; Gundel et al. 1993; Davidson 1996; Blackwell 1998, 2001, 2003; Quesada & Blackwell 2009; Mayol 2010; Blackwell & Quesada 2012; among many others). Specifically, as noted above, speakers tend to use overt SPs in conversation to carry out certain communicative functions such as to express emphasis or contrast (e.g. ***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). Other researchers discuss overt pronoun use

in terms of politeness and hedging (e.g. Stewart 2003), arguing that *yo* and *tú* can be used as hedging devices in order to negotiate face.

Furthermore, Quesada and Blackwell (2009) examined first-person singular SPs in relation to five pragmatic constraints that restrict their use among both Spanish L2 learners and native speakers: *salient referent, epistemic parenthetical* (which favor null SPs), *switch focus, contrastive focus, and pragmatic weight* (which favor overt SPs). In the following subsections I will discuss the notions of *saliency*, *switch focus*, *contrastive focus*, *pragmatic weight*, and *epistemic parenthetical*, as they pertain to the constraining of SP usage and which have been highlighted by various authors (e.g. Bosch 1983; Enríquez 1984; Gundel 1999; Quesada & Blackwell 2009). I will also discuss the findings of Quesada and Blackwell's (2009) analysis, which reveals the ways in which L2 and L1 speakers' SP behavior is regulated by discourse-pragmatics. As will be highlighted, their research, among that of many others, demonstrates that SPs are not simply used in free variation and are not "optional", despite what has been suggested in traditional grammars and textbooks (Quesada & Blackwell 2009:128) but rather to achieve specific communicative goals and are conditioned by pragmatic factors in Spanish.

3.2.1 Referent saliency and the cognitive status of 'in focus'

The notion of cognitive attention states in discourse may explain the behavior of subject expression. In particular, the state of *in focus* is described by Gundel (1999) as a state in which the attention of the interlocutors can be assumed to be focused on a given entity or referent due to its salience in the discourse. Regarding SP usage in particular, if the intended referent is salient and is the focus of attention at a given point in the discourse, this may influence the use of null SPs. In fact, Gundel et al. (1993) reported that the *in focus* status correlates with the use of unstressed and null pronouns cross-linguistically. A related way of describing this notion is in

terms of Bosch's (1983) *aboutness* hypothesis which states that the most salient entity at any point in the discourse is always the object that the discourse is about. Therefore, this entity is referred to with anaphoric pronouns. Blackwell (2003) demonstrates that in Spanish minimal forms such as null subjects (as well as clitic pronouns) typically achieve such anaphoricity, as seen in the following example with a series of null SPs (p. 126):

(3) *…Y que por esa Carmen Prendes, me acuerdo que* ***Ø*** *estaba que ahora ya* ***Ø*** *está de pelo bla:nco,* ***Ø*** *sale en la televisión blanca de pelo bla:nco, muy vieja ya, muy mayor.*
'And because of that Carmen Prendes, I remember that (she) was there and now (she)'s already got white hair, (she) appears on television white with white hair, [and is] very old now, very elderly.'

In (3), the use of null SPs reflects the basic pattern of anaphora in that the null forms are coreferential with the preceding subject, Carmen Prendes, who has in focus status. The idea of salient or in focus referents is also related to the notion of *accessibility* (e.g. Givón 1983; Ariel 1994). Specifically, referents that are the most accessible in the listener's mind are also those that are most salient and are thus likely to be marked with less coding material (in this case, null SPs) whereas less accessible and less salient referents will more likely be marked with more coding material (in this case, overt SPs) (Givón 1983:18). Accessibility is discussed in terms of distance, whereby shorter distances between referents and their antecedents imply high accessibility while longer distances imply low accessibility (Ariel 1994). Additionally, the prominence of antecedents plays a role in accessibility. According to Ariel (1994), more prominent antecedents such as subjects, topics, and humans are more accessible than non-subjects, non-topics, and non-humans. Thus, more accessible and salient referents favor null SPs since they were recently mentioned, are more prominent, and are in focus.

3.2.2 Switch focus

According to Quesada and Blackwell (2009), "overt SPs may be viewed as implicating a meaning (or reference) that a null subject would not (in most cases) encode" (p. 119). This idea relates to the switch in focus of attention from one entity to another and is based on Levinson's (2000) neo-Gricean M(anner)-Principle, which instructs speakers to use a marked (less common, more complex) linguistic form to implicate an alternative interpretation, that is, an interpretation different from that which would be implicated by an unmarked (more common, simpler) form. Therefore, overt SPs as the marked form may encode a different meaning from null SPs, which are unmarked. Quesada and Blackwell (2009) state that native Spanish speakers tend to use overts to shift the focus of attention to another referent when that referent is "(1) not currently in focus and (2) not the same referent as the one to which a null subject in the same context would refer" (p. 119). They illustrate switch focus in the following example where *yo* serves to shift attention from the speaker's friend to the speaker herself, which would not have been achieved by a null subject:

(4) *De la primera vez que [Ø] me enamoré no tiene mucho. [Ø] Fue el año pasado, de una persona que se llama Jonathan. El, [Ø] lo conocí porque [Ø] llegó un día a saludar a uno de mis mejores amigos y mi amigo entró.* ***Yo*** *estaba en la Preparatoria, y mi amigo entró a una clase y…* (Quesada & Blackwell 2009:120)

'Concerning the first time that (I) fell in love there isn't much to tell. It was last year, with a person named Jonathan. He, (I) met him because (he) came one day to greet one of my best friends and my friend came in. I was in high school, and my friend came in to a class and…'

Like salient/in focus referents discussed in the previous section, the idea of switch focus is related to *accessibility* in that subjects that make a switch in focus of attention have less accessible, less salient, and more distant antecedents and thus are more likely to be referred to with overt SPs.

3.2.3 Contrastive focus

As mentioned above, one of the basic functions of overt SPs in discourse is to express contrast. Related to this, the notion of focus has also been discussed in the literature and can be applied to SP expression. Focus is described by Hidalgo-Downing (2003) as the element or expression of the most salient and essential information in an utterance. In terms of information structure, the focus can be new, non-presupposed information as well as given information that has already been introduced into the discourse. According to Gundel (1999), elements with contrastive focus are meant to draw the hearer's attention to the elements and are prosodically emphasized with the purpose of contrasting them with other potential members of a contrast set. With regard to SP usage, this occurs when contrasting two (or more) subjects in discourse (e.g. ***Él*** *tiene dieciocho años y* ***yo*** *tengo veinticinco años* 'He is eighteen years old and I am twenty-five years old). This contrast can either be explicit or implicit and is described by Enríquez (1984) as a context in which the speaker intentionally counterposes the subject to one or more people for reasons of opposition, confirmation, or to adopt a different position. The following examples from Enríquez (1984:115, emphasis mine) demonstrate contrastive focus, particularly when certain entities of the contrast set are implied and not explicitly stated in the utterance itself:

(5) ***Yo*** *vengo solo [los demás no sé]*

'I come alone [the others I don't know]'

(6) ***Él*** *no se atreve a opinar [pero su hermano siempre]*

'He doesn't dare opine [but his brother always does]'

In (5) and (6), respectively, *Yo* 'I' is expressed as being in opposition to *los demás* 'the others', and *Él* 'he' is in opposition to *su hermano* 'his brother'. Quesada and Blackwell (2009:121) also illustrate the use of contrastive focus in their data in which the speaker refers both to himself and to a female referent with two contrastively focused pronouns:

(7) *... [Ø] pasamos mucho tiempo juntas además como que [Ø] es una retroalimentación,* ***yo*** *dependo de ella y* ***ella*** *de mí, [Ø] podemos hablar de de tontería y...*

'(we) spend a lot of time together and not only that, it's like we feed off of each other, I depend on her and she on me, (we) can talk about about nonsense and…'

Moreover, while many authors would argue that overt SPs are pragmatically obligatory in contrastive contexts, some researchers have challenged this view (e.g. Schwenter 2002; Amaral & Schwenter 2005) and argue that the use of certain adverbials can alternatively aid in carrying out such contrastive functions without the need of an overt SP. For example, in their analysis of spoken corpus data, Amaral and Schwenter (2005) point out that adverbials such as *aquí* 'here', *por mi parte* 'as for me', *honestamente* 'honestly', and *personalmente* 'personally', among others, that co-occur with null SPs can serve to express a contrast between subject referents that would otherwise be conveyed by an overt SP--e.g. *Mis padres veranean en la playa. <u>Por mi parte</u>,* ***Ø*** *prefiero ir a la montaña* 'My parents spend the summer in the beach. As for me, (I) prefer to go to the mountain' (p. 121) or *Los amigos de Ana siempre llegan tarde a las fiestas. <u>Honestamente</u>,* ***Ø*** *preferimos llegar temprano.* 'Ana's friends always arrive late at parties. Honestly, (we) prefer to arrive early' (p. 123). Therefore, Amaral and Schwenter's (2005) analysis demonstrates that it is not always the case that speakers convey contrastive focus only

through the use of overt SPs. Similar to their examples, I have observed cases of contrastive focus in the Roswell data in which a null SP occurs, as in (8) below where the adverbial *en lo personal* 'personally' is used:[14]

(8) *Bueno, <u>en lo personal</u>, **Ø** no **soy** muy religiosa pero…por lo que veo, la gente sí va bastante a la iglesia, aquí.* [F30Mex]

'Well, personally, I'm not very religious but…from what I see, people do go quite a bit to church, here'

3.2.4 Pragmatic weight

The notion of pragmatic weight is described by Davidson (1996) as a way in which speakers "increase their 'stake' in whatever they are saying, either in an argument or in a statement of belief" (p. 551). According to Davidson, this is achieved by the use of an overt SP, which increases the speaker's commitment to an utterance and also expresses greater personal involvement. Davidson observed that the use of overt SPs, particularly first person singular forms in Spanish, are used to add pragmatic weight typically with verbs of opinion, belief, and knowledge such as *pensar* 'think', *creer* 'believe' and *saber* 'know' as seen in (9) below:

(9) *Bueno, eso lo piensas, pero luego, cuando llega el momento, **yo** creo que es un poco distinto, /,eh?*

'Well, [you] think that, but later, when the moment arrives, *I think* that [it]'s a little different, eh?' (Davidson 1996:556, emphasis mine)

The speaker is expressing his/her point of view with the verb *creer* and is using the overt *yo* to emphasize and to increase his/her commitment to the state of belief, which is different from that of the other interlocutor. Quesada and Blackwell (2009:122) have also observed pragmatic

[14] For further illustrations of this type in the present data, see Chapter 4.

weight with verbs of emotion such as *sentir* 'feel' and *querer* 'want', as seen below where the speaker is expressing how she/he feels:

(10) *... primero [Ø] fue una amistad así muy bonita y como de tres, cuatro meses. De hecho* ***yo*** *siento que a veces [Ø] me hacía más caso a mí que a su novio,…*

'first it was a very nice friendship of about three, four months. Actually, I feel like sometimes she paid more attention to me than to her boyfriend'

3.2.5 Epistemic parentheticals

Finally, epistemic parentheticals are also expressions that are relevant to SP use. Thompson and Mulac (1991) classify epistemic parentheticals as verbs that are used with an absence of their truth-functional content, taking on a more abstract meaning and serving to convey information regarding how speakers position themselves in relation to their utterance. According to Davidson (1996), verbs that are commonly used in this way are verbs of knowing, seeing, or watching, which he has observed for Spanish, as illustrated in the example below (p. 558, emphasis mine):

(11) *Pero lo aprobó, lo aprobó con un trabajo,* ***¿sabes?*** *Porque,* ***no sé*** *es que hay mucha diferencia, yo creo, entre el Árabe de primero y el de segundo.*

'But [I] passed it, [I] passed it with a project, [you] know? Because, [I] don't know, [it]'s that there's a big difference, I believe, between first year Arabic and second year.'

In (11), the verb *saber* 'know' is not being used with its literal, truth-functional meaning but rather as a fixed expression in the first occurrence (¿*sabes*? '(you) know?') and functions as a hedge or mitigation of the strength of the assertion in the second (*no sé* '(I) don't know'). According to Quesada and Blackwell (2009), Spanish epistemic parentheticals such as *no sé* '(I)

don't know' or *digo* '(I) mean' or '(I)'m saying' are used as evaluative commentary or asides that reveal speakers' reflections on the content of the utterance. They further explain that the fact that they are deemphasized expressions accounts for the common use of null SP expression. The authors illustrate epistemic parentheticals in their data with the following examples (p. 122):

(12) *Pues [Ø] busca, adelante". **[Ø] No sé**, pues [Ø] era muy cómodo de ...*

'Well (she/he) continues to search. (I) don't know, well (it/I/she/he) was very comfortable…'

(13) *[Ø] hablo por mis hermanos que también **[Ø] digo**, este, mi mamá es enfermera, una ...*

'I'm speaking for my siblings/brothers that also, (I)'m saying, umm, my mom is a nurse, a…'

In carrying out their study and exploring the five aforementioned notions of saliency, switch focus, contrastive focus, pragmatic weight, and epistemic parenthetical, Quesada and Blackwell analyzed spoken narratives from beginner, intermediate, and advanced L2 learners of Spanish as well as L1 Spanish speakers from Mexico and found that learners generally improved their knowledge of the constraints as they advanced from beginner to advanced levels. Beginner and intermediate learners showed an excessive use of both overts and nulls. For example, they used null SPs in contrastive focus contexts and overused overts in salient referent contexts. However, advanced learners demonstrated more extensive knowledge of the pragmatic constraints, and their use of SPs was much more similar to that of the L1 Spanish speakers in the study. Namely, they preferred nulls in salient referent contexts and overts for contrastive focus, switch focus, and pragmatic weight. As one can see from the above discussion, variation in Spanish subject expression is highly regulated by discourse-pragmatic factors.

However, one major methodological issue that is not typically addressed in the pragmatics literature on SPs is how exactly some of the above-mentioned pragmatic constraints are determined by the analyst. That is, how does one know if a given token is a case of *salient referent* or *pragmatic weight*, for instance? Without objective measures to determine such uses, impressionistic and inconsistent data analysis becomes a concern. Regarding *salient referent*, it is difficult to know when a particular antecedent is or is not "in focus" in the listener's mind. One proposal might be to consider as salient any referent that was mentioned relatively recently in the discourse (e.g. 3 clauses back). The constraint of *switch focus*, while being discussed as SPs whose referents were different from the one that was salient in the preceding discourse (Quesada & Blackwell 2009:125), nevertheless lacks a reliable operationalization without the initial establishment of what constitutes *salient*. To take another example, how is *pragmatic weight* determined? This is a difficult constraint to quantify and operationalize because we do not have an operational definition of what counts as *pragmatic weight* upon analyzing speech data. Moreover, some of these categories overlap with each other. For instance, a case of *salient referent* could simultaneously be a case of *pragmatic weight*. Thus, it is difficult to distinguish each of these constraints on SP use as separate or unique. While it is crucial to take into account such discourse-pragmatic functions with regard to subject expression, the primary focus of the current study will be on the specific language internal (grammatical) and language external (social) factors that influence SP variation.[15] I now move on to an in-depth discussion of the variationist sociolinguistic approach to subject expression in Spanish, the main perspective that will be adopted in the present investigation.

[15] For a preliminary discourse-pragmatic treatment of SPs in Spanish spoken in Georgia, see Limerick (2018).

3.3 Variationist approaches

Variationist sociolinguistics is a quantitatively driven field that seeks to understand and explain patterns of variation in language by exploring the effect of multiple linguistic and social factors. Variation theory maintains that language is inherently variable and is systematically conditioned by both internal (linguistic) and external (social) factors (Labov 1966). According to Labov's (1966) Principle of Accountability, to begin to understand such variation one must first carry out a qualitative analysis of the data to find cases of all the variants that are part of the grammatical system of the linguistic variable. The notion of 'linguistic variable' from a variationist perspective is conceived of as "two or more ways of saying the same thing" (Tagliamonte 2012:4), such ways being referred to as different 'variants' of the linguistic variable. According to Sankoff (1988), variants should have the same referential meaning in discourse. In the case of SPs as linguistic variables, the main variants are the overt SP and the null SP.[16] The process of locating the different variants of a linguistic variable is necessary to understand how a particular variant functions within the grammatical subsystem and not simply in isolation. In other words, one must locate and analyze not only one variant of interest (e.g. the null pronoun) but analyze all variants within the pronoun system a whole in order to understand what guides their variation, that is, how a speaker's choice between the variants is systematically governed by internal and external factors (Tagliamonte 2012:10). A subsequent key process involves determining in which contexts variation is possible between the two (or more) variants and in which contexts it is not, a process known as "circumscribing the variable context" (Poplack & Tagliamonte 1989:60). For pronoun variation, this means determining all contexts in

[16] Another important variant for subject expression more generally is the lexical subject for third-person reference (e.g. *mi mamá* 'my mom'). See Chapters 4 and 5 for further discussion of lexical subjects in the present data.

which an overt SP occurred but could have been omitted (the null SP), and also cases in which a null SP occurred, but an overt SP could have been produced.[17] Once the variable context is established, one can move forward with the quantitative analysis, which involves examining the linguistic and social factors that influence the choice of one variant over the other.[18] I will now briefly outline such influencing factors as reported in previous studies.

3.3.1 A brief overview of SP rates and factors conditioning SP variation

Regarding monolingual Spanish, a fundamental aspect of previous research (among other aspects) has been to examine the overall distribution of overt and null SPs, reporting the percentage of null SPs produced by speakers vs. the percentage of overt SPs produced, the latter percentage now commonly known as the *pronoun rate* (Otheguy, Zentella, & Livert 2007; Carvalho, Shin, & Orozco 2015). Once pronoun rates began to be established across regional varieties, it was revealed that speakers of different Spanish dialects used different rates of pronouns, thus overall overt pronoun frequencies are now commonly delineated using regional/geographical distinctions. For example, it has been demonstrated that, in general, varieties from Mexico and Spain tend to have the lowest overt pronoun rates overall (16-25% overt), South American varieties such as Argentina, Colombia, and Chile exhibit mid-range rates (30-40% overt), and Caribbean speakers (e.g. Puerto Rico, Dominican Republic) have the highest pronoun rates out of all Spanish dialects (45-55%).[19]

Some of the specific factors shown to influence SP variation across most dialects of Spanish are person/number of the verb (Silva-Corvalán 1994a), same vs. switch reference

[17] See Chapter 4 for a detailed discussion of the variable context for SPs.
[18] For explanations of the quantitative methods used in variationist research and with the current data, see Chapter 4.
[19] See, for example, Solomon (1999) (Mexico), Ranson (1991) (Spain), Barrenechea & Alonso (1977) (Argentina), Orozco (2008) (Colombia), Cameron (1993) (Puerto Rico).

(Bayley & Pease-Alvarez 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 of referent (Cameron 1992), reflexivity of the verb (Otheguy, Zentella, & Livert 2007), priming (Travis 2007), as well as social factors including age and gender (Carvalho & Child 2011).

Table 3.1: Factors influencing SP variation cross-dialectally

Factor	**Example study**
Person/number	Silva-Corvalán (1994a)
Same vs. switch reference	Bayley & Pease-Alvarez (1996)
Clause type	Morales (1997)
Tense-Mood-Aspect (TMA)	Cameron (1994)
Lexical content	Travis (2007)
Specificity	Cameron (1992)
Reflexivity	Otheguy, Zentella, & Livert (2007)
Priming	Travis (2007)
Age	Carvalho & Child (2011)
Gender	Carvalho & Child (2011)

However, there are additional influencing variables in Spanish-English contact varieties including length of residency (LOR), age of arrival (AOA), and degree of use, proficiency, and exposure to English (Otheguy, Zentella, & Livert 2007). For example, Otheguy, Zentella, and Livert (2007) found that Spanish speakers living in NYC for longer periods of time (born/raised there) had higher pronoun rates (38%) than newcomers to the city (30%), suggesting a significant influence of English on NYC Spanish as well as a dialect contact influence from high-pronoun Caribbean varieties. Moreover, Shin and Otheguy (2013) employed a broader analysis of the role of social factors on increased pronoun rates in NYC, namely social class and gender, finding that more affluent communities showed significant increases of overt SPs while less affluent communities were more resistant to such change due to loose social networks in the

former group and more dense networks in the latter. Further, the study revealed that women lead this change, which is consistent with previous research on gender and language change (e.g. Eckert & McConnell-Ginet 2003; Romaine 2003). Specifically, women showed higher overt pronoun rates than men, which the authors attributed to women's more extensive contact with second generation Latinos, including their children, who exhibit higher pronoun rates. The next section discusses the possibility of English influence on Spanish subject expression in detail.

3.3.2 Investigating English contact through pronoun use

Contact varieties sometimes show higher overt pronoun rates than monolingual varieties, which is commonly attributed to English contact, as overt SPs are nearly obligatory in English but not in Spanish. Therefore, increased exposure to English has been argued to engender higher usage of overt pronouns in Spanish (e.g. Lapidus & Otheguy 2005; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012; Abreu 2012; Shin & Montes-Alcalá 2014). However, other researchers have not observed increased overt pronouns in U.S. contact varieties (e.g. Silva-Corvalán 1994a; Flores-Ferrán 2004; Travis & Torres Cacoullos 2010; Limerick 2017). The reason for this discrepancy has been explained in terms of differential social classes and social networks (Shin & Otheguy 2013; Shin & Van Buren 2016). Specifically, Shin and Otheguy's (2013) analysis of different nationality groups in NYC Spanish revealed that more affluent speakers, who tend to have looser social networks (Cubans and Colombians), had higher SP rates than less affluent speakers with more dense social networks (Mexicans, Dominicans, and Puerto Ricans). The latter group tends to have more tight-knit social networks and consequently maintain their linguistic patterns. Shin and Van Buren's (2016) study of US-born bilingual

Mexican children in Washington/Montana[20] corroborates this argument, finding little to no change in subject expression when compared to monolingual adults. Crucially, these authors posit that since the children were part of a more tight-knit community of farmworkers and were mostly in contact with monolingual Spanish-speakers, they retained their rates and usage patterns of SPs. Drawing from Milroy and Milroy (1992), Shin and Van Buren (2016:187) also tentatively attribute changes/lack of changes in subject expression in the broader U.S. Spanish literature to rural geography and poverty, explaining that these characteristics are linked to tighter social networks. They highlight that studies in New Mexico (e.g. Torres Cacoullos & Travis 2011) as well as their own study (Washington/Montana), both of which found no evidence for contact-induced change, investigate more rural communities than those in which change has been documented, such as NYC, Chicago, and Los Angeles.

Furthermore, what constitutes evidence for language change (contact-induced or not) is debated among scholars. Poplack and Levey (2010) posit that changes in SP frequencies do not constitute evidence for change, whether it be contact-induced or otherwise. They affirm that in order to consider contact-induced change, analysts must consider the ranking of factors that are probabilistically shown to favor or disfavor SP occurrence, that is, *constraint hierarchies*. Torres Cacoullos and Travis (2010:189) agree by stating that "reliance on overall rates of use to determine contact-induced change is problematic because it is well known that regional dialects vary enormously" and also suggest that divergences in grammatical patterning across dialects are necessary to substantiate contact-induced change.

[20] Shin and Van Buren refer to the speech community's location as Washington/Montana because their participants generally reside in Washington but spend their summers in Montana. Montana is also where their interviews took place (p. 174).

Other scholars attribute changes in SP usage to English more indirectly, through a process of simplification on part of the bilingual, whereby they lighten the cognitive load of managing two linguistic systems by simplifying certain morphosyntactic or discourse constraints (e.g. Silva-Corvalán 1994b; Sorace 2004). That is, speakers may exhibit a weakening of certain constraints due to bilingualism in general, which indirectly is an influence from English in the U.S. context. Specifically, speakers have been shown to simplify pragmatic constraints on SP usage. For instance, various scholars have argued that Spanish-English bilinguals show a weakened sensitivity to switch reference whereby they exhibit an increased use of overt SPs in same reference contexts, an environment where nulls are more strongly favored in monolingual varieties (e.g. Silva-Corvalán 1994a; Shin & Otheguy 2009; Limerick 2017). This issue will be discussed in greater detail below. Other researchers, in contrast, have found no such evidence for weakened sensitivity to switch reference or to other constraints in U.S. varieties (e.g. Travis 2007; Torres Cacoullos & Travis 2011).

Aside from simplification, Shin (2014) has argued for the opposite process for bilinguals in NYC, *complexification*, whereby they show a strengthening of constraints. Specifically, Shin found that US-born bilinguals were systematically guided by tense-mood-aspect (TMA) of the verb in their use of SPs (particularly third-person singular forms), but that first generation newly arrived immigrants were not. In other words, TMA was not a statistically significant predictor of pronoun variation for immigrants but emerged as significant for US-born speakers. Shin (2014:304-305) interpreted such emergence as a process in which the grammar of US-born speakers became more complex (adding an additional significant factor), known as grammatical complexification. Furthermore, she argues that this process has potentially occurred in tandem with grammatical simplification (discussed above), whereby the complexification or increased

sensitivity to one factor (TMA) is possibly an effort to compensate for the simplification or decreased sensitivity to another factor (e.g. switch reference). Other studies that have found a strengthening of TMA include those of Otheguy and Zentella (2012) and Torres Cacoullos and Travis (2011). Finally, an additional contact feature reported in the literature is the increased use of overt SPs for nonspecific third-person plural subjects (*ellos*) (Lapidus & Otheguy 2005). Lapidus and Otheguy (2005) found that New York born and/or raised speakers used these forms significantly more often than both newcomers to the city and established immigrants, suggesting contact influence from English among the New York born and/or raised.

In light of the above discussion thus far on the investigation of English contact, then, what would (non) contact effects look like in the current data? In other words, what role, if any, does contact play in the variation/change concerning pronoun usage? Contact effects for SP use in the data would include the following elements:

a. Substantially increased SP frequencies when compared to monolingual varieties (Otheguy & Zentella 2012)
b. Differences in constraint rankings when compared to monolingual varieties (Otheguy & Zentella 2012)
c. A simplification/weakening of constraints, such as switch reference (Shin & Otheguy 2009)
d. A complexification/strengthening of constraints, such as TMA (Shin 2014)
e. Increased use of overt nonspecific *ellos* (Lapidus & Otheguy 2005)

Non-contact effects, conversely, would include the following:

a. Decreased, equal, or marginally increased SP frequencies when compared to monolingual varieties (Flores-Ferrán 2004)

b. No differences in constraint rankings when compared to monolingual varieties (Torres Cacoullos & Travis 2010)
c. No evidence of simplification/weakening of constraints
d. No evidence of complexification/strengthening of constraints
e. Decreased, equal, or marginally increased use of overt nonspecific *ellos*

These types of evidence for/against contact-induced change in the present data will be explored in Chapter 6. Next, I discuss in further detail overall pronoun rates and conditioning factors of pronoun variation reported in the variationist literature.

3.3.3 Overall pronoun rates

As a point of reference and for comparison, I present below the reported overall overt pronoun rates for monolingual varieties (see Table 3.2) as well as Spanish-English contact varieties (see Table 3.3). It is important to note the cross-dialectal variability in rates among monolingual Spanish as well as between monolingual and contact varieties.

Table 3.2: Pronoun rates of monolingual varieties

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

Table 3.3: Pronoun rates of contact varieties

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

Regarding contact Spanish, one can see from a comparison of Tables 3.2 and 3.3 that SP rates are sometimes higher, but sometimes equal or lower than rates for monolingual varieties, as highlighted above. For instance, higher rates were observed for New York Dominicans (48%) than were for Dominicans in Santiago (39%). Likewise, New York born/raised Spanish-speakers showed an increase (38%) when compared to New York Newcomers (30%). In contrast, we see in other cases that pronoun rates are virtually the same in contact varieties when compared to monolingual varieties, for instance in the case of North Carolina (17%) and Georgia Mexicans (21%) when compared to Yucatan (16%) and Mexico City (22%). These rates will be revisited in Chapters 5 and 6 in order to discuss where the pronoun rate for the current data fits with respect to other varieties in the broader literature and, crucially, with respect to monolingual Mexican varieties.[22] I will now discuss the linguistic and social variables that have been shown to exert a significant influence on subject expression in previous sociolinguistic research.

[21] This was calculated as the average rate among all three generational groups from Silva-Corvalán (1994a:153).
[22] The data collection for studies of monolingual Mexican Spanish involved sociolinguistic interviews, and the pronoun rates were calculated from a total of 18 speakers in the case of Mexico City (Lastra & Martín Butragueño 2015) and 10 speakers in the case of Yucatan (Michnowicz 2015).

3.4 Linguistic variables

3.4.1 Person/number

In most Spanish varieties, both person and number of the verb highly influence SP variation. In fact, person/number has been shown to be the strongest predictor of variable subject expression cross-dialectally (Orozco 2015). Specifically, first-person singular and third-person singular verbs tend to favor overt pronouns (e.g. Silva-Corvalán 1994a; Flores-Ferrán 2002; Shin 2012; Lastra & Martín Butragueño 2015; Limerick 2017). In fact, most studies have found that all singular forms in general are more likely to appear with overt SPs compared to plural forms, which tend to favor null SPs (Orozco 2015). Moreover, regional differences have sometimes been reported. For instance, Otheguy, Zentella, and Livert (2007) found that Caribbean newcomers to NYC preferred overt SPs with second-person singular verbs compared to third-person singular verbs, but that South American newcomers preferred overt SPs with third-person singular verbs. The general finding for singular and plural persons/numbers and their influence on subject expression has been explained by Cameron (1993) in the following way:

> If we conceive of plural subjects as sets, we find that discourse is typically structured so that the great majority of plural subjects occur in contexts where their set members are either explicitly or inferably present within the immediately preceding discourse. Such contexts favor null subject expression. Therefore, plural subjects are frequently null overall (p. 328, note 2)

Furthermore, the widest distinction generally found for person/number is that first-person singular forms (*yo*) favor overt SPs the most (as stated above) while first-person plural forms (*nosotros*) strongly disfavor overt SPs. Regarding the former pattern, the high use of *yo* has been attributed to the "egocentric nature of discourse" (Posio 2011:795). Similarly, as Morales (1986)

explains, by explicitly making reference to herself, the speaker's pragmatic need to stay overtly present in the conversation is fulfilled. Regarding the latter pattern, it could be the case that *nosotros* is rarely expressed due to its comparatively longer length as well as the fact that its verbal inflection is the least ambiguous (Bentivoglio 1987). Similarly, Orozco and Guy (2008) suggest that since it has the most morphologically distinct and regular verb forms, the use of *nosotros* would be redundant. This is particularly convincing given that the first-person plural verbal inflection (*-mos*) stays the same for all TMAs in comparison to first and third-person singular forms which become indistinct (i.e. morphologically ambiguous) in the imperfect, subjunctive, and conditional. Yet another explanation is given by Orozco and Guy (2008): the regular omission of overt SPs with certain fixed expression (e.g. *vamos* 'let's go') could have expanded and generalized to other first-person plural forms.

3.4.2 Switch-reference

Switch-reference, which considers same vs. switch reference from one subject to another, has also shown a strong influence on the manifestation of SPs cross-dialectically (e.g. Bentivoglio 1987; Cameron 1994; Silva-Corvalán 1994a; Bayley & Pease-Alvarez 1997; Travis 2005; Prada Pérez 2009; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015). Specifically, where there is a switch in subject referent, the SP is often overt, as seen in (14); when there is no switch, null SPs are preferred, as in (15).

(14) *no <u>están</u> esperando a que **tú llegues*** [F28Mex]

'they aren't expecting you to arrive'

(15) *<u>yo me relaciono</u> mu- muy mucho con mi hermanito y y y s- ...(1.5) **Ø paso** mucho tiempo allí* [M27Mex]

'I relate ve- very much with my little brother and and and s- …(1.5) I spend a lot of time there'

This pattern is generally thought to have a functional influence that has to do with referential tracking (Shin & Otheguy 2009). As Cameron (1994:40-41) explains, "expressed pronominal subjects compensate for the change of information state which occurs with a switch in subject reference." In other words, overt SPs tend to be used in such contexts in order to facilitate interpretation of the antecedent for the listener. According to Shin and Otheguy (2009), this is especially important for third-person referents for disambiguating between *él* and *ella*, as in the following example from their data:

(16) *Ella tenía su novio allá y **él** pensaba venir pero no le dieron la visa* (p. 120, emphasis mine)

'She had her boyfriend there and he was planning on coming, but they didn't give him the visa.'

These authors argue that since there are competing referents in (16) (*ella* and *su novio*), the use of *él* helps the listener to track the antecedent of *pensaba*.

Interestingly, as noted above, certain contact varieties have shown a weakening/simplification of this constraint (Flores-Ferrán 2002; Shin & Otheguy 2009; Michnowicz 2015). For example, Flores-Ferrán (2002) found that NYC-born Spanish speakers exhibit an increased use of overt SPs in same reference contexts, demonstrating their loss of

sensitivity to null pronoun usage in such contexts. Similarly, such non-canonical patterns are found in the present study, as illustrated in (17) and (18):

(17) *él sufrió ese cambio donde* ***él***, ***se sentía*** *que era de aquí pero a la vez no era de aquí* [F56Mex]

'he went through that change where he, felt that he was from here but at the same time he wasn't from here'

(18) *yo vivía aquí,* ***yo me iba*** *a Alpharetta, caminando,* ***yo me iba*** *a Sandy Springs* [M32Mex]

'I lived here, I would go to Alpharetta, walking, I would go to Sandy Springs'

Furthermore, some researchers also include an additional category, *partial switch*, which analyzes cases of switch in subject where the subject is coreferent with the immediately preceding object (direct/indirect object, object of preposition, etc.), as in (19).

(19) *…en México me inculcó mi mamá que* ***tenía*** *que ir a la iglesia…* [M52Mex]

'…in Mexico my mom instilled in me that I had to go to church…'

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

3.4.3 Tense-mood-aspect (TMA)

The tense-mood-aspect (TMA) of a verb has also been shown to condition SPs. Certain TMAs favor overt SPs while others favor nulls. For instance, Silva Corvalán (1982) found that imperfects and conditionals favor overt SPs while presents and preterits are more likely to appear with nulls, as was also found by other researchers (e.g. Cameron 1994; Travis 2007; Carvalho & Bessett 2015). To explain such correlations, it has been proposed that imperfects and conditionals favor overt SPs due to their potential ambiguity. Since their first and third-person singular verb forms are morphologically indistinct, the use of overt SPs would serve to disambiguate the referents of such forms, an explanation that forms part of the Functional Hypothesis (Hochberg 1986). However, other studies have found no such correlation (e.g. Enríquez 1984; Bentivoglio 1987; Ranson 1991).[23]

An alternative explanation for the TMA effect, one that is not related to ambiguity, was been proposed by Silva-Corvalán (2001). She discusses imperfects and preterits in relation to discourse functions, suggesting that more overt SPs are used with imperfects due to the backgrounded nature of imperfect aspect, and that fewer overt SPs are used with preterits since they tend to foreground events. That is, for events that are more backgrounded the focus is on the subject (hence the use of an overt SP) whereas the focus is more on the action with more foregrounded events (hence the lack of an overt SP). However, this hypothesis has been rejected in recent work. For instance, Shin (2014) argues rather for the aforementioned ambiguity explanation, finding that imperfects favored overt SPs particularly in contexts of switch reference and with competing referents, both contexts in which referential tracking is more

[23] See further discussion in the *Morphological ambiguity* section directly below.

difficult. Given her findings, Shin argues that her analysis lends no support to Silva-Corvalán's proposal.

3.4.4 Morphological ambiguity

Similar to, but more specific than TMA, many researchers have shown that, when examining only particular verb forms that are morphologically ambiguous (as opposed to TMAs as general categories), ambiguous verb forms favor overt SPs while unambiguous forms prefer nulls (e.g. Hochberg 1986; Travis 2007; Prada Pérez 2009; Erker & Guy 2012; Lastra & Martín Butragueño 2015; Michnowicz 2015). Ambiguous verb forms are those that have indistinct person inflections in the first and third-person singular, specifically in the imperfect, subjunctive (both present and past), conditional, and pluperfect. In contrast, unambiguous forms have distinct inflections and include all other TMAs (e.g. present indicative, preterit, present perfect, etc.). The idea is that since ambiguous forms do not distinguish between first and third-person, speakers use overt SPs to clarify the referent for the listener. The referent of *tenía*, for instance, could be *yo*, *él*, *ella*, or *usted*, but this distinction may be unclear without an overt SP. However, some scholars have found the opposite effect and argue that true ambiguity is rare and that it is the context that clarifies the particular antecedent. For instance, Ranson (1991) actually found fewer overt SPs with ambiguous verb forms relative to unambiguous forms and posits that contextual markers such as previous mention of the referent or background knowledge help to clarify the speaker's intended referent when the SP is null. Further, she argues that a lack of such contextual markers may play a more important role in explaining the use of overt SPs, as she found that overt SPs were more frequently used when no contextual markers were present, thereby compensating for such lack of context. Ranson's analysis contradicts the Functional

Compensation Hypothesis (Hochberg 1986), which predicts that the use of overt SPs will increase with morphologically ambiguous verb forms.

3.4.5 Verb class

Several researchers have found that verb class (i.e. lexical content) can also determine how a subject pronoun is manifested (Bentivoglio 1987; Silva-Corvalán 1994a; Travis 2007; Otheguy & Zentella 2012; Orozco 2015). In general, it has been noted that verbs of psychological/mental activity (e.g. *creer* 'believe', *pensar* 'think'), verbs of communication (e.g. *decir* 'say', *hablar* 'speak'*)*, and copulas (e.g. *ser*, *estar* 'be'*)* tend to be expressed with overt pronouns, with psychological verbs showing the highest probability. On the contrary, motion verbs tend to disfavor overts (Bentivoglio 1987; Silva-Corvalán 1994a; Travis 2007). Regarding psychological verbs, it has been hypothesized that overts are frequently used because these verbs tend to express the point of view of the speaker and because of the implied contrastive function that is often carried out in such contexts (Silva-Corvalán 1994a). The speaker "asserts their role in the utterance" by using an overt SP (Travis 2007:117). With regard to the preference for overt SPs with communication verbs, particularly *decir*, Travis (2007:117) has postulated that this is perhaps related to the epistemic function of *decir* to express an opinion (e.g. *yo digo que* 'I say that'), similar to the aforementioned effect for psychological verbs. Additional categories have also been employed for verb class, such as stative and activity verbs (e.g. Orozco & Guy 2008; Erker & Guy 2012; Otheguy & Zentella 2012; Orozco 2015). These studies have generally found that stative verbs favored overt SPs while activity verbs were more likely to appear with nulls.

3.4.6 Verbal mood

The specific mood in which the verb appears and its influence on subject expression has not been extensively studied in previous variationist research aside from the widely-studied

factors of TMA and morphological ambiguity discussed above (the only exception, to my knowledge, is Lastra and Martín Butragueño 2015). The subjunctive mood has been examined in relation to specific tenses and aspects, as we have seen, but not so much in direct relation to the indicative and imperative moods in general.[24] Lastra and Martín Butragueño (2015) have begun studying this variable in Mexico City Spanish, finding that non-indicative moods, which included the subjunctive and imperative in their study, disfavored overt SPs. In fact, imperatives were categorially null in their data, which is not unexpected given the very strong disfavoring of overt SPs with imperatives reported across studies.[25] Regarding the indicative mood, while showing a higher frequency of overt SPs compared to imperatives, it showed neither a statistically significant favoring nor disfavoring effect on subject expression. These researchers suggest that the disfavoring effect for subjunctive forms is connected to the subjunctive's typical subordinate function. However, previous literature that examines the effect of subordinate clause types on SP variation generally shows that, rather than disfavoring overt SPs, they either favor them or show a neutral effect (e.g. Morales 1997; Shin & Montes-Alcalá 2014; Orozco 2015; Limerick 2017). In addition, Lastra and Martín Butragueño (2015:46) link this subordinate function to the "likely short distance of co-referring previous noun phrases", which would thus promote null SPs. Nevertheless, they provide no evidence or explanation as to why pronouns in subordinate clauses are likely to have close antecedents. From the perspective of the current

[24] The analysis of the subjunctive only in relation to specific indicative forms (e.g. preterit, imperfect, present, etc.), and not indicative forms as a whole, may obscure patterns that emerge from simply examining subjunctive vs. indicative forms overall. This is because variation has been widely observed within such indicative forms, such as the favoring of overt SPs with the imperfect but a disfavoring of overt SPs with the preterit (e.g. Silva-Corvalán 1982; Cameron 1994; Travis 2007; Carvalho & Bessett 2015). Thus, it is of interest to understand the patterns of SP variation when considering the subjunctive and indicative moods more generally.

[25] Non-indicatives continued to show a disfavoring effect even when imperatives were removed (Lastra & Martín Butragueño 2015:45). Thus, imperatives alone were not driving this pattern and the subjunctive continued to disfavor overt SPs.

analysis, an alternative explanation may be found in connecting the subjunctive with the null subject behavior of imperative moods. Specifically, the subjunctive's imperative-like nature in certain contexts may promote null SPs (e.g. *Quiero que* ***Ø estudie*** *inglés* 'I want him to her/him to study English').

3.4.7 Specificity

The use of a specific referent, as opposed to a more general or nonspecific referent, also tends to condition whether or not a pronoun is overt or null. The following two examples from the Roswell data illustrate this distinction between specific and nonspecific with regard to the subject *tú*:

(20) *...no sé si lo* ***conozcas*** *o escuchaste hablar de él* [F28Mex]

'…I don't know if you are familiar with it or you heard people talk about it'

(21) *...porque* ***vienes*** *sin conocer a nadie* [M32Mex]

'…because you come without knowing anyone'

In (20), the second-person singular form is used by the speaker to refer specifically to the listener; however, in (21) the subject referent *tú* is referring to a nonspecific entity ('one'), (i.e. the *impersonal tú*), where the speaker is talking about one coming to Georgia and is not referring to the interviewer in particular.

Previous research has shown that the conditioning effect of specificity on SP variation can depend on the dialect. For example, Cameron's (1992, 1993) work on San Juan and Madrid Spanish demonstrates that non-specific subject referents favored overt SPs in San Juan while specific referents favored overt SPs in Madrid. Regarding Mexican Spanish, both Solomon (1999) and Michnowicz (2015) found that, like Madrid Spanish, overt SPs were favored by specific referents. However, specificity is not found to be significant in all dialects (see, e.g.,

Otheguy, Zentella, & Livert 2007; Lastra & Martín Butragueño 2015); for instance, Otheguy, Zentella, and Livert (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. Additionally, Lastra and Martín Butragueño (2015) did not find a significant effect for specificity.[26]

Specificity also applies to third-person plural subjects (*ellos/ellas*) and has been discussed in this regard in some studies (e.g. Lapidus & Otheguy 2005; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012). Specific vs. nonspecific reference for third-person plural subjects has been defined in terms of the presence vs. absence, respectively, of an explicit antecedent in the linguistic context (see Otheguy & Zentella 2012:255). Although nonspecific *ellos* tends to be categorically null (e.g. ***Ø Dicen*** *que es bueno* 'They say that it's good'), and thus excluded from many variationist analyses, Lapidus and Otheguy (2005:164), for example, have shown that, despite it being relatively rare, speakers produced it overtly in 107 out of 2,834 instances. They also reported a significant increase of these forms among NYC-born speakers compared to newcomers, which they interpreted in support of an English contact hypothesis.

3.4.8 Polarity

Polarity considers affirmative vs. non-affirmative clauses and their conditioning on SP variation. While still relatively understudied, this variable has occasionally been analyzed in previous research, with some researchers finding a significant effect (Lastra & Martín

[26] Such lack of effects refers to specificity when considered overall, and not in relation to particular persons/numbers (*tú* or *ellos/ellas*). Otheguy, Zentella, and Livert (2007), as well as Otheguy and Zentella (2012), did in fact find a significant effect for specificity when considering its interaction with person/number. For instance, the former study found that, among Mainland newcomers, nonspecific *tú* disfavored overt SPs whereas specific *tú* showed little effect.

Butragueño 2015; Geeslin & Gudmestad 2016) and others reporting a lack of effect (Travis & Torres Cacoullos 2012; Torres Cacoullos & Travis 2015).[27] For example, Lastra and Martín Butragueño (2015) found that non-affirmatives (negatives and interrogatives) disfavored overt SPs and that affirmatives showed a neutral effect. These authors give a possible explanation for this pattern in terms of the frequent clustering together of negative clauses in their data, which tend to be coreferential. They hypothesize that "if negated clauses cluster together, it is possible that co-reference across these negated clauses contributes to their disfavoring effect on overt SPPs" (Lastra & Martín Butragueño 2015:46). In other words, there could be an interaction between the switch reference and polarity variable such that negative clauses disfavor overt SPs more frequently in coreferential contexts than in switch reference contexts.

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

Furthermore, Travis and Torres Cacoullos (2012), while not finding a significant effect for polarity overall, did find effects upon considering particular verb classes/lexemes. Their study revealed interesting patterns in the distinction between cognitive and non-cognitive verbs and their influence on polarity, namely that negation favored overt SPs, but only for non-

[27] Some differences in terminology and methods among these studies should be noted: Lastra and Martín Butragueño (2015:46) use the term *enunciative type* to refer to polarity (positive vs. negative), but with the addition of interrogative tokens being included in a *non-affirmative* category. Other researchers cited in this section use the terms *polarity* (Travis & Torres Cacoullos 2012) and *verbal negation* (Geeslin & Gudmestad 2016) and simply use the classification of affirmative vs. negative sentences. See Chapter 4 for further methodological details.

cognitive verbs. Moreover, although a significant effect for cognitive verbs was not found overall, the researchers did find an effect when considering only the particular verbal lexeme *creo*, specifically that negation highly disfavored overt SPs (e.g. ***Ø no*** *creo*) (Travis & Torres Cacoullos 2012:741, note 23).

To summarize, the findings across studies for polarity have been somewhat conflicting, with some finding it to be significant and others lacking an effect. Additionally, the general finding that negation disfavors overt SPs (Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016) may only apply to cognitive verbs, based on evidence from Travis and Torres Cacoullos's (2012) analysis. The present study intends to study this variable further, including its potential interaction with other independent variables such as switch reference, person/number, and verb class.

3.4.9 Style

The linguistic style of the speaker has also been found to constrain subject expression (e.g. Ávila-Jiménez 1996; Lastra and Martín Butragueño 2015). Ávila-Jiménez (1996), for instance, reported that more casual speaking favored overt SPs in Puerto Rican Spanish. Similarly, Lastra and Martín Butragueño (2015) found that Mexico City speakers tended to use more overt SPs for more conversational styles, particularly when they had longer speech turns in the latter part of their interviews, which also contained a higher use of argumentation, as opposed to shorter speech turns at the beginning of the interview (less conversational), for which they used fewer overt SPs. Although these authors do not explain possible reasons for such patterns, I suspect that it is related to more careful speech being used near the beginning of the interview and less careful (more casual) speech as the interview progresses. Careful speech styles, being subject to more self-monitoring and perhaps focus on using more prescriptive grammar norms,

would be more likely to comprise null pronouns (recall the above discussion in Section 2.3 on traditional grammarians' views on the overuse of overt SPs). More casual speech, on the other hand, without such restrictions, would be more likely to employ overt SPs.

Speech style is relatively understudied, however, which motivates further analysis as part of this dissertation. Particularly, the call from Lastra and Martín Butragueño (2015:49) for future research to investigate the interaction between style with person/number is of interest in the present study. Their suggestion to examine this interaction is based on opposite patterns for style being found in previous research that included only first-person singular SPs (e.g. Blanco Canales 1999; Travis 2007) while their analysis included all persons/numbers.

3.5 Additional variables[28]

3.5.1 Clause type

Another factor conditioning pronoun variation is the type of clause in which the SP appears. For example, Morales (1997) found that overt SPs are more likely to occur in subordinate clauses—particularly object relative clauses—than they are in main clauses. In addition, Otheguy and Zentella (2012) found that main clauses favored overts while coordinate clauses favored null SPs for NYC Spanish.

3.5.2 Reflexivity

Whether a verb is used reflexively or not is also a conditioning factor of subject expression that has been discussed in the literature (e.g. Otheguy, Zentella, & Livert 2007; Carvalho & Child 2011). For example, Carvalho and Child (2011) found that verbs that contain a reflexive pronoun disfavor overt SPs whereas verbs without a reflexive pronoun neither favor

[28] Variables in this section include those that have been analyzed in previous SP research, but which are not included in the current study. See Chapter 4 for the motivations regarding the choice of variables included.

nor disfavor overts. According to these researchers, "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" (Carvalho & Child 2011:20).

3.5.3 Priming

Priming takes into account whether the previous verb's SP was overt or null. Priming is defined by Travis (2007) 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" (p. 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 SP in a preceding clause. Likewise, it has been shown that null SPs exhibit the same effect, whereby the use of a null pronoun primes the use of a subsequent null pronoun (e.g. Cameron 1994; Flores-Ferrán 2002; Travis 2007).

3.5.4 Lexical frequency

Finally, the lexical frequency of verbs and its effect on pronoun variation has been explored in the literature (e.g. Erker & Guy 2012; Bayley, Greer, & Holland 2013). Erker and Guy (2012) measured this variable discretely, considering frequent versus infrequent verbs used by speakers and their effect on SP use. Rather than observing a direct influence on SP production, the researchers found that lexical frequency more directly interacted with other independent variables (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 their effects. For instance, they found that first and third-person verb forms of higher frequency were associated with a decrease of overt SPs whereas second-person high frequency forms were associated with an increase of overt usage. Moreover, Bayley, Greer, and Holland (2013) also

studied lexical frequency, but, unlike Erker and Guy (2012), did find a direct connection between frequency and subject expression. Specifically, they found a favoring of overt SPs with infrequent verbs and a disfavoring of overts with frequent verbs. Despite finding this significant effect, Bayley, Greer, and Holland highlight that the effect was relatively small and that other factors were much more important, such as person/number and switch reference.

Table 3.4 below summarizes the general findings for each of the linguistic variables discussed above. The following section will discuss the influence of social variables on subject expression attested in the literature.

Table 3.4: Summary of general findings for linguistic variables

Variable	**Overts favored**	**Example study**
Person/number	Singular verbs	Otheguy & Zentella (2012)
Switch reference	Disjoint reference	Bayley & Pease-Alvarez (1997)
Tense-Mood-Aspect (TMA)	Imperfect, conditional	Travis (2007)
Morphological ambiguity	Ambiguous forms	Prada Pérez (2009)
Verb class	Psychological verbs	Silva-Corvalán (1994a)
Verbal mood	Indicative	Lastra & Martín Butragueño (2015)
Specificity	*Mexico, Spain:* specific reference *Puerto Rico:* nonspecific reference	Michnowicz (2015), Cameron (1992) Cameron (1992)
Polarity	Affirmative	Geeslin & Gudmestad (2016)
Speech Style	Casual	Ávila-Jiménez (1996)
Clause type	Main clauses	Otheguy & Zentella (2012)
Reflexivity	Non-reflexive verbs	Carvalho & Child (2011)
Priming	Previous overt SP	Travis (2005)
Lexical frequency	Infrequent verbs	Bayley, Greer, & Holland (2013)

3.6 Social variables

3.6.1 Age

Age is one of the main social variables that has been shown to influence SP usage. Carvalho and Child (2011) found a favoring of null pronouns among younger speakers in Uruguay (ages 16-29) while older speakers favored overt SPs, a trend they hypothesize is influenced by the Portuguese substratum (p. 23). Similarly, Orozco and Guy (2008) observed that older Colombians favored overt SPs while adolescents favored nulls, which they explain as a possible change in progress attributed to greater access to education by young people compared to prior generations. They also hypothesize that these results are due to influence of the Colombian highlands dialect. More recently, Lastra and Martín Butragueño (2015) found the same effect whereby older speakers favored overt SPs and younger speakers were more likely to produce nulls. These researchers raise the issue of a possible change in progress toward less overt pronoun use in Mexico City. 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. In the U.S., younger speakers may be using more overt SPs due to English influence. However, previous research on U.S. contact varieties does not show a direct correlation between age and pronoun use (e.g. Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012). This issue will be further explored in Chapter 5.

3.6.2 Gender

Numerous authors have found that females and males differ in their SP frequencies such that women use more overt SPs than men (e.g. Bayley & Pease-Alvarez 1996; Solomon 1999; Carvalho & Child 2011; Otheguy & Zentella 2012; Shin & Otheguy 2013; Alfaraz 2015), while

others found gender to be nonsignificant (e.g. Orozco & Guy 2008; Lastra & Martín Butragueño 2015; Michnowicz 2015). For instance, in their study of Uruguayan Spanish in contact with Portuguese, Carvalho and Child (2011) found that females favored overt SPs whereas males favored nulls. Similarly, Shin and Otheguy (2013) observed a higher rate of overt pronouns among women in NYC, which they attribute to women's extensive contact with U.S. born bilinguals. Bayley and Pease-Alvarez (1996) hypothesized that women use narrative discourse more often than men and thus more SPs, due to the fact that this type of discourse tends to favor the use of more overt SPs. For Mainlanders in NYC, the finding that women were more likely to use overt SPs than men reflects a change in progress (Otheguy & Zentella 2012: 118-120). Moreover, Alfaraz (2015) suggests that this gender pattern, particularly in the Dominican Republic, is linked to prestige.

3.6.3 Educational level

A relatively understudied variable, educational level has occasionally been found to condition subject expression whereby less educated speakers favor overt SPs while those with more education favor nulls. For example, Ávila-Jiménez's (1996) study of Puerto Rican Spanish found that overt SPs were disfavored among the college educated and favored among those with lower education levels. The few additional studies that have examined this variable (Otheguy & Zentella 2012; Lastra & Martín Butragueño 2015), while not finding a statistically significant effect, have reported trends consistent with Ávila-Jiménez's (1996) analysis. That is, they observed higher overt SP rates for those with lower education levels.[29] Lastra and Martín Butragueño (2015) suggest that this pattern could indicate that null SP use has more prestige. In

[29] Orozco (2015) also analyzed education level and found a lack of statistical significance, but he did not report the SP frequencies for each level of education.

a similar vein, such a pattern could be related to prescriptive grammar norms for SPs (which presumably are taught in schools), which emphasize to only use overt SPs when necessary (e.g. for emphasis or contrast) and to avoid the overuse of pronouns (Butt & Benjamin 2004). Thus, speakers with higher education levels may be more likely to follow such norms, using fewer overt SPs.

3.6.4 Socioeconomic status (SES)

Socioeconomic status (SES) is another understudied variable in relation to pronoun use (notable exceptions are Otheguy and Zentella 2012, Alfaraz 2015, and Orozco 2015). While neither Orozco (2015) nor Alfaraz (2015) have found an effect for SES, Otheguy and Zentella (2012) found that speakers with lower SESs favored overt SPs while higher SES speakers favored null SPs among the Mainland population in NYC. Otheguy and Zentella used a combined measure of education and occupation to examine this variable (see Chapter 3 for a detailed explanation of their operationalization). Although they do not explain what may have motivated such effect, I presume that educational level, being part of measuring SES, partially explains this result as speakers with lower education levels were found to favor overt SPs.

3.6.5 More on assessing English contact effects: Length of residency (LOR), age of arrival (AOA), and English proficiency

As mentioned above, other important factors to consider regarding potential influence of English on Spanish are LOR in the U.S., suggesting greater exposure to English, and AOA to the U.S., indicating speakers' susceptibility to change. Drawing from Veltman (2000), Otheguy, Zentella, and Livert (2007:779) state that older arrivals who speak Spanish more frequently and fluently should show greater resistance to English influence. In addition to finding that longer LORs in NYC correlated with higher frequencies of overt SPs, the authors found that younger

arrivals, being more susceptible to English influence, exhibited higher pronoun rates than older arrivals, who tend to use less English and whose Spanish is more proficient, thereby making them less susceptible to the influences of language contact.

Additionally, English proficiency has been examined as an independent social variable in U.S. contact varieties of Spanish, with some researchers finding a positive correlation between English proficiency and overt pronoun use (e.g. Otheguy & Zentella 2012) and others finding no such correlation (e.g. Silva-Corvalán 1994a). Another factor to explore for examining contact effects, which has not been addressed in previous SP studies to my knowledge, is preferred media language. It could be the case, for example, that those who prefer English for media exhibit a higher use of overt SPs than those who prefer Spanish. This factor will be explored in this dissertation (see Chapter 4 for a methodological discussion and Chapter 5 for the findings related to this factor).

3.7 Summary

This chapter has addressed issues that have been studied from multiple perspectives in the literature on subject expression in Spanish, including the main factors that promote the presence or absence of overt SPs by Spanish-speakers. The first of these perspectives was traditional grammar. From this view, overt SPs are to be used to make a contrast among subjects, emphasize subjects, or to resolve potential ambiguity of the subject referent. Traditional grammarians also emphasize that overt SPs should not be overused; that is, they should not be used when there is no intention to make a contrast, emphasize, or clarify the referent. The second perspective addressed in the chapter was that of discourse-pragmatics. Researchers from this view also highlight the emphatic and contrastive uses of SPs, but go far beyond these functions in their analyses. For example, it is argued that overt SPs can be used as hedging devices or to

add pragmatic weight to an utterance. Moreover, certain constraints such as the salience and accessibility of the referent, a switch in focus from one referent to another, and the use of epistemic parentheticals or fixed expressions can also guide the use of SPs. Regarding the discourse-pragmatics perspective, one methodological issue was raised in the chapter, namely that of a lack of objective measures for determining what constitutes each of the aforementioned constraints upon analyzing data.

Finally, the variationist sociolinguistic perspective was discussed, which is the one being applied in the current investigation. Variationist theory and its application to the study of subject expression was introduced. Additionally, the frequency of use of overt pronouns as well as the linguistic and social factors that govern SP variation was addressed. These factors included, among many others, the person/number of the verb, a switch in subject referent, TMA, verb class, polarity, and the speaker's gender, age, and socioeconomic status. The study of SPs in Spanish-English contact situations was also discussed. Various studies were reviewed that had the objective of assessing the role of English and its possible influence on Spanish regarding pronoun use, both in terms of pronoun rates and constraints. It was highlighted that there are conflicting findings in the literature, with some finding increased pronoun rates and/or differential constraints among speakers with more English exposure, and others finding little difference between contact and non-contact varieties. The next chapter will present in detail the methodology employed in this dissertation to analyze subject expression in Roswell including the fieldwork conducted as well as the qualitative and quantitative analyses carried out.

CHAPTER 4

METHODOLOGY

This chapter presents the methodological aspects of the present investigation. First, I discuss the data collection process, the nature of the data employed in this dissertation, and the sociodemographic information of the participants. Next, I outline how the data were analyzed, beginning with the qualitative analysis of the interviews and the circumscription of the variable context for subject expression, and subsequently presenting both the independent variables included in the analysis and their operationalization. Finally, I discuss the nature of the statistical analyses carried out as well as my predictions for the results of the study.

4.1 Data collection

During the summer of 2015, I conducted sociolinguistic interviews with Spanish speakers who were either living or working in Roswell, GA at the time of (and prior to) the interviews. Sociolinguistic interviews 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). The sample of speakers for the present investigation consists of 20 first-generation Mexican immigrants (see section 4.2 below for further details concerning the speakers).

Regarding the recruitment of my participants, the majority took place at a local plaza containing several Latino businesses, including a grocery store, clothing stores, barber shops, jewelry stores, restaurants, and many others (see Figures 4.1 and 4.2 below). In addition, I recruited participants at a local church that had a Spanish ministry and at an ESL school. The

process of recruitment was different at each place. At the Latino plaza, I went into each business and introduced myself to the owners and employees, briefly explaining my project and that I was interested in raising awareness of the diversity of languages and cultures in Roswell. I explained to them that I was carrying out interviews with people in the Latino community, and I asked them if they would be interested in participating. There were three main outcomes: an acceptance and scheduling of the interview, an acceptance and completion of the interview at the time of recruitment, or a rejection. The interviews mostly took place at the participants' respective places of business. In other cases, I approached people who were outside the plaza, and we proceeded with the interviews in the parking lot.[30] For recruitment at the church, I first scheduled a meeting with the pastor to discuss my project and objectives and he offered to announce it to the congregation the following service. I attended the service and afterwards spoke with those interested in the project and scheduled interviews. Some interviews took place at the church and others at local restaurants. Finally, for recruitment at the ESL school, I scheduled meetings with the office administrator and the director to discuss my project. They then explained the details of my study to the students and provided me with names and contact information for those interested. I then contacted each student to schedule interviews, which took place at the school.

The above recruitment processes involved a combination of sampling methods that have been used in sociolinguistics research: (a) judgment sampling (also known as "snowball" or "friend-of-a-friend" sampling) and (b) random sampling. Judgment sampling, which involves utilizing the researcher's community contacts, who then recommend additional participants (Hoffmon 2013), is the method I employed both at the church and at the ESL school. This is the

[30] Though not an ideal location, finding an adequate place to conduct interviews is one of the many difficulties researchers encounter in the process of collecting sociolinguistic data (see, e.g., Tagliamonte 2006:17-27 and Hoffmon 2013 for discussions regarding fieldwork challenges).

most common method for sociolinguistic interviews (Hoffmon 2013). On the other hand, random sampling involves seeking out people unknown to the researcher (or the researcher's contacts) to ask them for an interview, such as going door-to-door. Although used in classic sociolinguistic studies such as Labov's earlier work in New York City (Labov 1972, cited in Tagliamonte 2012:103), it is not typically utilized in current studies or is modified in some way (see Tagliamonte 2012:100-104). For the current study, it was modified in order to specifically find Spanish-speakers. This was the case for the majority of interviews conducted in the Latino plaza since I focused on a part of the city with a large proportion of Latinos and simply approached them without being referred to them by other contacts. This combined method of using social networks (i.e. judgment sampling) as well as random sampling has also been used in previous research, such as the compilation of Tagliamonte's Toronto English Corpus (Tagliamonte 2003-2006, cited in Tagliamonte 2012:103).

Figure 4.1: Roswell Latin Plaza

Figure 4.2: Quinceañera store (inside Roswell Latin Plaza)

The selection of speakers for interviewing was based on the requirements that they be native Spanish-speakers, at least 18 years old, and that they lived or worked in Roswell.[31] The interviews were recorded using a TASCAM audio recorder and a head-mounted microphone, and they lasted between 30 minutes and one hour. They were informal, conversational, and addressed topics of personal history, local community life, differences between the speakers' home countries and the U.S., and experiences adapting to life in Roswell, among others,[32] following the interview protocol from the Roswell Voices Project (Kretzschmar et al. 2007).[33]

[31] Due to the specific restrictions of my IRB protocol, I did not interview anyone under 18 years of age.
[32] For further discussion of sociolinguistic interview methods, see, e.g., Tagliamonte (2006, 2012) and Hoffmon (2013).
[33] See sample interview in the Appendix.

The interview protocol had many similarities with that employed by Otheguy and Zentella (2012) for their corpus of interviews with 140 Spanish speakers in NYC. For instance, these researchers conducted interviews entirely in Spanish, noting that many of the speakers also used English loan words and code-switched to English on a number of occasions (p. 22). Additionally, some of the main topics discussed in their interviews included the speakers' first day in New York, visits back to their home countries, differences between life in NYC and Latin America, and work-related issues, among others (pp. 39-40). Information concerning the speakers' sociodemographic and language background and experience was also collected via a questionnaire after each interview (pp. 22-23). Although a written questionnaire was not utilized for the current study, each speaker's demographic background and language experiences were addressed as part of the interview.

4.2 The speakers

This dissertation examines subject expression using a sample of 20 Mexican speakers.[34]. The speakers' sociodemographic backgrounds can be summarized in the following way (summarized in Tables 4.1 and 4.2): The speakers were born in various regions of Mexico: Mexico City (8), Acapulco, Guerrero (2), the state of Guerrero (1),[35] Juando, Mexico (1), the state of Zacatecas (1), Cuernavaca, Morelos (1), the state of Morelos (1), Tampico, Tamaulipas (1), San Juan del Río, Querétaro (1), Monterrey, Nuevo León (1), the state of Colima (1), and the state of San Luis Potosí (1) (shown in Figure 4.3). They consist of 12 females and 8 males, and their ages range from 25 to 60. Additionally, their LORs in the U.S. range from 2 to 25 years

[34] This sample is part of a larger corpus consisting of speakers of different national origins (Colombia, Venezuela, El Salvador, Guatemala, Boliva, Ecuador, Honduras, Cuba). In order to maintain a relatively balanced sample in terms of sociodemographic factors (e.g. gender, age, LOR, AOA, English proficiency) and since the focus of this dissertation is on Mexican Spanish in particular, 20 speakers were selected from this corpus.
[35] Some speakers did not specify their city of origin.

(average = 12 years), and their AOAs range from 11 to 56 (average = 27). In terms of education levels, they range from primary school to university. The speakers have a variety of occupations, nearly half of them being small business owners (e.g. clothing stores, barber shops, jewelry stores, restaurants). Finally, their English proficiency ranges from very poor to good. English proficiency was measured using speakers' self-ratings on a scale from 1 (very poor) to 5 (very good). The average rating for all 20 speakers is 2.75 (see section 4.4.1.2 for further discussion of this factor).

Figure 4.3: Map of Mexico

Table 4.1: Speaker demographics (female speakers)

Speaker	City/State of Origin	LOR	AOA	Education	Occupation	English proficiency (scale of 1-5)
F39Mex	Mexico City, D.F	14	25	Partial law school	Owner of clothing boutique	2
F49Mex	Juando, Mexico	7	42	Primary school	Restaurant-food prep	3
F56Mex	Acapulco, Guerrero	25	31	Secondary school	Owner of tax business	4
F34Mex	Mexico City, D.F	10	24	University	Owner of clothing store	4
F25Mex	Zacatecas (state)	13	12	High school + Cosmetology school	Restaurant worker	4
F26Mex	Monterrey, Nuevo León	12	14	High school (U.S.)	Secretary	4
F52Mex	Mexico City, D.F	2	50	High school	Owner of sewing business	1
F43Mex	San Luis Potosí (state)	24	19	High school + Cosmetology school	Cosmetologist	3
F60Mex	Colima (state)	4	56	High school	Hairstylist	1
F32Mex	Acapulco, Guerrero	7	25	University	Owner of jewelry store	2
F30Mex	Mexico City, D.F	10	20	High school + Cosmetology school	Esthetician	3
F28Mex	Guerrero (state)	2	26	Master's	Cashier	3

Table 4.2: Speaker demographics (male speakers)

Speaker	City/State of Origin	LOR	AOA	Education	Occupation	English proficiency (scale of 1-5)
M51Mex	Cuernavaca, Morelos	10	41	Primary school	Landscaping	1
M41Mex	Mexico City, D.F	13	28	Partial university	Owner of clothing boutique	3
M34Mex	Morelos (state)	10	24	University	Owner of computer repair shop	3
M33Mex	Mexico City, D.F	12	21	Partial University	Owner of appliance store	3
M32Mex	Tampico, Tamaulipas	16	16	Secondary school	Manager of grocery store	3
M27Mex	Mexico City, D.F	16	11	High school (U.S.)	Auto body repair	4
M43Mex	San Juan del Río, Querétaro	25	18	Partial high school	Carpenter	2
M52Mex	Mexico City, D.F	15	37	Partial secondary school (2 years)	Owner of bakery	2

In order to more precisely determine the speakers' overall degree of contact with English (aside from factors such as LOR), I have adopted Travis and Torres Cacoullos's (2013:183) 'Contact with English' (CWE) Index. This involves quantifying speakers' responses to questions concerning first and preferred language, where English was learned, level of education attained, self-rating of English and Spanish, language spoken with family, among friends and at work, and preferred language for media (radio, TV, newspapers, books). These researchers assigned a score of 1, 2, or 3 for each of the above factors, with 1 indicating more contact with Spanish, 2 indicating a response of 'both', and 3 indicating more contact with English. A score was given to each speaker and also averaged for their entire sample of 41 speakers from their New Mexican Spanish-English Bilingual (NMSEB) corpus. They reported an overall average CWE score of

2.09 among the 41 speakers, explaining that this is approximately at the midpoint between Spanish and English. In other words, speakers in their corpus as a whole show more or less the same degree of contact with Spanish as they do English. They also reported a range among the speakers' individual scores from 1.67 to 2.75. Thus on the individual level, some speakers had greater or lesser degrees of contact with each of the langauges. These authors highlight that these scores can be used to test an English-contact hypothesis (i.e. those with a greater degree of contact with English will exhibit grammatical convergence [Travis & Torres Cacoullos 2013:183]). For the current investigation, I employed a modified version of this index that includes the following information: English proficiency (scale 1-5), social contacts (Spanish-speakers, English speakers, both), preferred media language for television, radio, newspapers, and music (Spanish, English, both), and whether or not they had any education in English (none, ESL school in the U.S., middle/high school in the U.S.). See Table 4.3 below for a summary of this information for each speaker. The application of the CWE Index was carried out in the following way:

a. English proficiency: A score of 1 was assigned for English proficiency self-ratings of 1-2, a score of 2 was assigned for ratings of 3 (middle of the 1-5 scale), and a score of 3 for ratings of 4 (none of the participants rated their English as a 5).
b. Social contacts: Spanish-speakers (1), both (2), English-speakers (3)
c. Preferred media language: Spanish (1), both (2), English (3)
d. Education in English: None (1), ESL school in the U.S. (2), middle/high school in the U.S. (3)

Average scores were then calculated for each speaker and for the set of speakers as a whole, resulting in an overall average CWE score of 1.63 for the 20 speakers (ranging from 1 to

2.75). This average suggests that there is generally a greater degree of contact with Spanish than with English among the speakers. This score, in conjunction with the additional information in Table 4.3, offers an overall picture of the speakers' language experiences and contact with English, which will be used in the interpretation of the forthcoming results as they relate to the question of contact-induced language change. Specifically, the CWE scores will serve as an additional measure to test English contact in that they will be used to determine whether those with greater degrees of contact with English, according to the CWE Index, show differential pronoun behavior compared to those with lower degrees of contact with English.

Table 4.3: Factors related to Contact with English Index

Speaker	**English proficiency (scale of 1-5)**	**Majority of social contacts**	**Preferred media language**	**Education in English**	**CWE Index**
F39Mex	2	Spanish-speakers	Both	None	1.25
F49Mex	3	Spanish-speakers	Spanish	None	1.25
F56Mex	4	Both	Spanish	None	1.75
F34Mex	4	Both	Both	University (Mexico)	2.25
F25Mex	4	Spanish-speakers	Both	Middle & High school (U.S.)	2.25
M41Mex	3	Spanish-speakers	English	None	1.75
M34Mex	3	Both	Both	ESL school (U.S.)	2
M33Mex	3	Spanish-speakers	Both	University (Mexico)	1.75
M32Mex	3	Spanish-speakers	Both	None	1.5

M27Mex	4	Both	English	Middle & High school (U.S.)	2.75
M43Mex	2	Spanish-speakers	Spanish	None	1
M52Mex	2	Spanish-speakers	Spanish	ESL school (U.S.)	1.25
M51Mex	1	Spanish-speakers	Spanish	None	1
F26Mex	4	Both	Spanish	High school (U.S.)	2.25
F52Mex	1	Spanish-speakers	Spanish	None	1
F60Mex	1	Spanish-speakers	Both	None	1.25
F43Mex	3	Spanish-speakers	Both	None	1.5
F32Mex	2	Spanish-speakers	Both	None	1.25
F30Mex	3	Spanish-speakers	English	None	1.75
F28Mex	3	Spanish-speakers	Both	ESL school (U.S.)	1.75

4.3 Circumscribing the variable context

As discussed in previous chapters, the alternation between overt SPs and null SPs is the primary focus of the present analysis, as exemplified in (1) below. The first clause illustrates the use of an overt SP (*yo* 'I') while in the second clause the speaker omits the SP.

(1) ***yo*** (OVERT) *nunca <u>he comprado</u> casa allá en México…sí,* ***Ø*** (NULL) *nunca <u>he comprado</u> una casa en México* [M52Mex]

'I have never bought a house there in Mexico…yeah, I have never bought a house in Mexico'

In order to analyze subject expression in the interviews, the audio data were first transcribed. I then extracted all the finite verbs from the interviews in order to locate each instance of SP usage.[36] In order to isolate only cases in which variation between an overt and null SP can occur in Spanish, I have excluded the following types of tokens that fell outside the variable context:

1. verbs within subject headed relative clauses

(2) *mi mami este es una gran mujer ah que* ***Ø*** *también nos* ***ha inculcado****…*

'my mom umm is a great woman uh that has also instilled in us…' [F39Mex]

2. verbs appearing with full noun phrases

(3) *Mi esposa* ***es*** *de, nacida en Virginia*

'My wife is from, born in Virginia' [M27Mex]

3. existential structures, such as *haber* and *ser*

(4) ***Ø hay*** *salvadoreñas o hondureñas, que se casan con mexicanos*

'there are Salvadorans or Hondurans, that marry Mexicans' [M32Mex]

(5) ***Ø es*** *que no, no tengo un solo patrón*

'it's just that I don't, I don't have just one boss' [M41Mex]

4. *hacer* + time expressions

(6) ***Ø hace*** *diez años vine*

'ten years ago I came' [M34Mex]

5. verbs with inanimate referents (Cameron 1995; Flores-Ferrán 2004; Otheguy & Zentella 2012)[37]

[36] I also examined lexical subjects. See section 4.3.2 for a discussion of the methodology employed.

[37] Although verbs with inanimate subject referents can appear with overt SPs (e.g. ***ellos*** *[los tacos] son buenos),* their occurrence is extremely rare (Otheguy & Zentella 2012:240-241). There were no such cases in the current data, thus warranting their exclusion.

(7) I: *Me encantan <u>los tacos</u> al pastor* > R: *Sí también,* ***Ø son*** *buenos*

'I: I love tacos al pastor > R: Yeah they are good too' [F30Mex])

6. impersonal *se* expressions

(8) *en la casa* ***Ø se habla*** *español*

'at home Spanish is spoken' [F43Mex]

7. imperatives

(9) *siempre me dijo, "****Ø estudia****,* ***Ø agarra*** *un libro…"*

'He always told me, "study, get a book…" [F52Mex]

8. verbs in 'pseudo-cleft' environments (Otheguy & Zentella 2012:242-243)[38]

(10) *la primera en graduarse de la universidad* ***fui yo***

'the first to graduate from university was me' [F28Mex]

9. set phrases/discourse markers where an overt or null SP was categorical (Silva-Corvalán 1994a; Bayley & Pease-Alvarez 1997; Flores-Ferrán 2004; Otheguy & Zentella 2012), including the following that were found in the present data:

(11) *¿Qué sé* ***yo****?*

'I don't know' (literally 'What do I know?')

(12) *que* ***yo*** *sepa*

'as far as I know'

(13) ***tú*** *sabes*

'you know'

[38] 'Pseudo-cleft' environments, according to Otheguy and Zentella (2012:242-243), involve constructions in which "a constitutent in the sentence (often the subject) receives special focus due to its position in a clause at the end of the sentence and its association with a free relative clause at the beginning of the sentence." In such constructions, the pronoun cannot be null; all such cases in the present data were overt and therefore excluded. However, postverbal SPs in general, aside from their appearance in pseudo-cleft constructions, were included as is common practice in variationist subject expression research.

(14) ¿(*me) entiendes* ∅?

'Do you know what I mean?' (literally 'Do you understand (me)?')

(15) ∅ *digámoslo así*

'we'll say it that way'

(16) ∅ *digamos*

'let's say'

(17) ∅ *(te/le) digo*

'I mean/I tell you'

(18) ∅ *(te) repito*

'I repeat (to you)'

(19) ∅ *no sé*

'I don't know'

(20) ∅ *no sabría decirte/le*

'I wouldn't know how to tell you'

(21) ∅ *mira*

'look'

(22) ∅ *¿mande?*

'Excuse me?'

(23) ∅ *¿sabes?*

'You know?'

(24) ∅ *¿ves?*

'You see?'

In the current data, speakers did not alternate between an overt and null SP in any of the above cases; that is, these structures exhibited either categorically overt SPs or categorically null SPs.[39] Thus, they were excluded from the analysis. Below is a passage from one of the interviews that illustrates verbs that were included (shown in bold) and those that were excluded (underlined). For instance, (25) was included since the overt *yo* could have been omitted (***Ø*** *los oigo que*), but (29) was excluded because *hay* cannot occur with an overt SP. The set phrase in (50) was also excluded since the overt *¿me entiende **usted**?* never appeared in the data.[40]

R: *…eh por Atlanta,* (25) ***yo los oigo*** *que* (26) ***dicen*** *que por allá* (27) *hay más…* (28) *hay más como* (29) *hay delincuentes…* (30) *hay allí este… la policía* (31) *es más, X luego aquí muchos no* (32) *traen licencia y eso, entonces la policía* (33) *está más acechando por aquellos lados… y todo eso* (34) *es que* (35) *hace que la gente* (36) *se venga más para acá… aparte pues eh…* (37) ***ellos se sienten*** *como rechazados… en en ciertos lugares, y* (38) ***yo he visto*** *que sí* (39) ***tienen*** *razón…me* (40) *ha tocado a mí ir a, algunas tiendas donde, no más porque lo* (41) ***ven****, que* (42) *es hispano, luego ya no lo* (43) ***atienden****, so- no le* (44) ***quieren*** *dar información, o, o le* (45) ***contestan*** *de malas, a veces ni le* (46) ***contestan****, entonces* (47) ***yo pienso*** *que sí,* (48) ***yo me he sentido*** *así también, sí desgraciadamente pues*

I: [unintelligible]

R: *¿verdad? pero, aquí* (49) *se **sienten** como que en su mundo,* (50) *¿me entiende? como que en su tierra pues, en su país* [F60Mex]

R: '…uh in Atlanta**, I hear** them that **they say** that there are more… there are more like criminals…there are, uh…the police are more, [unintelligible], then here many don't have a license and that, so the police are more vigilant in those areas… and all that it's just that it makes people come more toward here…besides well uh…**they feel** like rejected…in in certain places, and **I've seen** that **they are** right…it has happened to me going to, some stores where, just because **they see** one, that one is Hispanic, then **they don't attend** to one, so- **they don't want** to give one information, or, or **they respond** to one disrespectfully, sometimes **they don't even respond**, so **I think** so, **I've felt** that way too, yeah unfortunately'

[39] The structures *tú sabes* and *¿sabes?* (both in the list of set phrases above) were treated as two separate set phrases in the current analysis, the latter being an interrogative with rising intonation and clearly distinguishable from the former. Each of these cases were produced with categorical subject expression in the present data. That is, neither of them exhibited variation.

[40] While some Spanish-speakers may use this phrase with an overt pronoun, it was excluded from the analysis since it showed categorical null subject expression in the present data. Following Tagliamonte (2012:235-239), exclusions were ultimately based on the lack of variation observed in this particular dataset rather than being based exclusively on what is (not) possible for Spanish more generally based on previous research or intuition.

I: [unintelligible]

R: 'Right? But, here **they feel** like in their world, do <u>you know</u> what I mean? like in their land, in their country'

4.3.1 Contrastive contexts

As discussed in Chapter 3, contrastive contexts involve the counterposition of the subject to one or more entities for reasons of opposition, confirmation, or to adopt a different position, and may include the use of implicit contrast in which one of the entities is not explicitly stated in the utterance (Enríquez 1984). Essentially, the speaker is differentiating the subject from other referents or making a comparison between two (or more) people in some way, as in (51) below. Here the speaker distinguishes herself from her husband by stating that she stayed in Mexico while her husband travelled to the United States.

(51) *<u>mi esposo</u> tuvo que viajar a Estados Unidos, estuvo aquí, mucho tiempo, **yo** me quedé allá, en México*

'my husband had to travel to the United States, he was here, for a long time, I stayed there, in Mexico' [F52]),

Contrastive contexts are treated differently across SP studies in the variationist literature, with some researchers excluding them (e.g. Silva-Corvalán 1994a) and others including them (e.g. Otheguy & Zentella 2012). I have decided to include them in the present analysis since, although they favor overt SPs (as in 51 above), there are a number of cases in the current data in which null pronouns can be observed.[41] In other words, contrastive tokens are not categorically overt, as illustrated with the following examples:

[41] 18 out of 50 (36%) cases of contrast were null. In the current data, contexts with the explicit mention of or the possibility of opposing referents were identified as contrastive. These often involved features such as negation, the use of certain conjunctions (e.g. *pero* 'but'), opposite verbal semantics (e.g. *quedarse* 'stay' vs. *viajar* 'travel'), and

(52) *Bueno, <u>en lo personal</u>, **Ø** no **soy** muy religiosa pero…por lo que veo, la gente sí va bastante a la iglesia, aquí.* [F30Mex]

'Well, personally, I'm not very religious but…from what I see, people do go quite a bit to church, here'

(53) *nada más nos este, dedicamos, de lo que, bueno **yo** <u>en lo personal</u>, me **dedico** a lo, lo que es…* [F49Mex]

'We just uh, are devoted to, of what is, well I personally, am devoted to what, what is….'

(54) *cuando uno en su posición de inmigrante, llega a este país pues realmente, eh, buscas las oportunidades y, realmente no son muchas, eh, al menos <u>en mi caso</u>, **Ø trabajé** de, desde limpiando casas* [F32Mex]

'when one in their position as an immigrant, arrives to this country well really, uh, you look for opportunities and, there really aren't many, uh, a least in my case, I worked from, from cleaning houses'

(55) *todos nos catalogan como delincuentes, que venimos, a quitar el trabajo de, de de, pues de los que son nacionalizados aquí, de, de los que tienen más derechos porque es su país, entonces, yo <u>en mi caso</u>, **yo** no **digo** que vengo a, a quitarles un trabajo, yo estoy aquí porque…* [F60Mex]

'everyone categorizes us as criminals, that we come, to take the job from, from from, well from those who are nationals here, from, from those who have more rights because it's their country, so, I in my case, I'm not saying that I'm coming to, to take away a job from them, I'm here because…'

---

the use of certain adverbials (e.g. *en lo personal* 'personally'; *en México* 'in Mexico') (See Amaral & Schwenter 2005 and Travis & Torres Cacoullos 2012).

(56) *allá hay que pagar para estacionarse porque nunca hay espacio, y acá pues,* **Ø** *todavía no* ***tenemos*** *ese problema.* [F32Mex]

'there one has to pay to park because there's never space, and here well, we don't have that problem yet'

(57) *Por ejemplo en México* **Ø** *no* ***tenemos*** *la costumbre de hacer ejercicio, y yo eso lo noto en, en la gente de acá* [F32Mex]

'For example in Mexico we don't have the habit of exercising, and I notice that in, in people from here'

(58) *cuando me encuentro, a un americano, tienen esa hermosa costumbre, de saludarte,* ***nosotros*** *en México no* ***tenemos*** *esa bonita costumbre* [F32Mex]

'when I encounter, an American, they have that nice habit, of greeting you, we in Mexico don't have that nice habit'

In examples (52), (54), (56), and (57), the pronouns are null despite the contrast. Consistent with the analysis of Amaral and Schwenter (2005), we can see that, instead of overt pronouns indicating the contrast, adverbials such as *en lo personal, en mi caso, acá*, and *en México* function to highlight the change in referent. In (52) the speaker is distinguishing herself from others by indicating that, in contrast to people that go to church frequently, she is not religious. Her use of the adverbial *en lo personal* highlights this contrast (compare with overt variant in 53). In (54) the speaker is indicating contrast with *en mi caso*, saying that at least in her case, there were not many job opportunities and that she worked (perhaps unlike others) cleaning houses (compare with overt variant in 55). Further, (56) illustrates a comparison between Mexico (*allá*) and Georgia (*acá*) whereby the speaker is making a contrast between having to pay to park due to lack of space in Mexico and not having such a problem in Georgia. The locative adverbial

acá (and not *nosotros*) serves to indicate this contrast. Likewise, a locative adverbial is seen in (57) (*en México*) and a comparison between places is made. Here the speaker indicates a contrast by stating that in Mexico they do not have the habit of exercising like people do in Georgia. As in (56), the overt *nosotros* is not used to signal a switch in referent, but compare with the overt variant in (58). Thus, the above examples illustrate that overt SPs are not categorical (nor obligatory) in contrastive contexts, at least not in the current data. They also corroborate Amaral and Schwenter's (2005:125) argument that "it is inaccurate to say that SPPs are obligatory in contrastive contexts". Interestingly, some cases of contrast were also found in the current data with neither the presence of adverbials nor overt SPs, a pattern that, to my knowledge, has not been discussed in previous research. In fact, Amaral and Schwenter (2005) argue that expressions like adverbials are *necessary* in contrastive contexts if the SP is null in order to maintain pragmatic felicity. However, I maintain that this need not be the case as we see in the following examples where both overt SPs and adverbial expressions are absent:[42]

(59) *tal vez, la otra actividad más importante sería, el, conmemorar el ... las fiestas de diciembre, tal vez sea, este el nacimiento de Jesús o el año nuevo allí…* **Ø** *no* ***somos*** *como los, tal vez la cultura china que tiene más actividades en todo el transcurso del año según la religión…* [M34Mex]

'perhaps, the other most important activity would be, the, commemorating the … the December holidays, perhaps it's, uh the birth of Jesus or the new year or something…we aren't like the, perhaps the Chinese culture which has more activities throughout the course of the year according to the religion…'

[42] Although similar examples are presented in works such as Travis and Torres Cacoullos (2012) and Otheguy, Zentella, and Livert (2007), there is no discussion of these particular examples by the authors with regard to the (non) presence of adverbials in particular.

(60) *yo creo que a los seis años en lugar de estar jugando una pelota* ***Ø estábamos****, trabajando para ganar un poco de dinero* [M34Mex]

'I think at six years old instead of playing with a ball we were, working to make a little bit of money'

(61) *No soy egoísta pero tampoco* ***Ø soy*** *de las personas que, que este, que piense que hay que dejar todo lo que uno tiene* [F49Mex]

'I'm not selfish but I'm also not like people that, that uh, that think that one has to leave everything that one has'

(62) *pues no, no este,* ***yo*** *no* ***soy*** *de las personas que, que les diga a los demás "deja tus cosas y vente para acá"* [F49Mex]

'well no, no uh, I'm not like people that, that say to everyone else "leave your things and come here"

In (59) the speaker is saying that people in Roswell ([*nosotros*] *somos*) do not have many activities throughout the year, making a contrast with Chinese culture. Example (60) demonstrates another case, in which the speaker explains that instead of playing ball, "*estábamos trabajando*" (we were working). Here the contrast is made implicitly between other kids her age and the speaker. There is not an overt SP, nor is there an adverbial expression to indicate the contrast of the first-person plural subject. Finally, both elements are absent in (61) in which the speaker is distinguishing herself from others who think a certain way (also compare with overt variant in 62). Therefore, examples 59-62 demonstrate that (albeit infrequently) contrast can be expressed between subjects felicitously in the absence of either overt SPs or adverbials, in contrast to the claim made by Amaral and Schwenter (2005). In light of the above discussion and illustrations of contrast, such tokens were included in the present investigation.

4.3.2 Lexical subjects

In addition to SPs, I have also examined lexical subjects (e.g. *mi madre* 'my mother'). As mentioned in Chapter 3, most variationist studies of subject expression do not include lexical subjects. Due to such lack of attention in previous research, I include them in the present investigation in order to gain a better understanding of their variation with SPs. In order to isolate only the cases in which variation can occur between all three subject forms (lexical subjects, overt SPs, and null SPs), I have excluded lexical subjects in the following contexts: (a) when they introduce a new referent into the discourse (i.e. first mention cases), (b) when the previous mention of the referent is at a distance such that a pronoun would not suffice to identify the referent (5+ clauses back), (c) when there is a competing referent, and (d) when the subject refers to a collective referent. Each of these contexts is illustrated below:

(63) First mention:

*después este **un amigo** me invitó a...este, a trabajar con él en...(1.5) en poner carpeta, en los apartamentos* [M27Mex]

'after that uh a friend asked me to…uh, work with him in…(1.5) in putting carpet, in apartments'

(64) Distance from previous mention:

*hace dos años <u>mi madre</u> fallece, de cáncer, ujum, este tengo, todos mis hermanos, tienen esposas e hijos, sí, en total somos una familia de 35, personas, entre hermanos, nietos, uhm, esposas, esposos, ajá, sí, sí es una familia de, de, ajá, pero ya no existe **mi madre** y mi padre tampoco* [F52Mex]

'two years ago my mother dies, from cancer, uh huh, uh I have, all my brothers, have wives and children, yeah, overall we're a family of 35, people, among siblings,

grandchildren, um, wives, husbands, uh huh, yeah, yeah it's a family of, of, uh huh, but my mother no longer exists and neither does my father'

(65) Competing referent:

mi abuelo tenía este, descendencia de españoles...ah...(1) ***mi papá*** *viene de una familia de nueve hermanos, él es uno de los mayores...(2) y bueno mi papá continúa con…* [F39Mex]

'my grandmother had uh, Spanish descent…uh…(1) my dad comes from a family of nine siblings, he is one of the oldest…(2) and well my dad continues with…'

(66) Collective referent:

entonces ***la policía*** *está más acechando por aquellos lados… y todo eso es que hace que* ***la gente*** *se venga más para acá* [F60Mex]

'so the police are more vigilant in those areas… and all that it's just that it makes people come more toward here'

For first mention cases, the speaker introduces a referent for the first time in the discourse. In (63) the speaker's friend (*un amigo*) is a new referent that had not been previously mentioned in the interview. His friend could not have been referred to with *él* or **Ø** because the listener would, presumably, not have been able to interpret the antecedent. Thus, first mention cases, generally speaking, can only be realized as lexical subjects or else the listener is likely not able to identify the intended referent.[43] In other cases, such as (64), a lexical subject that is *not* a new referent was used, but its previous mention occurred at a distance, with multiple intervening

[43] There were a few exceptions to this first mention generalization observed in the current data. For example, one speaker referred to her son for the first time by using *él.* This was possible since her son was part of the extralinguistic context (sitting beside her). However, due to only finding two or three cases of such reference in the entire dataset, first mentions were excluded since, for the most part, they did not exhibit variation. Such examples of pronouns for first mention have also been attested in previous research (e.g. Blackwell 1998:616).

clauses between the target structure and its previous reference.[44] Thus, the use of *mi madre* (bolded) could not have varied with *ella* or ∅ because the use of a pronoun by the speaker would likely make it difficult for the listener to recover the intended antecedent (*mi madre*). Recall from Chapter 3 the notion of accessibility in relation to pronominal subjects (Givón 1983; Ariel 1994). The greater the distance between referents and their antecedents, the lower the accessibility and salience of such referents. Thus, less accessible subject referents are more likely to be marked with more coding material to enable the identification of the referent (Givón 1983), which in (64) is a lexical subject instead of a pronoun due to the degree of distance from the previous mention of *mi madre*. In other words, the distance (5 clauses back) of the first case of *mi madre* (underlined) from the target structure makes a lexical subject necessary for the subsequent reference to the speaker's mother. Other cases of distant previous mention involved a greater number of clauses intervening between the reference and the target (e.g. 10+ clauses), including cases in which the previous mention was found in an entirely different speech turn.

Furthermore, competing referents also make the use of lexical subjects necessary. The speaker in (65) produced the subject form *mi papá* and could not have used a pronoun to refer to her father because of the presence of the previous subject *mi abuelo*. The use of *él* or ∅ in place of *mi papá* would indicate reference to her grandfather. That is, although a pronoun could have occurred in place of *mi papá*, it would not have referred to the same subject (her father). On the other hand, the subsequent mention of *mi papá* in this example (underlined) was included in the

[44] Lexical subjects were the preferred variant of the speakers when their previous coreferential mentions were at a distance of 5+ clauses back in the discourse. Therefore, for the purposes of the present investigation, when such distance was present in the interviews, the lexical subject was excluded since variation with a pronoun was not observed. This pattern for distance is consistent with Dumont's (2006) analysis of lexical subjects in that she found that a greater number of intervening clauses between two coreferential subjects resulted in a higher use of lexical subjects and a lower use of pronouns, especially at a distance of 10+ clauses (p. 289).

present analysis since it does not have a competing referent, is not a first mention, and does not have a distant previous mention. Specifically, there is continuity in reference in that the previous subject *él* also refers to the speaker's father. In this case, the speaker could have alternatively said ***él*** *continúa* or **Ø** *continúa*. Finally, the use of collective referents, as in (66), presents a case where variation with overt and null SPs is not possible (Otheguy & Zentella 2012). In this example, neither *la policía* nor *la gente* could have been replaced by *ella* or **Ø** since collective subjects cannot be referred to with pronouns in Spanish.[45] In sum, due to the (near) categorical use of lexical subjects in the above contexts, these cases were excluded. The following section outlines the independent linguistic and social variables included in the current study.

4.4 Linguistic and social variables

Given that one of the primary objectives of this dissertation is to investigate issues of contact-induced language change and the potential influencing role of English, I have chosen to include independent variables that would facilitate comparability between my results and studies of monolingual Mexican Spanish. Therefore, I have included the same set of linguistic variables that were observed by Lastra and Martín Butragueño (2015) in their study of Mexico City Spanish:[46]

[45] The only exception to this is when a subsequent reference to the collective is made with third-person plural inflection and/or a plural SP (e.g. *La gente tiene que aprender el idioma y (**ellos**) **tienen** que respetar la ley* 'people have to learn the language and they have to respect the law').

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

a. Person/number

b. Switch reference

c. Tense-mood-aspect (TMA)

d. Morphological ambiguity

e. Verb class

f. Verbal mood

g. Specificity

h. Polarity

i. Style

The reason for choosing their study as a primary comparison is mainly due to the regional compatibility between data sets: though not a perfect comparison, a number of speakers in the current data are from Mexico City (N = 8). Other studies of subject expression in varieties of monolingual Mexican Spanish employ speakers from the Yucatan Peninsula (Solomon 1999; Michnowicz 2015) as well as Veracruz (Orozco 2016). However, none of the speakers from the present data are from these regions, and thus these particular studies were not chosen as a primary point of comparison. Nevertheless, they are certainly useful for comparisons with monolingual Mexican Spanish more generally. While many of the above variables have been shown to constrain subject expression repeatedly across numerous studies and dialects, others such as polarity and style are much less studied, offering the opportunity for this dissertation to shed more light on their influence.

With regard to the social variables, I have included the following:

a. Age

b. Gender

c. Age of arrival

d. Length of residency

e. English proficiency

f. Preferred media language

g. Educational level

h. Socioeconomic status (SES)

The selection of these variables was primarily based on their inclusion in previous studies of subject expression in Spanish, especially those in situations of language contact, as outlined in Chapter 3. Certain variables in particular, such as preferred media language, educational level, and SES, were also chosen due to there being relatively little research with regard to their influence on subject expression. The following sections discuss the manner in which each linguistic and social variable was operationalized in the present study.

4.4.1 Coding of variables

For all tokens within the variable context (N = 4,705), I first coded whether each verb appeared with a null SP, an overt SP, or a lexical subject to account for the dependent variable in the study. Subsequently, I coded for the independent variables using their respective categorizations as discussed in the following subsections.

4.4.1.1 Linguistic variables

Person/number

The present study includes five person/number categories: first-person singular (*yo* 'I'), second-person singular (*tú* 'you' [informal]), third-person singular (*él*/*ella* 'he/she, *usted* 'you'

[formal]),[47] first-person plural (*nosotros*/*nosotras* 'we' [masc./fem.]), and third-person plural (*ellos*/*ellas* 'they' [masc./fem]), as exemplified below:

(67) First-person singular:

*…también este **estuve** trabajando este como pasante de abogado, y bueno **yo creo** que siempre es bueno trabajar* [F39Mex]

'also uh I was working uh as a paralegal, and well I think it's always good to work'

(68) Third-person plural:

*el año que entra **vienen**, o sea puedo verlos… porque **ellos vienen*** [M32Mex]

'next year they're coming, so I mean I'm able to see them…because they come'

Switch reference

Switch reference comprises three categories in the current analysis: Same reference, switch reference, and partial switch. Adopting a more extensive method from previous SP studies (e.g. Orozco and Guy 2008; Otheguy and Zentella 2012; Lastra and Martín Butragueño 2015), I have considered not only same or switch reference between subjects, but also cases of coreference with preceding objects (direct/indirect object, object of preposition, etc.) within switch reference contexts, which are referred to as partial switches. Examples from the data that illustrate each of the three categories are shown below:

[47] *Usted* was included in the third-person singular and not the second-person singular category because this variable was coded according to morphological person/number of the verb, not person/number of the subject referent (see Otheguy & Zentella 2012:155).

(69) Same reference:

*yo tengo un buen concepto de esta ciudad, me gusta <u>me siento</u> cómoda no me **siento** discriminada no me **siento** rechazada por mi color de piel, por el, por el idioma que **yo hablo*** [F32Mex]

'I have a good concept of this city, I like it I feel comfortable I don't feel discriminated against I don't feel rejected for my skin color, for the, for the language that I speak'

(70) Switch reference:

*pues, noso- <u>yo tenía</u> siempre la inquietud de de oír lo que **ellos estaban** diciendo y no **entendía*** [F60Mex]

'Well, we- I always had the curiosity of of hearing what they were saying and I didn't understand'

(71) Partial switch:

a. *yo tengo un <u>hijo</u>, pero **él está**, está en X, en Tijuana, en Tijuana ahí vive* [M43Mex]

'I have a son, but he is, he's in [unintelligible], in Tijuana, in Tijuana he lives there'

b. *...tengo, varios <u>amigos</u>, pero **son** del trabajo* [M43Mex]

'…I have, several friends, but they're from work'

In (69), since there is continuity in the subject referent (*yo*) from the first case of *siento* (underlined) to the second and third cases, and then from *siento* to *yo hablo*, all three bolded tokens were coded as same reference. In (70), there is a switch in subject from *yo* to *ellos* and then another switch back to *yo* (*<u>entendía</u>*). Thus, both of these tokens were coded as switch reference. Finally, (71a) exemplifies a partial switch, in which there is a switch in subject from *yo* to *él*, and *él* is also coreferential with the preceding object *hijo*. The same is illustrated in (71b) (a switch from *yo* to *ellos*, with *amigos* as the preceding coreferential object).

Additionally, following Otheguy and Zentella (2012:259), I took the interviewer's speech into account regarding coreferentiality and considered the previous reference made by the interviewer when analyzing the respondent's first token of their speech turn. That is, I coded same reference for any case in which the interviewer's final reference was the same as the respondent's first reference, as seen in (72), and switch reference when the interviewer's final reference was different.

(72) I: *OK, um, <u>vive</u>, entonces ahora en Roswell, ¿verdad?*

'OK, um, so you live, in Roswell now, right?'

R: *Uh-huh sí, me **traje** a mi familia completa… que tengo, son... tres hijos…tres hijos en total* [M52Mex]

'Uh-huh yeah, I brought my whole family…I have, they're…three children…three children total'

For example, since there is no switch in reference in (72) between the respondent's self-reference *traje* and the interviewer's reference *vive*, this could influence the use of a null pronoun by the respondent.

Tense-Mood-Aspect (TMA)

Nine TMA categories were included: Present indicative, preterit, imperfect, perfect (including present perfect and pluperfect), present subjunctive, past subjunctive, synthetic future, periphrastic future, and conditional. The imperative mood was not included since it was categorically null. In other words, there were no imperative verbs in the data that appeared with an overt SP. To exemplify the coding of TMA, tokens of the imperfect and preterit are provided below.

(73) Imperfect:

*Aquí, **yo** no **conocía** nada… y **tenía** que estar este… **tenía** que estar este... encerrado aquí* [M32Mex]

'Here, I wasn't familiar with anything…and I had to be uh…I had to be uh…enclosed here'

(74) Preterit:

*Claro que sí, **yo estudié** una licenciatura, en administración turística en la Universidad de Guadalajara en México, también **estudié** una maestría en España*…[F28Mex]

'Of course, I studied a Bachelor's degree, in tourism administration at the University of Guadalajara in Mexico, I also studied a Master's in Spain…'

Morphological ambiguity

Morphological ambiguity comprised two categories: ambiguous and unambiguous. Ambiguous morphology included first and third-person singular verbs that were either in the imperfect, subjunctive (present or past), conditional, or pluperfect since verb inflections with these particular persons/numbers and TMAs are the same for both the subject referents *yo* and *él/ella/usted*. For example, the antecedent of the imperfective form *tenía* could be either *yo*, *él*, *ella*, or *usted*, but this distinction is unclear without the proper context or without an overt SP. On the other hand, the unambiguous category included all other verbs since there is indeed such morphological distinction (e.g. *tengo* 'I have', *teníamos* 'we had'). Examples of ambiguous and unambiguous tokens from the current data are shown below in (75) and (76), respectively.

(75) Ambiguous:

*pero, **yo** allá siempre **comía** frijoles con queso… beans…y tortilla... that's it…Era lo único que **comía**…*[M32Mex]

'but, there I always ate beans with cheese…beans…and tortillas…that's it…it was the only thing I would eat…'

(76) Unambiguous:

*Entonces **ella se casó**, eh y así sucesivamente **nosotros empezamos** a venirnos, uno por uno, y llegar y llegar allí a su casa, a buscar trabajo y este, y todos **empezamos** a casarnos* [F49Mex]

'Then she got married, uh and like that successively we started to come, one by one, and arrive and arrive there at her house, to look for work and uh, and we all started to get married'

Verb class

Following the categorization of Bentivoglio, Ortiz, and Silva-Corvalán (2011), verb class consisted of four categories: Mental processes (e.g. *creer* 'believe', *pensar* 'think', *saber* 'know', *entender* 'understand', *querer* 'want'), stative verbs (e.g. *ser* 'be', *estar* 'be' *vivir* 'live', *tener* 'have'), *verba dicendi* (i.e. communication, e.g., *decir* 'say', *hablar* 'speak', *preguntar* 'ask'), and activity verbs (e.g. *jugar* 'play', *hacer* 'do/make', *estudiar* 'study', *ir* 'go'). The examples below illustrate tokens from each verb class category in the present data:

(77) Mental:

Yo creo *que, uhm, no es fácil, porque dejas todo, tu, tu familia, tu casa, tu, tus costumbres* [F30Mex]

'I think that, um, it's not easy, because you leave everything, your, your family, your home, your, your customs'

(78) Stative:

también ***tenemos*** *amistades este brasileñas, queremos que nuestros hijos siempre…* [F39Mex]

'we also have umm Brazilian friends, we want our children to always…'

(79) *Verba Dicendi* (communication):

rara vez ***hablo*** *con algunos de mis hermanos de allá, pero, es cada mes, cada mes y medio.* [M43Mex]

'I rarely talk to any of my siblings from there, but, it's every month, every month and a half'

(80) Activity:

aquí sí, my…aquí es el middle school, y high school…y ***estudié*** *un poquito más mhmm.* [F25Mex]

'here yes, my…here it's middle school, and high school…and I studied a little bit more mhmm'

Example (77) illustrates mental process verbs in the data with the cognitive verb *creer* 'believe/think'. The use of stative verbs is exemplified in (78) with the verb *tener* 'have'. Further, the verbal semantics in (79) represents the category of *verba dicendi* with the

communication verb *hablar* ‘talk/speak’, and (80) demonstrates the use of activity verbs (*estudiar* ‘study’).

Verbal mood

Verbal mood included two categories: Indicative and subjunctive.

(81) Indicative:

nuestro trabajo lo ***hacemos*** *con mucho gusto* [F39Mex]

‘our job we do it with a lot of pleasure’

(82) Subjunctive:

quiero que ***ellos*** *no* ***pierdan*** *el español* [M41Mex]

‘I want them not to lose Spanish’

In (81), the present indicative mood is used by the speaker with the verb *hacer* ‘do’, and in (82) the present subjunctive is produced with the verb *perder* ‘lose’. Verbs in the imperative mood, as indicated above, were excluded due to being categorically null.

Specificity

Specificity consisted of specific and nonspecific reference. This factor was operationalized taking into account the criteria set forth by Otheguy and Zentella (2012:255) for the factor they call “definiteness”: A referent is considered to be definite/specific as long as there is an explicit antecedent in the linguistic context (with the exception of *alguien* ‘someone’, *alguno* ‘some’, *nadie* ‘nobody’, *quien* ‘who’, and *uno* ‘one’). In contrast, a referent is considered indefinite/nonspecific when an explicit antecedent is not mentioned. Nonspecific referents also include second-person singular forms that do not refer specifically to the listener. The examples below illustrate specific and nonspecific reference:

(83) Specific:

*<u>mis padres</u> nunca quisieron una casa... grande, porque **ellos decían** que, eso rompe l- la unión familiar* [M33Mex]

'my parents never wanted a…big house, because they said that, that breaks th- the family bond'

(84) Nonspecific:

*…en la secundaria nos **dan** algo, algo mínimo, casi nada de inglés*...[F52Mex]

'…in middle school they give us something, something minimal, almost no English…'

(85) Nonspecific:

*allá también hay más transporte y aquí, eh no, **tienes** que o tener un carro o, o usar el servicio de taxi que es súper costosísimo* [F32Mex]

'there there's also more transportation and here, uh there's not, you have to either have a car or, or use the taxi service which is super expensive'

The token *ellos decían* in (83) was coded as specific since it contains the explicit antecedent *mis padres*. However, (84) was coded as nonspecific due to lack of an explicit antecedent. That is, there is not a specific subject for *dan* mentioned in the surrounding discourse, such as *los maestros* 'the teachers'. The use of *tienes* in (85) is also nonspecific since the speaker is not addressing the interviewer specifically but uses the impersonal *tú* form.

Polarity

Following Lastra and Martín Butragueño (2015), polarity included two categories: Affirmative and Non-affirmative, the latter including both negative clauses and interrogatives.[48] The three examples below illustrate each of these categories.

[48] As noted in Chapter 3, Lastra and Martín Butragueño (2015) use the term *enunciative type* for this variable.

(86) Affirmative:

*un día típico para mí, **yo me levanto** temprano, eh, **superviso** la cocina*

'a typical day for me, I get up early, uh, I supervise the kitchen' [F32Mex]

(87) Non-affirmative (negative):

*hace muchos años, no, <u>no</u> **enseñaban** inglés en las escuelas en México* [F34Mex]

'many years ago, they didn't, didn't teach English in schools in Mexico'

(88) Non-affirmative (interrogative):

*¿Qué **fue usted** a conocer por ahí?* [F30Mex]

'What did you go see there?'

Affirmative clauses are shown in (86) in which the speaker produced *yo me levanto* 'I get up' and *superviso* 'I supervise'. In (87), a non-affirmative clause, specifically the use of negation, is illustrated (***no** enseñaban* 'they didn't teach'). Finally, (88) exemplifies another non-affirmative token, which in this case is a question (interrogative) asked by the interviewee.

Style

The operationalization of style was based on a distinction between (a) the first part of the interview in which the speaker had shorter turns and where there were more dialogic pairs (more interview-like) and (b) the second part of the interview which involved longer speech turns by the interviewee (more conversation-like) (Lastra & Martín Butragueño 2015). Tokens in the former category were coded as First part, as in (89), and those in the latter were coded as Second part, as in (90). Specifically, First part tokens were those that occurred within the first 10-15 minutes of each interview when speakers were asked about their demographic background and when the questions were less open-ended. This is when the speakers' responses were generally much shorter. After the first 10-15 minutes, the questions were more open-ended and the

speakers produced much longer responses. The interview flowed more like a conversation at this point, and tokens were coded as Second part. This manner of operationalizing style is based on a distinction of there being, on one hand, well-defined roles between the interviewer and interviewee with several brief dialogic pairs at the start of the interview, and on the other hand, more diffused interviewer/interviewee roles with more extensive turns as the interview progresses (Lastra & Martín Butragueño 2015).

(89) First part:

I: ¿Vive usted aquí en Roswell?

'Do you live in Roswell?

*R: Eh todo el- durante diez años **viví** en Roswell… y ahora recientemente en se- septiembre pasado,*

'Uh the whole- for ten years I lived in Roswell…and now recently in Se- last September,'

I: Uh-huh.

*R: eh me **mudé** a Sandy Springs,*

'uh I moved to Sandy Springs,'

I: ¿A dónde?

'Where?'

R: Sandy Springs

'Sandy Springs'

I: Ah OK, ¿Hace cuánto tiempo?

'Ah OK, how long ago?'

R: De septiembre, last September…it's not even a year so mhm.

'since September, last September…it's not even a year so mhm'

I: Um y ¿en qué ciudad nació en México?

'Um and what city were you born in in Mexico?

R: La Ciudad de México, Mexico City, uh-huh la ciudad. [F34Mex]

'Mexico City, Mexico City, uh-huh the city'

(90) Second part:

I: ¿Diría usted que tiene una preferencia entre español e inglés?

'Would you say that you have a preference between Spanish and English?'

R: ***Yo prefiero*** *español pero...(2) eh mi cerebro- es lo que le* ***explicaba*** *a mi mamá porque mi mamá* ***dice*** *"Carolina, hablas como, like… spanglish", no le gusta a mi mamá cuando* ***yo hablo*** *por teléfono, le da mucho coraje... y me- cuando X, porque mis tías les pasó lo mismo,* ***empezaron*** *a olvidar el español, y pasa, pasa porque como ya no lo* ***usas****, hasta en tu cerebro porque, yo, antes… cuando recién* ***llegué****, todo lo que* ***yo quería*** *decir, primero lo* ***pensaba*** *en español, y después lo* ***traducía*** *en mi cerebro, y después lo* ***empezaba*** *a decir, ¿OK?... pero entonces* ***empecé****, te digo, muchos años* ***estaba*** *todo el tiempo con americanos americanos, entonces me* ***di*** *cuenta o mi cerebro empezó a... pensar en inglés y hablar en inglés* [F34Mex]

'I prefer Spanish but…(2) uh my brain- it's what I was explaining to my mom because my mom says "Carolina, you talk like, like…Spanglish", my mom doesn't like it when I talk on the phone, it makes her really mad…and I- when [unintelligible], because the same thing happened to my aunts, they started to forget Spanish, and it happens, it happens because if you don't use it anymore, even in your brain because, I, before…when I had just arrived, everything I wanted to say, first I would think about it in Spanish, and then I translated it in my brain, and after I started to say it, OK?...but then I started, as I told you, for many years I was with Americans Americans all the time, so I realized or my brain started to…think in English and speak in English'

4.4.1.2 Social variables[49]

Age was coded using three categories: 20-34, 35-49, and 50+. Gender was coded as male or female. Length of residency (LOR) includes three categories: 1-7 years, 10-14 years, and 15-25 years. Age of arrival (AOA) also comprises three categories: 11-19, 20-28, and 30-56. Regarding English proficiency, speakers' self-ratings on a scale from one (very poor) to five (very good) were used, following the methodology of Travis and Torres Cacoullos (2013).[50] The participants' preferred language for media (television, music, radio, and newspapers) was coded using three categories: Spanish, English, or both (cf. Travis & Torres Cacoullos 2013). With regard to educational level, four categories were included: elementary, secondary, college, and graduate, following Otheguy and Zentella (2012). Finally, socioeconomic status (SES) was operationalized following Otheguy and Zentella (2012:271), who used a combined measure based on both education and occupation of the speakers. Specifically, speakers were given points ranging from 1 to 4 in the following way:

Education: elementary (1), secondary (2), college (3), graduate (4)

Occupation: unskilled (1), skilled blue collar, clerical (2), store owner, manager, white collar (3), professional, business owner (4).

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

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

The speakers' SES was then based on combining the point values for each of the two categories and classifying them into four SES categories: (a) 1-2 points, (b) 3-4 points, (c) 5-6 points, and (d) 7-8 points. The following section discusses the statistical analyses carried out for this dissertation.

4.5 Statistical methods

To determine the statistical significance of the linguistic and social variables and the relative weight of each factor regarding its conditioning on SP use, I carried out mixed-effects multivariate analyses using the program Rbrul (Johnson 2009) with the inclusion of the speaker as a random effect. A multivariate analysis allows us to determine the relative effect of multiple factors at once. Specifically, it will facilitate the production of ranking for the statistically significant factors based on the amount of variation explained by each factor. The results of the multivariate analysis are what will aid in determining, for example, if it is the person/number of the verb that most strongly predicts when speakers will produce overt SPs and when they will omit them. Moreover, the quantitative results assist in creating a ranking of values within a single factor group or variable (e.g. person/number)—i.e. values within a factor group that favor overt SPs (e.g. first-person singular) vs. those factors that disfavor SPs (e.g. third-person plural). In addition, the inclusion of the speaker as a random effect in the multivariate analysis is a way of controlling for the individual speaker and ensuring that the results obtained are generalizable to the data set as a whole, and that the patterns are not due to particular speakers skewing the results (e.g. Bayley, Greer, & Holland 2013; Shin 2014; Michnowicz 2015; Shin & Van Buren 2016). Essentially, a random effect such as the speaker or verb lexeme, another common factor employed as a random effect (see, e.g., Erker & Guy 2012; Orozco 2016; Shin & Van Buren 2016), takes into account individual factors and their potential effects on the variation observed

in the statistical model. Including the speaker as a random effect "takes into account that some individuals might favor a linguistic outcome while others might disfavor it, over and above (or 'under and below') what their gender, age, social class, etc. would predict" (Johnson 2009:365).

Furthermore, in order to explore the relationship among the independent variables in the data and to spot potential interactions between them, I have employed cross-tabulations and conditional inference tree analysis. These particular types of analysis will reveal which variables interact with each other to constrain subject expression. In other words, they helps us to determine when a single independent variable (e.g. person/number) is constrained by another independent variable (e.g. switch reference) in its conditioning of SP variation, rather than having an effect on the data set as a whole. Thus, for example, person/number and switch reference could interact in such a way that third-person singular tokens favor overt SPs, but only (or to a greater degree) when there is a switch in subject referent (not for cases of *same* reference). Cross-tabulations and conditional inference trees help us to reveal such complex relationships in the dataset. In the following section I discuss my predictions regarding the linguistic and social variables and their influence on subject expression in Roswell.

4.6 Predictions

This section presents predictions regarding the overall overt pronoun rate for Mexicans in Roswell as well as the factors that constrain SP usage in this variety of Spanish. Recall the first three research questions from Chapter 1:

1. What is the overall distribution of subject expression in terms of overt and null pronoun usage? How does the rate of overt SP usage compare to that of non-contact varieties of Mexican Spanish?

2. How do contextual/language internal factors (e.g. grammatical person/number, tense-mood-aspect, switch reference) condition SP variation and how do such patterns compare to previous research?

3. Do language external factors such as English proficiency, age, sex, length of residency in the U.S., and age of arrival to the U.S. play a role in how the speakers use SPs?

Based on the findings of previous studies outlined in Chapter 3 (including research on Mexican Spanish in particular and Spanish in general) regarding both overall pronoun rates and correlations between independent variables and Spanish subject expression, I predict the following concerning SP use in the current data:

Pronoun rate and linguistic variables

a. Pronoun rate: Based on preliminary research of Roswell Mexican Spanish (Limerick 2017), which has reported a rate similar to that of monolingual Mexican Spanish (21%), I predict that speakers in the present analysis will produce overt SPs at a similar rate. Furthermore, given that the speakers are Spanish-dominant and that, in general, they have more contact with Spanish-speakers than English-speakers, they will not exhibit a substantial increase in their overall pronoun rate compared to non-contact varieties of Mexican Spanish.

b. Person/number: Singular verb forms, especially first-person singular and third-person singular forms, will favor overt SPs while plural forms will disfavor them (Silva-Corvalán 1994a; Flores-Ferrán 2002; Shin 2012; Lastra & Martín Butragueño 2015; Limerick 2017).

c. Switch reference: Switch reference contexts will favor overt SPs while same reference contexts will favor nulls. However, these contextual effects may be weaker than those observed in monolingual varieties given that previous research on contact Spanish has reported a reduced importance for this constraint for bilinguals (e.g. Flores Ferrán 2002, Shin & Otheguy 2009; Otheguy & Zentella 2012; Michnowicz 2015), including for Roswell Spanish (Limerick 2017). Particularly, this weakening will manifest as an increased use of overt SPs in same reference contexts (e.g. *Mi hijo tiene diez años…* ***él*** *sabe bien el inglés pero* ***él*** *no quiere estudiar español*). Additionally, partial switches (i.e. a switch in subject but no switch with the preceding object referent) will show little effect, that is, neither a statistical favoring nor disfavoring of overt SPs.

d. TMA: Imperfect, conditional, and subjunctive verb forms will favor overt SPs while preterit forms disfavor them. Present tense verbs and perfect forms will show little effect (Cameron 1994; Travis 2007; Carvalho & Bessett 2015). Since the imperfect, conditional, and subjunctive are morphologically indistinct in their first and third-person singular inflections, speakers will prefer overts to prevent ambiguity in referent identification (Hochberg 1986; Shin 2014). The other TMAs, having more distinct verbal inflections, will promote null SPs.

e. Morphological ambiguity: Morphologically ambiguous forms will be more likely to appear with overt SPs than unambiguous forms (Hochberg 1986; Travis 2007; Prada Pérez 2009; Erker & Guy 2012; Lastra & Martín Butragueño 2015; Michnowicz 2015). Like TMA, verb inflections that are less distinct (i.e. lack explicit person marking) will be more likely to appear with overts to disambiguate the subject referent (Hochberg 1986).

f. Verb class: Mental process and stative verbs will favor overt SPs while activity and *Verba Dicendi* (communication verbs) will disfavor them (Travis 2007; Orozco & Guy 2008; Erker & Guy 2012; Otheguy & Zentella 2012; Orozco 2015). Due to their implied constrastive function and semantics of expressing a point of view (Silva-Corvalán 1994a), mental process verbs will promote overt SPs, which are used to assert the role of the speaker (Travis 2007).

g. Verbal mood: Subjunctive forms will disfavor overt SPs and indicative forms will show little effect (Lastra & Martín Butragueño 2015). Such disfavoring of overts may be due to the subjunctive's subordinate function or imperative-like nature in certain contexts, as discussed in Chapter 3.

h. Specificity: This factor is more difficult to predict since it has not consistently been found to be significant across studies. Based on the findings for Mexican Spanish, however, I predict it to not have a significant effect on subject expression in general. Nevertheless, it may interact with person/number and show significant effects in this regard. For example, second-person singular specific forms may favor overt SPs while second-person singular nonspecific forms disfavor them (see Solomon 1999; Michnowicz 2015).

i. Polarity: Affirmatives will favor overt SPs while non-affirmatives disfavor them (Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016). As Lastra and Martín Butragueño (2015) posit, the disfavoring effect of overts with non-affirmative structures may be due to the clustering together of negative clauses that are simultaneously in coreferential contexts, thus promoting null SPs. In other words, polarity may be constrained by whether or not there is a switch in subject referent. (same vs. switch

reference). Moreover, the presence of preverbal elements (i.e. negation) could disfavor overt SPs (Geeslin & Gudmestad 2016).

j. Style: Second-part style (i.e. longer speech turns, more conversation-like) will slightly favor overt SPs while First-part style (i.e. more dialogic pairs, more interview-like) will slightly disfavor overt SPs (Lastra & Martín Butragueño 2015). While no explanation for the motivation of this pattern has been proposed, I would postulate that speakers are more likely to follow prescriptive grammar norms (e.g. avoid using overt SPs unnecessarily) and self-monitor their speech more in the beginning of the interview, but shift away from this behavior as the interview progresses into a more casual style.

Social variables

a. Age: There have been conflicting findings for age in the subject expression literature, with some finding it to be significant (Orozco & Guy 2008; Carvalho & Child 2011; Alfaraz 2015; Lastra & Martín Butragueño 2015) and others not finding an effect (Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012). The direction of effect has also varied, with some finding that older speakers favored overt SPs and younger speakers favored nulls (Orozco & Guy 2008; Carvalho & Child 2011; Lastra & Martín Butragueño 2015), and others finding the exact opposite result (Flores-Ferrán 2002). Nevertheless, if an effect is observed, I predict that, on one hand, older speakers will favor overt SPs based on previous findings for Mexican Spanish (Lastra & Martín Butragueño 2015). On the other hand, younger speakers in the current data who had schooling in the U.S. and whose English level is higher may exhibit a preference for overt SPs.

b. Gender: There have also been conflicting findings for the relationship between gender and subject expression. However, several studies have found it to be a significant predictor of pronoun use such that women are more likely to produce overt SPs than men (Bayley & Pease-Alvarez 1996; Solomon 1999; Carvalho & Child 2011; Otheguy & Zentella 2012; Shin & Otheguy 2013; Alfaraz 2015). Thus, if an effect is observed, I predict the same pattern. This gender effect has been explained in various ways. For example, it could suggest a change in progress (Otheguy & Zentella 2012), be linked to prestige (Alfaraz 2015), or, in the U.S., be due to women's extensive contact with U.S. born bilinguals who are high pronoun users (Shin & Otheguy 2013).

c. Age of arrival: Considering the inconsistent findings of AOA in previous research (e.g. not significant in NYC for Flores-Ferrán [2004], but significant in NYC for Otheguy & Zentella [2012]), in addition to the fact that no speaker in the current data arrived to the U.S. as a child (and none are U.S. born), I do not anticipate an AOA effect.

d. Length of residency: Because of the relatively wide distinctions in LORs in the current data (e.g. some with only two years in the U.S. and others with 25 years in the U.S.) as well as the previous findings that speakers with longer LORs tend to produce significantly more overt SPs than those with shorter LORs (Flores-Ferrán 2004; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012), I predict the same direction of effect for the present study.

e. English proficiency: Due to conflicting findings in the literature, namely that studies have found both higher (Otheguy & Zentella 2012) and lower overt SP frequencies (Silva-Corvalán 1994a; Torres Cacoullos & Travis 2010) among those with greater English proficiency, it is difficult to hypothesize a particular direction of effect.

f. Preferred media language: Those who prefer English for media may favor overt SPs and those who prefer Spanish may favor nulls. To my knowledge, this factor has not been examined in previous research on Spanish subject expression.

g. Educational level: Although education level is not commonly found to be significant, it is also highly understudied, thus motivating its inclusion in the current study. Lastra and Martín Butragueño (2015) found that, while not statistically significant, speakers in Mexico City with lower education levels showed higher frequencies of overt SPs than those with higher levels, which makes sense if one considers prescriptive grammar rules. Since prescriptive grammar norms for SPs (which presumably are taught in schools) emphasize to only use overt SPs when necessary (e.g. for emphasis or contrast) and to avoid redundant use of pronouns (Butt & Benjamin 2004:130), speakers with more education may be more likely to follow such norms. The findings of both Ávila-Jiménez (1996) and Otheguy and Zentella (2012) are consistent with this explanation in that they found a disfavoring of overt SPs among the college educated. I predict similar patterns for the present study.

h. Socioeconomic status (SES): This factor is also relatively understudied. Among the researchers that have analyzed it, some have found it to be a significant predictor and others found a lack of significance. Based on the findings of Otheguy and Zentella (2012) for Mainlanders in NYC, if an effect is observed in the current study, those with a lower SES will favor overt SPs while those with higher SESs will disfavor them.

The above predictions are summarized below along with a comparison with findings of previous studies of monolingual Mexican Spanish (See Table 4.4).

Table 4.4: Summary of predictions with comparisons to previous findings[51]

	Current study (Roswell, GA)	**Lastra & Martín Butragueño (2015) (Mexico City)**	**Orozco (2016) (Xalapa, Mexico)**	**Michnowicz (2015) (Yucatan, Mexico)**	**Shin & Van Buren (2016) (Washington/Montana)**
Linguistic variables	*Overt SPs will be favored with…*	*Overt SPs favored with…*	*Overt SPs favored with…*	*Overt SPs favored with…*	*Overt SPs favored with…*
Person/number	1sg; sg	1sg; 3sg	1sg; 3sg	2sg (definite); 1sg; 3sg;	1sg; 3sg
Switch reference	Change in subject referent	Change in subject referent; partial switch	--	Change in subject referent	Change in subject referent
TMA	Imperfect; conditional; subjunctive	Imperfect	Imperfect	--	Imperfect
Morphological ambiguity	Morphologically ambiguous forms	Morphologically ambiguous forms	--	Morphologically ambiguous forms	--
Verb class	Mental process; stative	[n.s.]	Copulative; perception; motion	[n.s.]	Mental process
Verbal mood	Indicative	Indicative	--	--	--
Specificity	[n.s.]	[n.s.]	--	--	--
Polarity	Affirmative	Affirmative	--	--	--
Style	Second-part (conversational, casual)	Second-part (conversational, casual)	--	--	--
Social variables	*Overt SPs will be favored with…*	*Overt SPs favored with…*	*Overt SPs favored with…*	*Overt SPs favored with…*	*Overt SPs favored with…*
Age	Older speakers	Older speakers	Older speakers	[n.s]	--

[51] [n.s] indicates that a particular factor was not significant, and a blank cell indicates that a particular factor was not included in the study.

Gender	Women	[n.s.]	Men	[n.s.]	--
Age of arrival (AOA)	[n.s.]	--	--	--	--
Length of residency (LOR)	Speakers with longer LORs	--	--	--	--
English proficiency	No particular hypothesis	--	--	--	--
Preferred media language	Speakers who prefer English	--	--	--	--
Educational level	Speakers with less than a college education	[n.s.]	--	--	--
Socioeconomic status (SES)	Speakers with a lower SES	--	--	--	--

4.7 Summary

This chapter has explained and illustrated the specific methodological steps that were carried out in order to analyze SP variation in the Spanish of Roswell, Georgia, beginning with the data collection process involving sociolinguistic interviews by means of both random sampling and judgment (i.e. "snowball") sampling at a Latino plaza, a church, and an ESL school. I then described the sample of 20 Mexican speakers employed for the present investigation including their sociodemographic backgrounds (gender, age, LOR, AOA, education, etc.) as well as their language experiences and degree of contact with English using a 'Contact with English' (CWE) Index (Travis & Torres Cacoullos 2013). The CWE Index suggests that, in general, the speakers have a greater degree of contact with Spanish than they do

English as measured by their English proficiency, social contacts, preferred media language, and education in English.

The qualitative and quantitative methods were then discussed. The qualitative methods involved circumscribing the variable context for subject expression, that is, locating tokens that exhibit overt/null variation in the interviews to analyze, and excluding those tokens for which there is (near) categorical subject expression (i.e. appearing exclusively with one of the two variants). There was also a discussion of contrastive contexts with several illustrating examples from the current data demonstrating the presence of overt/null variation in these contexts. Then, the use of lexical subjects was explained along with the methods for circumscribing the variable context and the illustration of lexical subject tokens. Next, I presented the set of linguistic and social variables that are employed in the present study including how each of them were coded or operationalized. The quantitative methods involve the use of Rbrul (Johnson 2009) to perform a multivariate analysis of the influence exerted by the linguistic and social variables on subject variation, which is the main tool being employed to assist in determining the statistical significance and amount of variation explained by each factor. Finally, the predictions regarding the forthcoming results were presented, particularly pertaining to the overt SP rate and the factors governing pronoun variation in Roswell.

The following chapter will discuss the results of various aspects of subject expression in the Roswell data, including the overall pronoun rate, correlations between the aforementioned independent variables and pronoun usage, the patterns observed for lexical subjects, and issues of potential contact-induced language change.

CHAPTER 5

RESULTS

In this chapter, I first present the results concerning the overall pronoun rates of Roswell Spanish, comparing them with other Spanish dialects. Next, I discuss the findings regarding the independent linguistic and social variables that had a statistically significant effect (and those that did not) on variable subject expression in this variety, again making comparisons with other varieties, primarily that of Mexico City Spanish, but also comparing and contrasting the findings with the broader sociolinguistic literature on Spanish subject expression. Aside from discussing the individual effects of the independent variables, I also present findings that highlight interactions between the variables. Finally, although pronominal subjects are the primary focus of this chapter, I also present and illustrate the data and findings for the analysis carried out on lexical subjects.

5.1 Overall pronoun rates

Table 5.1 below shows the frequencies of overt and null SPs produced by the 20 speakers. Out of 4,649 total verbs, 1,239 appeared with overt SPs and 3,410 appeared with null SPs, indicating an overall overt pronoun rate of 27% for Mexicans in Roswell.

Table 5.1: Distribution of overt and null SPs

	N Verbs	**%**
Overt SPs	1,239	27
Null SPs	3,410	73
Total	4,649	100

When we compare this rate to other Mexican dialects (see Table 5.2 below), we see that it is notably higher than most monolingual Mexican varieties (with the exception of Xalapa), making Roswell Spanish less conservative in terms of overall pronoun use.[52] For instance, we see a five percent rate increase between Mexico City (22%) and Roswell (27%). This comparison and its implications will be further discussed in Chapter 6. Next, I will discuss the linguistic and social variables that influence SP variation in the Roswell data.

Table 5.2: Comparison of pronouns rates between Roswell Spanish and monolingual Mexican varieties

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

5.2 Linguistic variables conditioning pronoun use

For the multivariate analysis in Rbrul, I included all nine linguistic variables discussed in Chapter 4: *Person/Number, Switch Reference, Tense/Mood/Aspect* (*TMA*), *Morphological Ambiguity, Verb Class, Polarity, Speech Style, Specificity,* and *Verbal Mood*. I will first note

[52] Some researchers use the term *conservative* to indicate a resistance toward the expected progression of language change, which in the case of SPs is a general diachronic change toward higher overt pronoun usage in Romance (Erker & Guy 2012). For example, Lastra and Martín Butragueño (2015) classify the relatively low 22% overt SP rate in Mexico City as a "conservative dialect" (p. 53). Other researchers use different terms, such as *archaic*, to describe relatively low SP rates (e.g. Erker & Guy [2012:531]). For the purposes of the present investigation, "less conservative" refers to Roswell speakers using more overt pronouns overall than expected.

some methodological modifications before moving on to discussing the effects of the variables. Regarding *Person/Number*, the third-person plural nonspecific tokens were discarded due to near categoricity of null pronoun use (98% null). Additionally, the low number of tokens for *usted* (N = 12) and *ustedes* (N = 1) were removed. Regarding *TMA*, several categories were collapsed into one named *all other TMAs* due to low token counts of these forms. These categories included the periphrastic future, conditional, present and past subjunctive, and present perfect. The three synthetic future tokens in the data (*podrán; tú tendrás; ¿qué te diré?*) were discarded. Lastly, I combined the categories of stative and communicative for *Verb Class* because they patterned very similarly, showing no statistically significant difference between them.[53]

In the Rbrul analysis, six variables out of the nine were selected as significant: *Person/Number, Switch Reference, Morphological Ambiguity, Verb Class, Polarity*, and *Style*. *TMA, Specificity*, and *Verbal Mood* were not significant. Table 5.3 below shows the variable hierarchy of the significant factors from strongest to weakest effect. The top-ranked factor, *Person/Number*, has the most predictive power for variable SP usage while the bottom factor, *Style*, is the least powerful. The variable hierarchy was determined based on factor weight ranges, which are an indication of the relative strength of each variable (Tagliamonte 2006, 2012). The higher the range is for a given variable, the greater amount of variance that is explained by the variable (see discussion below and Table 5.4 for the specific factor weights and greater detail of how the range was calculated).

[53] The rate of overt SPE was 25% for stative verbs and 27% for communicative verbs.

Table 5.3: Variable hierarchy as compared to Mexico City[54]

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

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

The above comparisons between the two variable hierarchies of Roswell and Mexico City follow the tradition of *comparative sociolinguistics*. The main objective of this type of methodology, according to Tagliamonte (2002), is to examine

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

> the connection (relationship) of linguistic variation in one body of materials to another. This requires a methodology that, first, enables the many different influences on linguistic features to be disentangled through systematic examination of their behavior, and, second, that situates and explains the linguistic features through comparison with like features in related varieties (p. 729).

One of the specific types of comparisons in utilizing this methodology involves the differences and similarities in the order of factors between two or more groups of speakers (illustrated here in Table 5.3 above). Different rankings between groups suggests that each group has a different variable grammar, that is, the specific set of probabilistic constraints (or variable rules) that determine which variant is more likely to be used for a given linguistic variable (Tagliamonte 2002; Meyerhoff 2009). Similarities or differences in variable grammars assist us in determining both parallels and divergences between one language variety and another and whether or not language change is occurring. Comparative sociolinguistics is a common practice in particular for studies of subject expression in contact situations in order to explore the role of language contact (e.g. Torres Cacoullos & Travis 2010; Otheguy & Zentella 2012; Michnowicz 2015; Prada Pérez 2015; Bessett 2018).

Aside from simply comparing the order of variables in the hierarchy, another type of comparison that has been utilized in previous SP research is the comparison between the range values for each variable (e.g. Otheguy & Zentella 2012; Michnowicz 2015; Prada Pérez 2015). For instance, Otheguy and Zentella (2012:169) found that the ranges for certain variables (e.g. Switch Reference, Clause Type) were smaller for New York born/raised bilinguals when compared to immigrant newcomers, which they argued to indicate a reduction in strength and weakened sensitivity to these variables for bilinguals.

Returning to Table 5.3, we see the ranges presented for each variable in both the Roswell Mexican Spanish variety and Spanish in Mexico City. By utilizing the range values, we can corroborate some of the observations above regarding the differences in ranking in addition to revealing further evidence of how the two Spanish varieties are different or the same. For instance, the range for *Morphological Ambiguity* in Roswell (20) suggests that it is a stronger constraint in this variety than it is for Mexico City speakers (13). On the other hand, *Polarity* is a weaker constraint in Roswell (10) as compared to Mexico City (23). As mentioned above, *Style* is ranked last for both varieties; in fact, we see an identical range value (8) in both Roswell and Mexico City. Moreover, an additional difference is revealed for *Person/Number* and its relation to *Switch Reference* upon comparing the ranges. Particularly, we see that *Person/Number*, although ranking first in both varieties, appears to be much stronger in Roswell (54) than in Mexico City (32). It should further be noted that the predictive power compared to *Switch Reference* more than doubles in Roswell (*Person/Number* = 54 vs. *Switch Reference* = 26, respectively) and that these two variables differ only slightly in relative importance for Mexico City (*Person/Number* = 32 vs. *Switch Reference* = 31). Thus, in terms of relative predictive power, *Person/Number* is more on par with *Switch Reference* in Mexico City while in Roswell there is a much greater distinction between the two variables. In other words, the variable grammar represented by Roswell speakers does not exhibit the same degree of strength to *Switch Reference* as the grammar of Mexico City speakers and instead shows a greater reliance on *Person/Number* as a constraining factor of SP variation.

Table 5.4 below presents the constraint hierarchies for each factor group in order to show more specifically how each individual variable constrains subject expression in the data. The first column shows each factor group along with their particular values, and the second column

presents the factor weights (FW) for each constraint from highest to lowest probability of appearing with an overt SP. When a FW is closer to 1, this indicates a relative favoring of overt SPs. When it is closer to 0, it generally indicates a disfavoring of overt SPs (see Tagliamonte 2006:145, 156). When FWs are close to .5, however, this indicates that the constraint has a neutral effect (i.e. neither a favoring nor disfavoring) (Michnowicz 2015:110). Thus, for example, Table 5.4 shows that the FW of .75 for the third-person singular constraint for the *Person/Number* factor group indicates a favoring of overt SPs while the FW of .21 for first-person plural tokens indicates a disfavoring of overt SPs. The FW of .47 for second-person singular indicates a neutral effect. Using the FWs, the range value was established for each factor group that indicates its relative strength, as seen in column one. The range was established by calculating the difference between the highest and lowest FWs (Tagliamonte 2006:242). Thus, the range for *Person/Number* is 54 (the difference between 75 and 21) and the ranges descend for each subsequent factor group. Furthermore, the third column presents the proportion of overt SPs in the data for each context and the fourth column indicates the total token count in each context, including both overt and null SPs. Finally, the fifth column shows the *p*-value for each factor group, indicating whether it is statistically significant. All *p*-values that are less than .05 indicate that a factor group is statistically significant (the first six factor groups in Table 5.4), and *p*-values higher than .05 indicate a lack of statistical significance (the last three factor groups in Table 5.4).[55] I will now discuss each factor group in turn.

[55] See Tagliamonte (2006) for further discussion of interpreting results of a multivariate analysis.

Table 5.4: Constraint hierarchy

Factor Group	Factor Weight	% Overt	Total N tokens	*p*-value
Person/Number				1.71e-59
3s	.75	38%	372	
1s	.69	36%	2565	
2s	.47	14%	489	
3pl	.39	12%	387	
1pl	.21	7%	836	
RANGE	54			
Switch Reference				6.39e-41
Switch	.64	35%	1945	
Partial switch	.49	28%	247	
Same	.38	20%	2457	
RANGE	26			
Morphological Ambiguity				5.54e-05
Ambiguous	.60	46%	533	
Unambiguous	.40	24%	4116	
RANGE	20			
Verb Class				1.95e-07
Mental	.57	37%	962	
Stative/communicative	.50	25%	1896	
Activity	.43	22%	1791	
RANGE	14			
Polarity				2e-04
Affirmative	.55	27%	4063	
Non-affirmative	.45	23%	586	
RANGE	10			
Style				8.45e-05
Second part (conversation-like)	.54	28%	2988	
First part (interview-like)	.46	24%	1661	
RANGE	8			

[TMA]				0.0992
Preterit	[.55]	29%	886	
All other TMAs	[.50]	28%	429	
Present	[.49]	25%	2584	
Imperfect	[.46]	30%	750	
[Specificity]				0.194
Specific	[.56]	28%	4213	
Nonspecific	[.44]	13%	436	
[Verbal Mood]				0.27
Indicative	[.54]	27%	4497	
Subjunctive	[.46]	24%	152	
Speaker (random) Std. Dev. .40				

5.2.1 Person/Number

As shown in Table 5.4, both third-person singular and first-person singular verbs favor overt SPs with the former constraint showing the strongest effect (FW = .75, see example (1) below). That these two forms favor overt SPs the most is generally consistent with previous research (e.g. Silva-Corvalán 1994a; Flores-Ferrán 2002; Blackwell & Quesada 2012; Shin 2012; Lastra & Martín Butragueño 2015; Limerick 2017). There are, however, exceptions to this pattern. For instance, Otheguy and Zentella (2012:163) found that these two forms showed a neutral effect (not significant) and were outweighed by third-person *usted* and *uno*, and by second-person singular verbs, all of which favored overt SPs among NYC immigrant newcomers.[56]

[56] I suspect that the mere inclusion of *usted* and *uno*, two forms that are typically excluded in subject expression studies, helps to explain why they outweighed first-person singulars (*yo*) and third-person singulars (*él/ella*) to such a high degree that these latter forms no longer showed an effect. *Usted* and *uno* tend to be overt at a rate substantially higher than that of *yo* and *él/ella* (e.g. *usted*: 56% [Carvalho & Bessett 2015]; 67% [Alfaraz 2015]; *uno*: 85% [Cameron 1992]).

(1) Third-person singular with overt SP:

me gusta que um, ya hable los dos idiomas aunque, le cuesta mucho trabajo, me gusta que, una vez que ***él*** *estudia aquí,* ***él*** *va a poder ir a trabajar a donde sea en el mundo* [F34Mex]

'I like that um, he already speaks both languages even though, it's hard for him, I like that, once that he studies here, he's going to be able to go and work anywhere in the world'

First-person plural verbs, in contrast, strongly disfavor overt SPs (FW = .21), followed by third-person plurals, as illustrated below:

(2) First-person plural with null SP:

I: *OK ¿tenías como quehaceres que hacer en la casa?*

R: *En México sí ah .. **íbamos** a la escuela .. **llegábamos** y mi mamá me decía…*' [M28Mex]

'In Mexico yeah uh .. we would go to school .. we arrived and my mom would say to me…

Finally, second-person singular verbs show a neutral effect, a pattern that diverges from the pan-Hispanic trend of singular forms favoring overt pronouns and plural forms favoring nulls (Bentivoglio 1987; Cameron 1992; Silva-Corvalán 1994a; Bayley & Pease-Alvarez 1996; Flores-Ferrán 2002; Otheguy, Zentella, & Livert 2007; Prada Pérez 2009; Carvalho & Child 2011; Abreu 2012; Alfaraz 2015; Orozco 2015; among others). However, this pattern follows other Mexican varieties in bucking this trend, namely first-generation Mexicans in NYC who show a neutral effect (Shin & Erker 2015) and both Mexico City (Lastra & Martín Butragueño 2015) and Xalapa (Orozco 2016) speakers who disfavor overt pronouns with second-person

singular verbs. Furthermore, mirroring what Orozco (2016) found for Xalapa Spanish, third-person singulars outweigh first-person singulars, an atypical pattern for Mexican Spanish in general. This could be due to the aforementioned functions of referential tracking and the need to disambiguate between third-person subjects when there is a switch in referent, as discussed in Chapter 3 (Section 3.4.2).

5.2.2 Switch Reference

In line with previous studies, *Switch Reference* contexts favor overt SPs, and same reference contexts disfavor overt pronouns (e.g. Silva-Corvalán 1994a; Bayley & Pease-Alvarez 1997; Travis 2005; Prada Pérez 2009; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015), as in the following examples:

(3) Switch Reference with overt SP:

*los conocemos en, los, nos <u>conocimos</u> solteros, **ellos** fueron a mi boda cuando nos casamos* [M41Mex]

'we meet them at, we met each other as single people, they went to my wedding when we got married'

(4) Same reference with null SP:

*…totalmente, el idioma, si no <u>hablas</u> inglés no **puedes** conseguir tan fácil un buen trabajo.*

[F26Mex]

'absolutely, the language, if you don't speak English you can't get a good job as easily'

Moreover, partial switches are neutral (FW = .49), which is also consistent with previous research (Orozco & Guy 2008; Otheguy & Zentella 2012; Orozco 2015). Specifically, overt SPs occur in this context 28% of the time, about the midpoint between 20% with same reference and

35% with switch reference. The idea is that overt SPs are produced less for purposes of referential tracking in partial switch contexts, due to an immediately preceding coreferential object, even less for same reference contexts since the subject was just mentioned, and are produced the most for reference tracking when both the previous subject and object were noncoreferential (Shin & Otheguy 2009). Such relation between pronoun use and referential tracking is also associated with the notions of *accessibility* and *salience* discussed in Chapter 3 (Givón 1983; Ariel 1994). In this respect, switch reference contexts make referents less accessible/salient while in cases of same reference the referent, just being mentioned in the previous clause, is most accessible. Thus, the former context promotes overt SPs while the latter promotes nulls.

5.2.3 Morphological Ambiguity

Also consonant with previous research (e.g. Hochberg 1986; Travis 2007; Prada Pérez 2009; Erker & Guy 2012; Lastra & Martín Butragueño 2015; Michnowicz 2015), ambiguous verb forms favor overt SPs with a FW of .60 while unambiguous forms promote null SPs with a FW of .40, as in the following examples:

(5) Morphologically ambiguous verb with overt SP:

*entonces me especialicé como ocho años... trabajando con gemelos, y **yo iba** a las casas de los americanos* [F34Mex]

'so I specialized like eight years…working with twins, and I went to the homes of Americans'

(6) Morphologically unambiguous verb with null SP:

*…programas de televisión **veo** en, inglés y español... este ambos, eh el notic- noticias en en inglés y español* [M33Mex]

'…TV shows I watch in, English and Spanish…uh both, uh the ne- news in in English and Spanish'

This finding appears to lend support to the Functional Hypothesis (Hochberg 1986), which predicts an increased use of overt SPs with ambiguous verbal morphology for first-person singular and third-person singular forms in order to compensate for the absence of explicit person marking. The percentage of overt SPs nearly doubles from 24% with unambiguous forms to 46% with ambiguous forms. Further study of the discourse context would be needed, however, to determine whether or not there is contextual ambiguity (i.e. ambiguity not resolved by the context; see Ranson [1991]).

For instance, Ranson (1991) explored the role of ambiguity in SP variation by examining various contextual markers in her quantitative analysis of Andalusian Spanish, finding that more overt SPs were produced when contextual markers that facilitated referent identification were absent from the discourse. In particular, this author examined both textual and nontextual reference, the former being when a subject is evoked through previous mention in discourse, and the latter being when a subject is evoked from the extralinguistic context (i.e. the current situation). Consider the following example, which illustrates nontextual reference (Ranson 1991:142):

(7) Nontextual reference:

Pónle la dirección. Pónsela tú, que conoce tu letra.

'Write down the address for her. You write it down for her, because she knows your handwriting.'

In situations like the one presented in (7), "two other participants refer to a third party who is present" (p. 141). Because of the nature of this situation, in which the speaker requests that the listener do something for a third party (in this case, the third party is the researcher), the referent of *conoce* can be clearly identified, despite the absence of prior mention of the subject in the discourse itself. As Ranson (1991) explains, "[w]hat distinguishes nontextual reference from textual reference in all cases is that the information needed to interpret person comes solely from the present situation, and the subject of the verb has not been mentioned in previous discourse" (p. 142). Thus, for textual reference, the subject is previously mentioned; additionally, there are no competing referents between the previous mention and the target reference, as in the example below (Ranson 1991:142):

(8) Textual reference

Ese fue también cabo. Yo fui cabo para un desfile…Como era alto, me ponían delante…

'That one, too, was leader. I was leader for a parade. Since I was tall they put me in front…'

In (8), the referent of *era* is *yo*, which was mentioned in the preceding clause. There are also no intervening subjects to whom the verb *era* could have referred, so although *era* is *morphologically* ambiguous, it can be argued that the contextual marker in the previous clause (the textual reference *Yo fui*) disambiguates the referent of *era*. Aside from the category of reference, the other main type of disambiguating contextual markers that Ranson highlights

pertains to the interlocutors' shared knowledge, which may either be textual or nontextual (e.g. general world knowledge or specific knowledge shared between the interlocutors). Therefore, in the absence of such contextual markers that assist in clarifying person (whether they are based on reference or shared knowledge), the subject referents of *morphologically* ambiguous verbs can also be considered *contextually* ambiguous in that identifying the referent is nearly impossible without any type of textual or nontextual markers.

5.2.4 Verb Class

Regarding *Verb Class*, mental process verbs slightly favor overt SPs (FW = .57) and activity verbs favor null SPs (FW = .43). Stative/communication verbs show neither a favoring nor disfavoring, with a FW of .50. The pattern that mental verbs favor overt SPs follows the trend in the variationist literature (e.g. Silva-Corvalán 1994a; Travis 2007; Torres Cacoullos & Travis 2010, 2011; Otheguy & Zentella 2012; Orozco 2015) and is perhaps attributable to the implied contrast expressed by the speaker and the fact that speakers convey their point of view with these types of verbs (Silva-Corvalán 1994a), as seen in the following example:

(9) Mental verb with overt SP:

…muchas actividades, el día de la Virgen de Guadalupe porque ***nosotros creemos*** *en la Virgen de Guadalupe, eh, en México*… [F34Mex]

'…many activities, the day of the Virgin of Guadalupe because we believe in the Virgin of Guadalupe, uh, in Mexico…'

The finding for activity verbs (example 10) also aligns with previous studies that have reported a disfavoring of overt SPs (e.g. Orozco & Guy 2008; Erker & Guy 2012; Otheguy & Zentella 2012; Orozco 2015).

(10) Activity verbs with null SP:

*Bueno me **levanto**, @, este, me **arreglo**, y **voy** a mi trabajo, **entro** a las diez, once de la mañana…* [F30Mex]

'Well I get up, @, uh, I get ready, and I go to my job, I go in at ten, eleven in the morning…

However, the effect of stative and communication verbs is inconsistent with previous research since these verb classes showed a neutral effect in the present study and a favoring of overt SPs in other studies (e.g. Travis 2007; Orozco & Guy 2008; Erker & Guy 2012; Otheguy & Zentella 2012; Orozco 2015).

5.2.5 Polarity

Despite its overall modest effect, the results for *Polarity* show that affirmative clauses slightly favor overt SPs and that non-affirmative forms (negation and interrogatives) slightly favor null SPs. This finding agrees with the results of Lastra and Martín Butragueño (2015) and Geeslin and Gudmestad (2016). The following two examples illustrate such patterns:

(11) Affirmative clause with overt SP

*Cuando **yo era** niño **yo empecé** a trabajar a los siete años* [M32Mex]

'When I was a kid I started working at age seven'

(12) Non-affirmative clause with null SP:

*…pero nada más tuve primer y segundo, ya **no acabé** el tercer año* [M52Mex]

'…but I only had first and second, I didn't finish the third year'

As discussed in Chapter 3, it is possible that subject coreference across contiguous negated clauses explains the favoring of null SPs (Lastra & Martín Butragueño 2015). It might also be the case in the current data that, as was found by Geeslin and Gudmestad (2016), the *Polarity*

effect is only operative for first-person singular forms. Moreover, this variable may be constrained by *Verb Class* (Travis & Torres Cacoullos 2012). Interactions between *Polarity* and other variables such as *Switch Reference*, *Person/Number*, and *Verb Class* will be discussed below.

5.2.6 Style

Finally, albeit the weakest effect, *Style* also significantly constrains pronoun use, with second-part style (more conversational, longer speech turns) slightly favoring overt SPs and first-part style (more interview-like, shorter speech turns) slightly favoring nulls (see examples 13 and 14 below). This pattern is consistent with the findings of Lastra and Martín Butragueño (2015). Insofar as second-part style represents a more casual style, this result is also consistent with Ávila-Jiménez's (1996) findings showing a favoring of overt SPs with more casual speech. If speakers are less self-aware and careful of their speech, they may not focus on prescriptive rules to not overuse overt pronouns (Butt & Benjamin 2004:160). As mentioned in Chapter 3, the effect of *Style* may be constrained by *Person/Number*. Such a potential interaction will be explored below.

(13) First-part style with null SP:

R: ahm ...(3) pues ***fui*** *criada más que todo por mi abuela… porque mi mamá…* [F25Mex]

'umm …(3) well I was raised mostly by my grandmother…because my mom…'

(14) Second-part style with overt SP:

…hasta que llegué a este país me siento así, porque ***yo digo****,* ***yo puedo*** *aprenderlo,* ***yo tengo*** *ganas de aprenderlo porque* ***yo puedo*** *llegar y pedir bien, como lo* ***hago yo*** *en mi país* [F52Mex]

'…since I arrived to this country I felt like that, because I say, I can learn it, I have a desire to learn it because I can arrive and order well, like I do in my country'

5.2.7 Comparison of constraint hierarchies with Mexico City Spanish

A further type of comparison with monolingual Mexican Spanish that is important to establish is one of constraint hierarchies, which is useful in more specifically determining the similarities and differences in the variable grammar of speakers (Tagliamonte 2002). Specifically, if we compare the direction of effect as well as the ordering of constraints within each variable to such patterns in Mexico City Spanish, we will see that there are virtually no differences between the varieties. In other words, although there are substantial differences in the *strength* of the variables as evidenced above by comparing the range values, the patterning within each variable and how subject variation is governed are virtually the same. Table 5.5 below shows the constraint hierarchies for Roswell Spanish (the same ones from Table 5.4 above) and the many parallels with the hierarchies established for Mexico City Spanish. With the exception of *Person/Number*, we see here that the constraint order is identical in both varieties. For instance, the constraints for *Switch Reference*, going from greatest likelihood to least likelihood of appearing with an overt SP, begin with *switch* with the highest factor weight, then *partial switch*, and finally *same*. Such parallel ordering continues for the other variables presented. Returning to *Person/Number*, we see a slight difference in the constraint ranking in

that third-person singular verbs supersede first-person singular verbs in Roswell while the inverse was observed in Mexico City.

In terms of the directions of effect overall, then, we see that they are very similar in both varieties, with overt SPs being favored in the following contexts: 1s and 3s persons/numbers; a switch in subject referent; ambiguous verb forms, affirmative clauses; second-part style. Thus, it is the overall strength of each variable that changes for Roswell speakers while the general probabilistic patterns that regulate subject expression remain the same. Further discussion of these similarities (as well as differences in terms of variable strength) will be presented in Chapter 6.

Table 5.5: Constraint hierarchies as compared to Mexico City[57]

Roswell, GA (present study)		Mexico City (Lastra & Martín Butragueño 2015:43)	
Factor Group	**Factor Weight**	**Factor Group**	**Factor Weight**
Person/Number		**Person/Number**	
3s	.75	1s	.58
1s	.69	3s	.57
2s	.47	2s	.40
3pl	.39	3pl	.33
1pl	.21	1pl	.26
Switch Reference		**Switch Reference**	
Switch	.64	Switch	.70
Partial switch	.49	Partial switch	.54
Same	.38	Same	.39
Morphological Ambiguity		**Morphological Ambiguity**	
Ambiguous	.60	Ambiguous	.61
Unambiguous	.40	Unambiguous	.48
Polarity		**Polarity**	
Affirmative	.55	Affirmative	.54
Non-affirmative	.45	Non-affirmative	.31
Style		**Style**	
Second part (conversation-like)	.54	Second part (conversation-like)	.54
First part (interview-like)	.46	First part (interview-like)	.46

5.2.8 The unexpected Tense-Mood-Aspect (TMA) finding

Aside from the significant variables, one of the most striking findings emerges when we consider what variables were not found to be significant, particularly *TMA*. This variable has

[57] Only common factors found to be significant in both studies are presented here.

been shown to be significant in an overwhelming number of studies in the variationist literature, and with a very clear distinction between two of its factors: imperfect forms, which typically favor overt SPs, and preterit forms, which typically disfavor such forms. There is generally a wide percentage difference of overt SPs in these two contexts in the literature, with imperfect forms having a much higher rate than preterit forms. The result in the current data, however, shows a drastic narrowing of this distinction with a virtually identical percentage of overt SPs in both contexts: 30% with imperfect and 29% with preterit (see Table 5.6).

Table 5.6: Rate of overt SPE according to *TMA* with comparison to previous studies

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

This divergence from previous data reveals two things: First, it evidences a loss of the *TMA* constraint in the Roswell data, particularly with regard to the imperfect and preterit environments.[58] Secondly, this finding contradicts Silva-Corvalán's (2001) hypothesis concerning the discourse functions of foregrounding and backgrounding discussed in Chapter 3. Specifically, Silva-Corvalán argued that more overt SPs are used with the imperfect because the focus with these forms is on the *subject* rather than the (backgrounded) action and that fewer overts are used with the preterit since the focus is more on the (foregrounded) *action* and less on

[58] A more in-depth explanation of the loss of *TMA*, beyond its mere lack of statistical significance, will be provided in Chapter 6.

the subject. Such pragmatic functions are not observed in the present data since the speakers did not exhibit differential pronoun behavior according to whether they used imperfect or preterit aspect. As noted above, the proportion of overt SPs was essentially the same in each context, only showing a one percent difference (30% for imperfect vs. 29% for preterit). The alternative explanation discussed previously for the *TMA* effect (Chapter 3) has to do with the potential ambiguity of the subject referent. Despite the lack of statistical significance for the *TMA* variable, the current results do not necessarily contradict such explanation since there was indeed a significant effect for the *Morphological Ambiguity* variable. As stated above, a closer look at the surrounding discourse would be necessary to discern contextual ambiguity (in the sense of Ranson [1991] illustrated above in section 5.2.3) from simple morphological ambiguity. At any rate, the present results lend less support to the foreground/background explanation, as argued by Silva-Corvalán (2001), since speakers clearly did not distinguish between the imperfect and preterit with regard to their pronoun usage.

5.2.9 Interactions between independent variables

Now that we have looked at the main effects of each independent variable, I will now discuss how the variables interact with each other to influence variable subject expression in order to get a more in-depth and nuanced picture of how a speaker's choice in pronoun use is determined. I have examined interactions by means of conditional inference trees (Tagliamonte & Baayen 2012) as well as cross tabs calculated using Rbrul. Conditional inference trees are useful statistical tools, using branching structures that display statistically significant interactions between the independent variables in a given dataset. They allow researchers to efficiently locate and visualize interactions in their data (Tagliamonte 2012; Tagliamonte & Baayen 2012; Schwenter 2017). The conditional inference trees were created using the recently developed

statistical program *Language Variation Suite* (Scrivner & Díaz-Campos 2016). The splits in the branching structures shown in the trees (see Figures 5.1 – 5.4 below) indicate (a) instances in which particular contexts for a given variable exhibit significantly different pronoun frequencies and (b) how this variable interacts with other variables. For example, it will be shown below in Figure 5.1 that *Verb Class* splits into mental verbs on one side, and all other verb types on the other side, indicating that mental verbs behave differently from other verbs. Additionally, the mental verbs branch splits off further to interact with *Polarity*, which then splits into affirmative and negative. What I will demonstrate is that the main effect of certain independent variables are not operative across the data set as a whole, but rather interact with other independent variables. The particular interactions tested for are based on previous SP literature that has drawn attention to the need for research on such interactions, namely those of *Polarity* with *Verb Class* (Travis & Torres Cacoullos 2012), *Polarity* with *Person/Number* (Geeslin & Gudmestad 2016), *Polarity* with *Switch Reference* (Lastra & Martín Butragueño 2015), and *Style* with *Person/Number* (Lastra & Martín Butragueño 2015). I will begin with the interaction between *Polarity* and *Verb Class*.

5.2.9.1 Polarity and Verb Class

Regarding *Polarity* and *Verb Class*, the conditional inference tree shown in Figure 5.1 below illustrates the following information: *Verb Class* is split between mental verbs on one hand (which favor overt SPs) and activity, communicative, and stative verbs on the other hand. This split indicates that mental verbs behave significantly different from the other verb classes, a result already seen in the Rbrul analysis above. What was not revealed above however, is the interaction we see between mental verbs and *Polarity* on the right side of the tree, which shows that mental verbs favor overt SPs when they are affirmative but not when they are negative.

Specifically, we see an overt rate of 41% in affirmative contexts and only 18% with negatives, as illustrated by the black shaded regions of the middle and right-hand bars of the tree, respectively. In contrast, the other three verb classes on the left side of the tree do not interact with Polarity in any way as we do not see any additional branches for affirmative and negative.

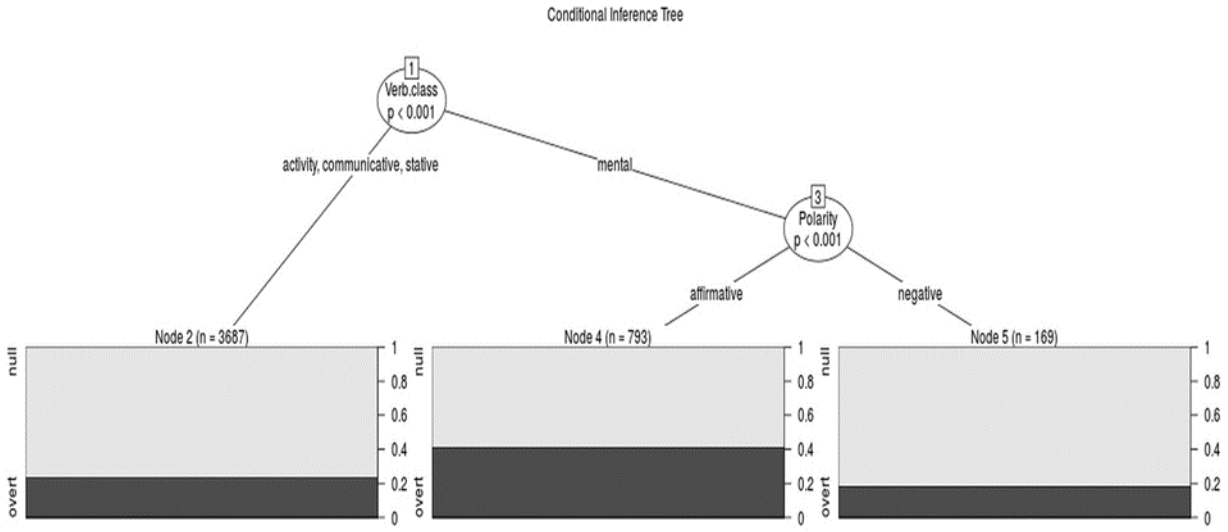

Figure 5.1: Conditional inference tree for *Polarity* and *Verb Class*

From the perspective of the *Polarity* effect discussed in the Rbrul analysis (overt SPs favored with affirmatives), the tree also shows that such a pattern applies only to mental verbs and is not operative for the other verb classes. The cross tab shown in Table 5.7 below shows the relationship between *Verb Class* and *Polarity* in more detail with the specific overt pronoun frequencies for each *Verb Class* and *Polarity* category. The cross tab indicates that it is the mental verbs that show the greatest distinction between affirmative and negative polarity (a difference of 23 percentage points) and that the other verb classes do not show such a distinction. In other words, *Polarity* is only operative for mental verbs. It may be the case that the particular

negative polarity token ***Ø no*** *creo* '(I) don't think' is driving this pattern, as was found by Travis and Torres Cacoullos (2012:741). Further research is required, however, to uncover potential verb lexeme effects.[59]

Table 5.7: Rate of overt SPE according to *Polarity* and *Verb Class*

	Activity	Communicative	Mental	Stative
Affirmative	22%	26%	**41%**	26%
Negative	26%	31%	**18%**	22%

5.2.9.2 Polarity and Person/Number

A significant interaction was also found between *Polarity* and *Person/Number*. Specifically, both first-person singular and third-person singular forms favor overt SPs when they are affirmative (38% overt) but not when they are negative (29% overt), as seen on the right side of the tree in Figure 5.2. We do not see any interaction with *Polarity* for the other persons/numbers. The left side of the tree simply illustrates that second-person singulars and third-person plurals are significantly more likely to appear with overt SPs than are first-person plural forms, a result previously observed in the multivariate analysis. Furthermore, we again see, as was shown with the interaction with *Verb Class* in the previous section, that the *Polarity* effect applies only to certain environments. In this case those environments are first-person and third-person singular verbs.[60]

[59] A separate analysis of the verb lexeme *creo*, the most frequent lexeme among the mental verbs, showed that the negative polarity token ***no*** *creo* was categorically null (0% overt) while the affirmative token *creo* preferred overt SPs (69% overt), a finding in line with the strong disfavoring effect of negative polarity observed by Travis and Torres Cacoullos (2012) in their analysis of *creo*. However, only 3 total tokens of ***no*** *creo* were produced in the present dataset, necessitating further research with a larger data set in order to provide more conclusive results.

[60] One question that arises from these results is whether the 1sg and 3sg tokens that are also morphologically ambiguous (i.e. imperfect, subjunctive) would make the polarity effect stronger, A separate analysis of these tokens, however, showed that there was not a significant interaction between *Morphological Ambiguity* and *Polarity* ($p >$

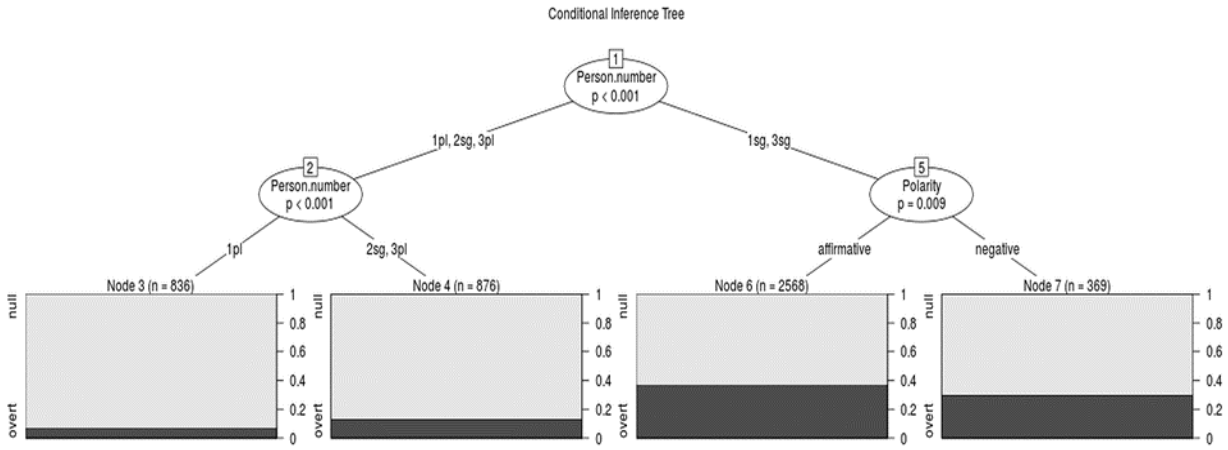

Figure 5.2: Conditional inference tree for *Polarity* and *Person/Number*

The cross tab in Table 5.8 below shows the differences in overt frequencies for each person/number according to *Polarity*. These rates further reveal that the polarity effect is greater for third-person singular verbs relative to first-person singulars, with a percentage difference of 12 and 7, respectively. The cross tab also confirms the lack of interaction between 1pl, 2sg, and 3pl and *Polarity* with very small percentage differences between affirmative and negative.

Table 5.8: Rate of overt SPE according to *Polarity* and *Person/Number*

	1sg	2sg	3sg	1pl	3pl
Affirmative	**37%**	13%	**39%**	7%	13%
Negative	**30%**	14%	**27%**	8%	10%

0.05). In other words, the polarity effect is no longer operative with cases of morphologically ambiguous 1sg and 3sg tokens.

The result that the polarity effect is operative for first-person singular forms but not for second-person singular forms is consistent with Geeslin and Gudmestad's (2016) findings for native speakers in their study.[61]

5.2.9.3 Polarity and Switch Reference

Regarding the interaction of *Polarity* with the *Switch Reference* variable, Figure 5.3 below shows that it is the noncoreferential environments (both *switch* and *partial switch*) that interact with *Polarity*. By contrast, the coreferential environments (i.e. same reference) do not show an interaction. In particular, when there is a switch in subject referent, affirmative clauses favor overt SPs over negative clauses. Moreover, affirmative contexts further interact with *Switch Reference*, showing a favoring of *switch* over *partial switch* contexts, a distinction not seen for negative contexts. The cross tab shown in Table 5.9 further reveals that it is *switch reference* contexts, as opposed to *partial switch*, that drive the polarity pattern (a wide distinction of 36% affirmative vs. 27% negative for *switch reference* as opposed to virtually no difference with 28% affirmative vs. 29% negative for *partial switches*).

[61] Geeslin and Gudmestad (2016) studied subject expression among L2 Spanish learners and compared them with native Spanish speakers from a variety of countries (Colombia, Argentina, Chile, Ecuador, Mexico, Spain, Puerto Rico, the Southwestern United States, and Uruguay).

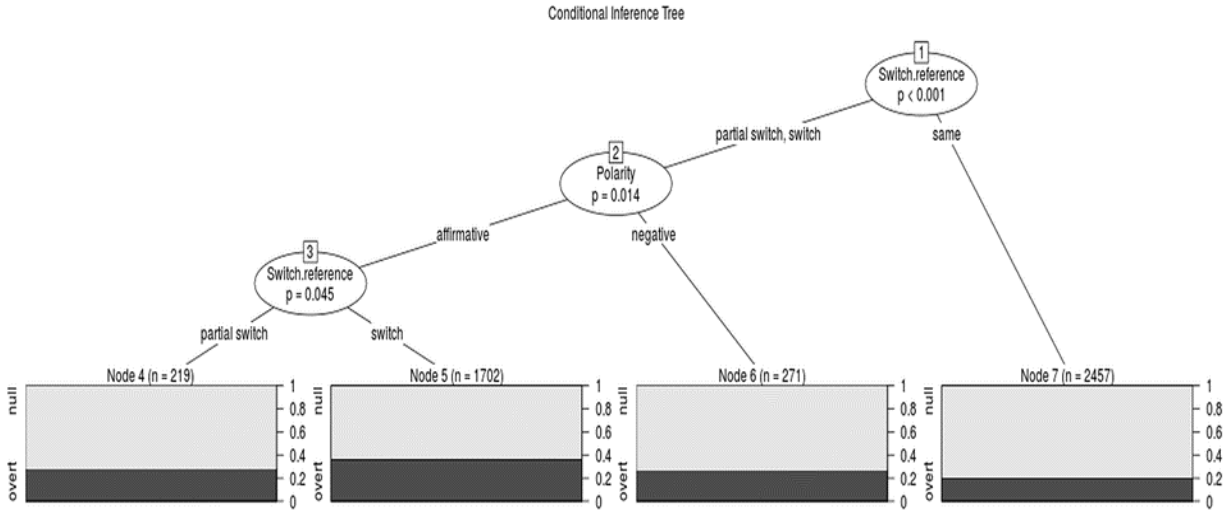

Figure 5.3: Conditional inference tree for *Polarity* and *Switch Reference*

Table 5.9: Rate of overt SPE according to *Polarity* and *Switch Reference*

	Same	Switch	Partial switch
Affirmative	20%	**36%**	28%
Negative	19%	**27%**	29%

As discussed in Chapter 3, Lastra and Martín Butragueño (2015) hypothesize that the disfavoring of overt SPs with negative clauses may be due to the clustering together of such clauses in primarily coreferential (i.e. same reference) contexts (a context which also disfavors overt SPs). This explanation is not borne out in the current data, however, since negative clauses that are in same reference contexts do not show a significantly reduced overt rate compared to affirmative clauses in same reference contexts. In fact, they are virtually the same, with 19% overt for the former and 20% for the latter, as seen in Table 5.9.

5.2.9.4 Style and Person/Number

As discussed in Chapter 3, the interaction of speaker *Style* and *Person/Number* deserves some attention (Lastra & Martín Butragueño 2015). Analysis of these two variables together in the current data indeed reveals a statistically significant interaction. As seen from the conditional inference tree below (Figure 5.4), the environment where *Style* intersects with *Person/Number* is particularly the use of first-person singular and third-person singular verbs. Specifically, third-person singulars behave significantly different from first-person singulars (41% overt for the former and 28% overt for the latter) for interview-like tokens, but not for conversation-like tokens. This means that, as the sociolinguistic interviews progressed, the difference between these two persons/numbers was neutralized; that is, third-person singular overt use decreased and first-person singular overt use increased. Thus, the observation that third-person singular outweighs first-person singular in the data (Rbrul analysis above) is only the case for interview-like (more careful) speech. One explanation for this effect is that perhaps speakers were more concerned with and more careful about facilitating referential tracking in the beginning stages of the interviews (thus more overt SPs with third-person singular which are less contrastive morphologically than first-person singular), and not as careful in the latter stages when they were paying less attention to their speech. Furthermore, a cross tab reveals that the overall *Style* effect (more overt SPs with conversation-like tokens) only applies to first-person singular and not to the other persons/numbers (See Table 5.10 below). This pattern is seen by the substantial difference in frequencies for first-person singular conversation-like (42%) vs. interview-like (28%) that is not seen for other persons/numbers. This could be due to the fact that more mental verbs (e.g. *creer* 'believe/think'; *pensar* 'think') tend to appear in the latter stages of the interview, particularly in which speakers are forming opinions on different topics. As explained

above in the discussion of the main effects in the Rbrul analysis, mental verbs favor overt SPs. The first-person singular verbs *creo* and *pienso* have the highest overt pronoun rates (69% and 67%, respectively), so the finding that the *Style* effect applies only to first-person singular forms may have to do with the more frequent occurrence of *creo* and *pienso* among the conversation-like tokens opposed to the interview-like tokens.

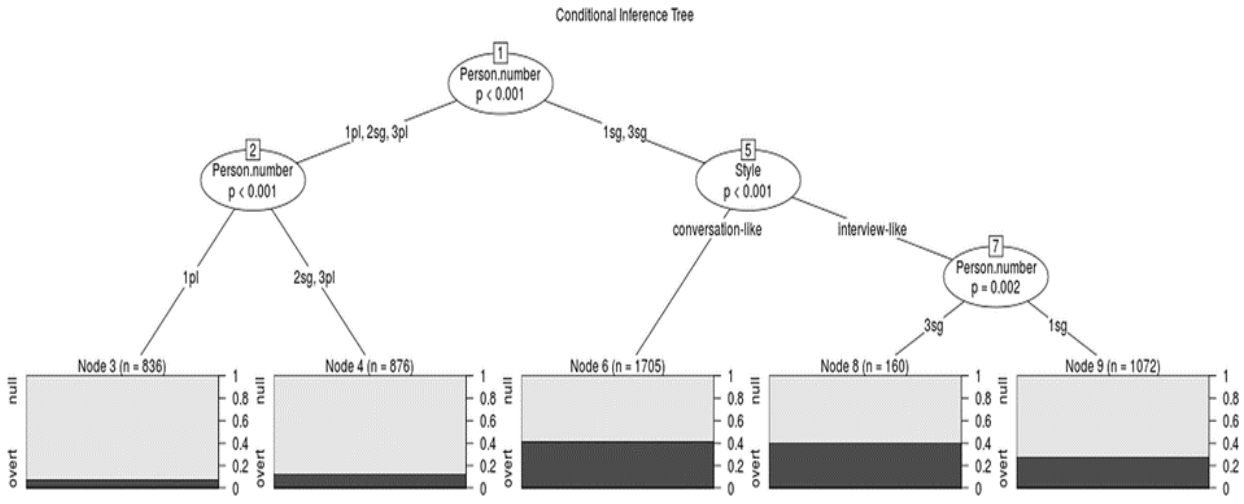

Figure 5.4: Conditional inference tree for *Style* and *Person/Number*

Table 5.10: Rate of overt SPE according to *Style* and *Person/Number*

	1pl	1sg	2sg	3pl	3sg
Conversation-like	8%	**42%**	14%	11%	36%
Interview-like	7%	**28%**	12%	17%	41%

5.3 Social variables conditioning pronoun use

Aside from the language-internal (linguistic) variables discussed above, there are also language-external (social) factors that are important to address. For the multivariate analysis in

Rbrul, one social factor was found to be a statistically significant predictor of SP variation: the age of the speaker. None of the other seven additional social variables included—*Gender, Length of Residency, Age of Arrival, English Proficiency, Preferred Media Language, Socioeconomic Status,* and *Education*—showed a significant effect on pronoun variation. Table 5.11 below shows the constraint hierarchy for *Age*.

Table 5.11: Constraint hierarchy for *Age*

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

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

Similar to the Mexico City dialect analyzed by Lastra and Martín Butragueño (2015), this could indicate a case of a change in progress toward lower overt pronoun usage in Roswell Mexican Spanish. From the perspective of apparent time (as opposed to real time), "generational differences are compared at a single point and are used to make inferences about how a change

may have taken place in the (recent) past. Age differences are assumed to be temporal analogues, reflecting historical stages in the progress of the change" (Tagliamonte 2012:43). From this view, increases or decreases in frequencies of a given linguistic feature according to age can be understood as a change in progress (Sankoff 2006). Thus, the statistically significant difference in pronoun rate between older and younger speakers in the current analysis can be taken as evidence of a change in progress toward the use of fewer overt SPs overall.

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

Aside from comparing Spanish in Roswell with other Mexican dialects, another useful comparison can be made by examining the similarities/differences in pronoun rates among speakers within the Roswell community based on their status as established immigrants vs. recent arrivals to the U.S., younger vs. older arrivals, and more proficient vs. less proficient in English.

Although the factors of *LOR, AOA,* and *English Proficiency* were not selected as significant, it can be very useful and informative to explore the general trends of these factors, particularly for shedding light on the issue of language contact. Table 5.12 below shows the overt SP rate for these different groups of speakers.

Table 5.12: Pronoun rates according to *LOR*, *AOA*, and *English proficiency*

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

Overall, when looking at the pronoun rates among the speakers' differing lengths of residency, ages of arrival, and English proficiency, the trends contradict a contact hypothesis based on English transfer. First, with regard to *LOR*, the highest pronoun users are those that have been in the U.S. for the shortest amount of time (1-7 years) while more established immigrants have lower pronoun rates. Secondly, regarding *AOA*, the highest pronoun users are those with the oldest ages of arrival (30-56) whereas younger arrivals have lower rates. Finally, *English Proficiency* is more difficult to interpret as there is not a linear trend. Nonetheless, we see no evidence in favor of English transfer: Speakers who rated their English as 1-2 (poor) and 4 (good) have virtually identical pronoun rates. Those with self-ratings of 3 (moderate) are the

lowest pronoun users. These trends contradict what some researchers have found for U.S. Spanish varieties under contact-induced change, namely higher pronoun rates among more established immigrants, younger arrivals, and those with higher English proficiency (e.g. Lapidus & Otheguy 2005; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012). A more in-depth discussion of the potential role of language contact will be presented in Chapter 6. The following section will discuss the analysis of lexical subjects in the Roswell data.

5.5 Lexical subjects

With regard to the analysis of lexical subjects (e.g. *mi mama* 'my mom'*)*, 362 total tokens were found in the 20 interviews. After discarding those outside the variable context, 39 tokens were left for analysis. Due to this relatively small token count, a multivariate analysis with lexical subjects was not carried out. Instead, I will present a qualitative analysis of these tokens.

While one of the primary uses of lexical subjects is to introduce new referents into discourse (Silva-Corvalán 1994a), there were several cases in which a lexical subject was used in the current data where a pronoun could have also occurred, thus moving beyond the mere introduction of new referents to additional uses (see Dumont 2006). The most immediate evidence of this is seen simply by the appearance of the same lexical subject in two consecutive clauses or otherwise close distances, an observation consistent with that of Dumont (2006), as seen in the following example:

(15) *<u>mi mamá</u> siempre fue ama de casa…**mi mamá tenía** más trabajo con nosotros cuidando, somos cinco hermanos* [M41Mex]

'my mom was always a housewife...my mom had more work caring for us, there are five of us [siblings]'

Thus, the second mention of *mi mamá* 'my mom' does not introduce the referent since the referent was already introduced by the speaker in the immediately preceding clause. Despite Silva-Corvalán's (1994a:148) assertion that "if the subject is coreferential with the subject of the preceding sentence, [....] a full subject NP is not acceptable in Spanish", we do see evidence of speakers producing lexical subjects in such contexts.[62] However, in most cases in the Roswell data the previous coreferential mention was in object position, as in (16):

(16) *es lo que le explicaba a <u>mi mamá</u> porque **mi mamá dice** "Carolina, hablas como, like… spanglish"*... [F34Mex]

'it's what I was explaining to my mom because my mom says "Carolina, you speak like, like...Spanglish"

Thus, these latter environments (where the previous clause contains a coreferential mention in non-subject position) are simply more common for lexical subjects than are the former environments (where the previous coreferential mention is a subject). The use of lexical subjects when their referent's first mention was in object position could be explained by the notion of *accessibility* and *salience* (see Chapter 3); referents in object position are thought to be less accessible and less salient than those in subject position (Ariel 1994). Therefore, speakers may use lexical subjects (instead of pronouns) in these contexts to bring the referent into higher salience. Similarly, another function of lexical subjects found in the current data was to establish the referent as a topic of more than one clause (see Silva-Corvalán 1994a), as in (17):

[62] Though no precise explanation of what it means to not be "acceptable" is given by Silva-Corvalán (1994a:148) regarding the lexical subject in this context, she discusses comparative examples of subject variation/obligatoriness in the same paragraph as being not "optional" or "allowed", leading one to interpret "acceptable" in these terms. There is no mention of the words "felicitous", "odd", or similar terms to denote pragmatic unacceptability. Thus, though it seems that the meaning of "acceptable" here is perhaps more in line with "grammatical", the more accurate explanation for the use of lexical subjects in this context and in terms of natural occurring speech is that lexical subjects are less *felicitous* or generally *disfavored* after their coreferential mention in the preceding clause.

(17) *Sí, muy importante más para la comunidad latina porque… **los latinos** son ah, una cultura que se caracteriza por ser... mm espiritual,*

'Yeah, very important more for the Latino community because...Latinos are uh, a culture that is characterized as being...mm spiritual'

I: Uh-huh.

R: En México, si no... me equivoco el noventa y ocho por ciento... son católicos,

'In Mexico, if I'm not...mistaken ninety-eight percent...are Catholic'

I: Sí.

'Yeah'

R: aquí lo que me he fijado es que... muchos mexicanos o latinos que llegan católicos, a veces se van convirtiendo en otras religiones… [F34Mex]

'here what I've noticed is that...a lot of Mexicans or Latinos that arrive Catholic, sometimes they end up converting to other religions'

In this example, the subject *los latinos* is used to establish Latinos as the topic to be continued in the subsequent clauses (*son católicos, latinos que llegan católicos, se van convirtiendo*). A further illustration of lexical subjects to establish a topic is the following:

(18) *R: ...yo siento que Roswell no es pobre, Roswell es rico…y este... ha cambiado bastante… en, en los policías también* [M32Mex]

'I feel like Roswell isn't poor, Roswell is rich...and umm...it has changed quite a bit...in terms of, the police too'

I: Mhm

*R: Ahorita **los policías** ya no son tan malos…Um, que ellos hacen su trabajo*

'Now the police aren't that bad anymore...Um, they just do their job'

I: Sí.

'Yeah'

R: Tienen que hacer su trabajo… [M32Mex]

'They have to do their job…'

The speaker in (18) first mentions *los policías* 'the police' as a non-subject, and then repeats it in subject position, thereby establishing the police as his topic. He then continues to refer to the police in the following two clauses with an overt SP and then a null SP. Furthermore, in establishing such topics in the above two examples, speakers may be simultaneously making the subject referents more salient/accessible, similar to the cases discussed above.

As Silva-Corvalán (1994a) notes, speakers can also make the communicative choice to highlight the subject referent by means of an expressed subject, for example in situations that are counter to expectation. Example (19) below from the present data demonstrates this function for the lexical subject *ese señor* 'that man'. In the preceding discourse, the speaker talks about a man he knows who never went to school and is illiterate. The first mention of the man is made by a noun phrase in object position followed by several cases of anaphora with overt and null SPs. The speaker then uses a noun phrase in subject position (lexical subject) to refer to the man presumably to draw the listener's attention toward the subject referent at a point in the discourse where unexpected information is introduced: despite the man not being able to read or write, *ese señor ha salido adelante* 'that man has done well for himself':

(19) *conozco a <u>un señor</u> que nunca fue a la escuela, él nunca- él no sabe X T:12:08 no sabe ni leer ni escribir, X T:12:13 ve, símbolos más o menos pero en sí él no sabe leer ni escribir, pero es-* ***ese señor ha salido*** *a, adelante es... su vida es muy interesante, este es muy trabajador y ya ha sacado su familia adelante y todo,*

'I know a man that never went to school, he never- he doesn't know [unintelligible] he doesn't even know how to read or write, [unintelligible] he sees, symbols more or less but he doesn't know how to read nor write, but th- that man has done well for himself he's...his life is very interesting, umm he's very hardworking and now he has done well for his family and everything'

Furthermore, although appearing only once in this particular context in the Roswell data, lexical subjects are used in question/answer sequences. According to Dumont (2006), this type of usage indicates a "repetition effect between speaker and interlocutor" (p. 286):

(20) *I: ¿De dónde es <u>tu esposa</u>?*

'Where is your wife from?'

R: ***Mi esposa es*** *de, nacida en Virginia* [M27Mex]

'My wife is from, born in Virginia'

Note that in (20) the interviewee could have used an overt or null SP but produced a lexical subject. The motive for this could be due to the interviewer's previous use of the same subject. Beyond simply a "repetition effect", then, this exemplifies an inter-speaker priming effect. That is, the use of *mi esposa* 'my wife' (and not *ella* 'she'*)* by the interviewee is influenced by the interviewer's immediately preceding use of *esposa* as part of the question. Further, it may also be the case that the use of lexical subjects is influenced by priming within a single speaker's turn, as is the case with pronominal subjects (e.g. Travis 2005, 2007). In other words, a speaker's own previous use of a lexical subject may lead them to repeat the lexical subject (instead of using a

pronoun) in subsequent clauses. A quantitative analysis of a larger data set would shed some light on this issue.

In summary, lexical subjects in the current data tend to occur immediately following or at close distances from their previous coreferential lexical noun phrases; this suggests that they exhibit additional functions aside from simply introducing new referents into discourse for the first time. One primary function of such occurrences may be to heighten the salience of their referents and also to establish the subject referent as the topic of subsequent discourse. Additional uses of lexical subjects in the data involved highlighting the referent in relation to unexpected situations as well as repeating a recently mentioned lexical noun phrase (a potential priming effect). Further examination of a more robust dataset of lexical subject tokens within the variable context is necessary to uncover such patterns in a more systematic way. In particular, a multivariate analysis that quantitatively examines specific variables (e.g. Person/Number, Switch Reference, TMA, etc.) in relation to variation between lexical subjects, overt SPs, and null SPs will be especially beneficial for future research.[63]

5.6 Summary

This chapter has reported the results obtained from carrying out several types of analyses regarding subject expression among Spanish-speakers in Georgia. First, the results for the overall pronoun rate showed that Mexican Spanish in Roswell exhibits a higher percentage of overt pronouns overall when compared to monolingual Mexican varieties. Second, the results of the multivariate analysis were presented, demonstrating that factors such as person/number, switch reference, morphological ambiguity, and the speaker's age, among others, played a significant

[63] As stated above, having only 43 tokens of lexical subjects within the variable context available for analysis makes a statistical analysis infeasible for the present investigation.

role in the variable use of subject pronouns in this variety. The variable hierarchy presented showed which variables had a stronger effect on pronoun use (*person/number*, *switch reference*) and which factors exhibited a weaker effect (*polarity*, *style*). In addition, similarities and differences from other varieties of Spanish were discussed in terms of the hierarchical position and strength of these factors, including factors that were either absent or emerging in Roswell Spanish. Third, the results obtained from cross-tabulation and conditional inference tree analyses were reported and illustrated, revealing that several independent variables interacted with one another, such as *Polarity*, *Verb Class*, and *Style*, among others. Discussions of how such interactions fit or did not fit with previous research were also presented. Lastly, the findings were exemplified and discussed from the qualitative analysis of non-pronominal forms, namely lexical subjects. The focus of this analysis was mainly on pragmatic and cognitive issues related to the use of lexical subjects, such as salience/accessibility of the referent, the establishment of new topics, referent prominence in situations counter to expectation, and potential repetition or priming effects. The following chapters will discuss in greater detail the implications of the results reported above, providing answers to the research questions posed in Chapter 1, and will summarize the primary findings of the present investigation.

CHAPTER 6

DISCUSSION

In this chapter, I will discuss the main findings of the present investigation. Revisiting each of the six research questions posed at the beginning of this dissertation (Introduction), I will now offer answers that aid in explaining the subject expression behavior in Roswell Spanish, including the implications regarding the role of language contact:

6.1 Research Question #1: Overall distribution of subject expression in Roswell Spanish

What is the overall distribution of subject expression in terms of overt and null pronoun usage? How does the rate of overt SP usage compare to that of non-contact varieties of Mexican Spanish?

Out of a total of 4,649 cases of variable subject expression observed, it was found that Roswell speakers produced 1,239 overt SPs (27%) and 3,410 null SPs (73%). With regard to the overall overt SP rate of 27%, this diverges from most other monolingual Mexican varieties in that it is a notable increase in overt pronominal expression. For instance, the rate found for Mexico City was 22% (Lastra & Martin Butragueno 2015) and the rate for NYC newcomers was 19% (Shin & Otheguy 2013) (see Table 6.1 below).

Table 6.1 (replication of Table 5.2): Comparison of pronoun rates between Roswell Spanish and monolingual Mexican varieties

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

Regarding the comparison between the pronoun rate of Roswell Spanish with the broader sociolinguistic literature on non-contact Spanish, as well as in countries beyond Mexico, Table 6.2 shows that Spanish in Roswell is situated among the relatively low pronoun rates of Spain and Mexico. It also reaches toward that of New York Newcomers overall (30%), but is substantially lower than varieties in Colombia, Argentina, and certainly the Caribbean. This finding is unsurprising given that Caribbean varieties typically exhibit high rates of overt pronouns and Mexican varieties have relatively low rates.

Table 6.2: Pronoun rates of monolingual varieties beyond Mexico

Variety	**PN Rate**	**Source**
Yucatan, Mexico	16%	Michnowicz 2015
Valladolid, Mexico	19%	Solomon 1999
Mexico City	22%	Lastra & Martín Butragueño 2015
Puente Genil, Spain	24%	Ranson 1991
Roswell Mexicans	**27%**	**Current study**
Barranquilla, Colombia	36%	Orozco & Guy 2008
Buenos Aires, Argentina	36%	Barrenechea & Alonso 1977
Santiago, Dominican Republic	39%	Olloqui de Montenegro 1987
San Juan, Puerto Rico	45%	Cameron 1993

Furthermore, in comparing Roswell Mexicans to other contact varieties, Table 6.3 shows us that Roswell Spanish falls approximately in the middle of the New Jersey (24%) and Los Angeles (29%) Mexican varieties.

Table 6.3: Comparison of pronoun rate in Roswell to other contact varieties

Variety	**PN Rate**	**Source**
New Jersey Mexicans	24%	Flores Ferrán 2007
Roswell Mexicans	**27%**	**Current study**
Los Angeles Mexicans	29%[64]	Silva-Corvalán 1994a
New York born/raised (Mexican, S. American, Caribbean)	38%	Otheguy, Zentella, & Livert 2007

As highlighted in Chapter 3, there is not a consistent pattern in overt SP rate differences in terms of contact vs. non-contact varieties. In other words, in comparison to non-contact

[64] This was calculated as the average rate among all three generational groups from Silva-Corvalán (1994a:153).

varieties, not all contact varieties show an increased rate of overt SPs; some stay about the same or even show lower rates than non-contact (monolingual) varieties. I will provide a more detailed discussion of the pronoun rate in Roswell and its relation to potential contact-induced change upon answering Research Question #4 below. The above discussion of overall SP frequencies highlights that there is much dialectal variation in the distribution of overt and null SPs, both among contact as well as non-contact varieties, and also reveals where the Roswell data fit with respect to other varieties. Overall frequencies tell us in general how often speakers are using both overt and null SPs, and can serve as guides to the further analysis of the particular contexts in which the pronouns are used. They can also serve as a point of departure for assessing issues of language contact and bilingualism; however, more revealing are the specific constraints that govern such pronoun distribution, which I turn to in the following section.

6.2 Research Question #2: The influence of linguistic (language internal) factors on SP variation

How do contextual/language internal factors (e.g. grammatical person/number, tense-mood-aspect, switch reference) condition SP variation and how do such patterns compare to previous research?

As discussed above, the linguistic variables (language internal) that influence subject expression in a statistically significant way in the Roswell data were shown to be the following (from strongest to weakest effect): *person/number, switch reference, morphological ambiguity, verb class*, *polarity*, and *speech style*. In particular, overt SPs were found to be favored in the following environments: first-person singular and third-person singular verbs; switch in subject referent; morphologically ambiguous verbs; mental process verbs; affirmative polarity

environments; and conversation-like speech style. In comparing these results with previous sociolinguistic studies, they are, generally speaking, consistent with what other researchers have found. The *Person/Number* of the verb is consistently the strongest predictor of SP variation in all varieties of Spanish previously studied (e.g. Bayley & Pease-Alvarez 1996; Flores-Ferrán 2002; Otheguy, Zentella, & Livert 2007; Prada Pérez 2009; Carvalho & Child 2011; Abreu 2012; Alfaraz 2015; Orozco 2015). The comparatively high use of overt SPs with first-person singular verbs, for instance, has been attributed to the egocentric nature of discourse (Posio 2011) as well as the pragmatic need for the speaker to remain overtly present in the conversation (Morales 1986). Regarding *Switch Reference*, this is also a variable with a very strong effect cross-dialectally (e.g. Silva-Corvalán 1994a; Bayley & Pease-Alvarez 1997; Travis 2005; Prada Pérez 2009; Torres Cacoullos & Travis 2010; Carvalho & Child 2011; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015). That overt SPs are more likely to be produced when there is a switch in the subject referent has been explained in terms of referential tracking, salience, and accessibility of the antecedent (Shin & Otheguy 2009; Givón 1983; Ariel 1994). When a switch in referent occurs, the antecedent tends to be less salient/accessible to the listener, thus promoting more overt SPs to facilitate referent identification. The third strongest variable, *Morphological Ambiguity*, also shows patterns consistent with previous research (e.g. Hochberg 1986; Travis 2007; Prada Pérez 2009; Erker & Guy 2012; Lastra & Martín Butragueño 2015; Michnowicz 2015) with the overt SP percentage with ambiguous verbs nearly doubling when compared to unambiguous verbs. As discussed in Chapter 5, this finding appears to corroborate Hochberg's (1986) Functional Hypothesis which predicts a compensatory use of overt SPs given the lack of overt person marking with ambiguous forms. A more in-depth contextual analysis is warranted, however, to explore whether there is *contextual* ambiguity in these cases or simply

morphological ambiguity (Ranson 1991). With respect to the *Verb Class* effect, it has been argued that mental process (cognitive) verbs tend to appear with overt SPs due to the fact that these verbs are often used to express point of view and an implied contrast (Silva-Corvalan 1994a). The favoring of overt SPs with this verb class has been observed repeatedly (e.g. Silva-Corvalán 1994a; Travis 2007; Torres Cacoullos & Travis 2010, 2011; Otheguy & Zentella 2012; Orozco 2015). Furthermore, the factor of *Polarity*, while less studied, corroborates the findings of the researchers who have analyzed it (Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016). Finally, the finding that conversation-like *Speech Style*, (as opposed to interview-like/less casual speech) favors overt SPs is consonant with the few studies that have examined it (Ávila-Jiménez 1996; Lastra & Martín Butragueño 2015). I proposed in Chapter 5 that perhaps when speakers are less self-aware of their speech in a more casual/conversation-like style, they may pay less attention to grammar norms that prescribe the minimal use of overt pronouns (e.g. Butt & Benjamin 2004). Three additional linguistic variables that were included in the present analysis but that were not found to be significant were *Specificity*, *TMA*, and *Verbal Mood*. One of these nonsignificant factors, *TMA*, deserves further discussion as the findings relate to the Roswell data since this was a rather unexpected result.

In the previous chapter it was demonstrated that the *TMA* constraint disappeared for Roswell Mexicans (it was not a statistically significant predictor of SP variation). Specifically, we observed that speakers showed a narrowing of the well-documented distinction in overt SP use between the imperfect and preterit aspects (a higher use with the former and a lower use with the latter). Such narrowing manifested as an increased use of overt SPs with preterit verbs, making the pronoun rate essentially the same in both contexts. When comparing to previous studies of Mexican Spanish, we did not see a difference in pronoun behavior in imperfect

contexts with regard to overt SP rate. The main distinction was seen in preterit contexts, in which we observed a substantial pronoun increase compared to rates in previous studies (e.g. 16% in Mexico City vs. 29% in current study).

A natural question that arises, then, is why speakers in the current data would use more pronouns with the preterit? Based on the framework of a weakened/loss of sensitivity to the *switch reference* constraint in previous SP studies of contact varieties of Spanish (e.g. Flores-Ferrán 2002, 2004; Shin & Otheguy 2009; Michnowicz 2015), it is possible that speakers in the current data are exhibiting a loss of sensitivity to *TMA* as a constraint regulating the variable use of overt and null SPs. In particular, parallel to the repeated finding in the contact literature of an increased use of overt SPs in *same* reference (coreferential) contexts (Flores-Ferrán 2002, 2004; Shin & Otheguy 2009; Otheguy & Zentella 2012), which narrows the distinction relative to switch reference (i.e. disjoint reference) contexts, Roswell speakers show an increased use of pronouns in *preterit* contexts compared with imperfect environments, neutralizing the overt/null SP distinction for imperfect vs. preterit contexts. Thus, Roswell bilinguals are experiencing a loss of sensitivity to *TMA* as part of their variable grammar of subject expression. That is, they are no longer guided by differences in TMAs for their variable pronoun use. Instead, they are treating preterit verbs no different from imperfect verbs, at least with regard to overt SP rate. Such similar behavior in pronoun frequencies in these two contexts lends little support to Silva-Corvalan's (2001) proposal of the pragmatic effects of SPs, namely that speakers tend to give a greater focus to the subject (and less to the action) when using the imperfect (thus more overt SPs), and that they tend to give less focus to the subject (and more to the action) when using the preterit (thus fewer overt SPs). Insofar as such pragmatic functions are in effect in Mexico City Spanish (and other dialects), the data from Roswell Spanish suggests that these speakers exhibit

a pronoun grammar that lacks these TMA constraints on their SP use. The factor of TMA, then, is neutralized and weaker in this variety.

Another question that arises in terms of the use of overt SPs with the preterit is in what other environments, simultaneously with the preterit, exhibit this rise in pronoun use? In other words, what other linguistic variables are at play that might further explain the increased pronoun usage with the preterit? After testing for interactions with *TMA*, it was found that the high use of overt SPs with preterits also coincided with first-person plural verbs; that is, first-person plural verbs in the preterit aspect are significantly more likely to appear with overt pronouns than first-person plurals in all other TMAs ($p < .05$). This result provides an answer to our question regarding the contexts in which we are seeing a rise of overt SPs with the preterit. A specific motivation for such pattern is unclear at this time; however, I suspect it may have to do with the notion of contrast and contrastive contexts (Chapter 4). At least with regard to the current dataset, I have observed that in contexts of preterit first-person plural verbs, the speaker also tends to be expressing a contrast, a context in which overt SPs are favored, as discussed in Chapter 4. Future studies will need to both corroborate this finding as well as explore potential functional motivations for the differential use of overt pronouns in such contexts. At any rate, from a broader perspective, there is clearly a loss of sensitivity to *TMA* as a governing constraint on SP variation in the present data.

Finally, it is worth exploring the question, in a broader sense, of why a *TMA* effect was not found in the Roswell data. One hypothesis is that while *TMA* was not significant as a *main* effect, it could be interacting with other variables such as *Switch Reference*, as was found by Shin (2014) in NYC. This would mean that *TMA* still functions as a conditioning factor of SP variation, but to a lesser extent and in a more indirect manner. In fact, an additional interaction

analysis was carried out to test the interaction between *TMA* and *Switch Reference*, revealing a significant interaction ($p < 0.05$). However, the same unexpected result for *TMA* remains. Speakers do not distinguish between imperfect and preterit verbs in terms of their overt SP frequencies, regardless of whether or not there is a switch in subject referent. Concerning the interaction effect in particular, *TMA* emerges as significant in switch reference contexts, but not in same reference environments, as seen in Figure 6.1.

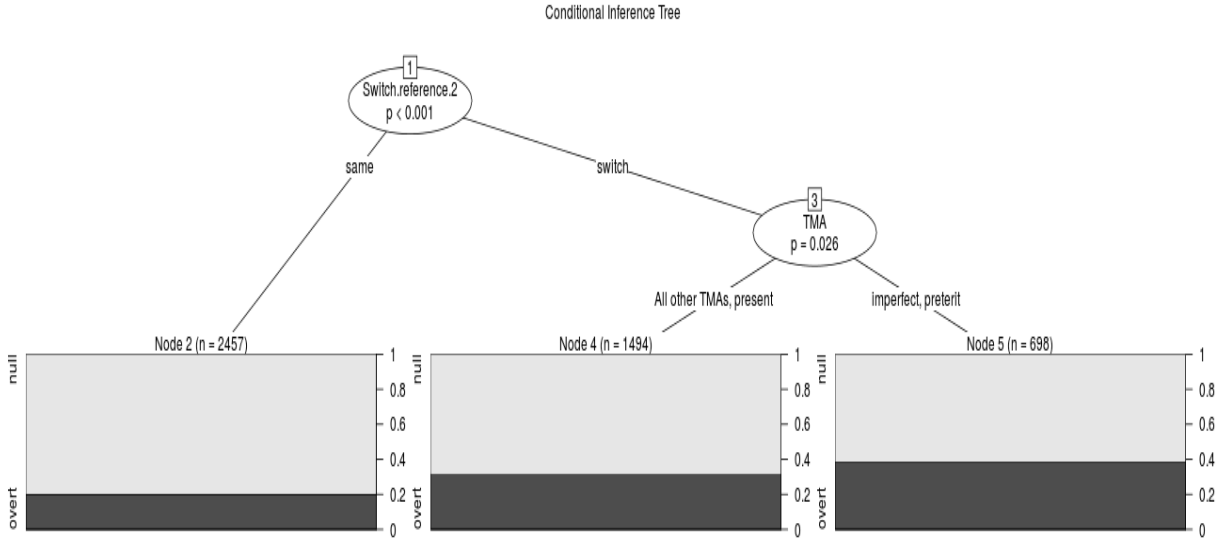

Figure 6.1 Conditional inference tree for TMA and Switch Reference

This *general* effect is consistent with previous studies that have tested this interaction (Cameron 1994; Shin 2014). However, the *specific* effect is quite different. That is, the distinction in pronoun rates being made in switch reference contexts is not between the imperfect and the preterit, but between past tense forms in general (imperfect/preterit both with high pronoun usage) and all other TMAs (lower pronoun usage) (see also Table 6.4 for comparison).

Table 6.4: Rate of overt SPE according to *TMA* and *Switch Reference*

	Imperfect	Preterit	Present	All other TMAs
Same	23%	22%	18%	20%
Switch	**38%**	**40%**	32%	35%

Thus, we are seeing an extension and generalization of the *TMA* effect to now include preterit contexts. At the same time, speakers are holding on to their relatively high overt SP use with imperfect (38%). For example, we are not seeing a decrease in the use of SPs in imperfect environments. Therefore, there could still be a disambiguating strategy at play, especially since the factor of *morphological ambiguity* was significant, as discussed in the previous section.[65] In sum, while Roswell Mexicans are not guided by *TMA* as a whole (absence of a main effect), they are guided by it in combination with other environments, namely when there is a switch in subject referent. This effect is unique in that speakers are simplifying the *TMA* constraint in these contexts by increasing their usage of overt SPs with the preterit. As for the motivations for such a distinct effect, this pattern could be an effect of bilingualism and/or dialect contact between Mexican varieties, but the particular influence remains to be seen.

6.2.1 Research Question #2a: Interactions among the linguistic variables

What do the interactions among the linguistic variables reveal about the Roswell data? In other words, what can the non-independent effects on SP variation show us concerning variable subject expression in the current data?

[65] A subsequent interaction analysis with *Person/Number* revealed that this *TMA* effect only applies to first-person singular and third-person singular forms ($p < 0.01$), which are morphologically indistinguishable in the imperfect, but not in the preterit. Further study of this interaction will be useful for future research.

In examining interactions between the independent linguistic variables, several factors demonstrated a non-independent effect on subject expression. Particularly, *polarity* interacted with several variables, namely *verb class, person/number,* and *switch reference*. Regarding *verb class*, the polarity effect (affirmatives favoring overt SPs) is only operative for mental process verbs and not for the other verb classes. It was proposed, though tentatively, that the particular verbal lexeme *creo* '(I) think' may be the driving force behind this interaction effect, a finding consistent with the analysis of Travis and Torres Cacoullos (2012). These researchers examined subject expression with the use of *creo* in particular and observed a strong disfavoring of overt SPs in negative polarity environments (i.e. ∅ *no creo*). This general interaction in the current study between *Polarity* and *Verb Class* also indicates that the verb class effect (mental verbs favoring overt SPs) only applies to negative polarity contexts.

With regard to *person/number*, the interaction shows the following patterns: the polarity effect only applies to first-person and third-person singular verbs. The finding concerning the significance of first-person singular verbs in particular is consistent with Geeslin and Gudmestad's (2016) analysis, in which they found that polarity was operative only for first-person singular and not for second-person singular forms among native speakers. Moreover, the effect of *person/number* is stronger in affirmative contexts than in negative contexts.

For polarity's interaction with *switch reference*, the polarity effect is only operative in switch reference (noncoreferential) environments (including both switch and partial switches). This finding contradicts the hypothesis proposed by Lastra and Martín Butragueño (2015) that negation favors null SPs due to its frequent co-occurrence with coreferential (i.e. same reference) contexts in their Mexico City data. That is, since the polarity effect in the current data does not apply to coreferential contexts, there is no relation between coreference and negative polarity

such that they jointly constrain SP variation. If this were the case, we would have expected to see significant differences in overt SP rates between negative/coreferential contexts and affirmative/coreferential contexts, which we did not observe.

Additionally, a significant interaction was found between the variables of speech *style* and *person/number* such that the effect of style (overt SPs favored with conversation-like [more casual] speech) only applied to first-person singular environments. This finding reveals that the speech style effect is not operative for other persons/numbers, an observation not yet attested in the literature since previous studies that have analyzed *style* have either not taken into account other persons/numbers aside from first-person singular (e.g. Blanco Canales 1999; Travis 2007) or have not tested their possible interactions with style (e.g. Lastra & Martín Butragueño 2015). It was suggested in the previous chapter that this particular finding could be due to the presence of more first-person singular mental process verbs (*creo* and *pienso*, in particular, which favor overt SPs) that coincided with more casual/conversation-like speech. In sum, we see that the variables of *polarity, verb class, person/number, switch reference,* and *style* do not exhibit independent effects on SP variation across the dataset, but rather intersect with each other to constrain subject expression. In other words, the effect of one variable (e.g. polarity) *depends on* the effect of another (e.g. verb class). These findings reveal a more nuanced picture of the Roswell data in that they reflect a more complex variable pronoun grammar of Roswell speakers that was previously obscured by simply examining the main effects of each of the linguistic variables (Research Question #2).

6.3 Research Question #3: The influence of social (language external) factors on SP variation

Do language external factors such as English proficiency, age, gender, length of residency in the U.S., and age of arrival to the U.S. play a role in how the speakers use SPs?

It was found that the only language external (social) factor that played a role in the usage of SPs for Roswell speakers was their age. Specifically, younger and middle-aged speakers slightly disfavored the use of overt SPs, producing them 24% of the time, while older speakers (50+) showed a favoring of overt SPs with 34% production. This finding corroborates previous research that has reported a general distinction between younger (fewer overt SPs) and older speakers (more overt SPs) (Orozco & Guy 2008; Carvalho & Child 2011; Alfaraz 2015; Lastra & Martín Butragueño 2015; Orozco 2015, 2016). This particular observation in the current data is argued to provide evidence of a potential change in progress toward lower overt pronoun use for Roswell Mexicans.

Furthermore, Drawing from Labov (2001:75), Lastra and Martín Butragueño (2015) propose that this pattern for age and SP usage could be a case of *retrograde movement*, meaning that, in terms of SPs, the change from higher to lower pronoun rates overall moves against the trend of the observed diachronic change in Romance toward *higher* pronoun rates (Lastra & Martín Butragueño 2015:49-50). It appears that this retrograde movement is also in effect among Mexicans in Roswell. Although it was noted that *overall* the Roswell Mexican variety is one that comparatively comprises more overt pronouns (Chapter 5) in relation to non-contact Mexican dialects, when we look more closely by examining generational (age) differences, there appears to be a shift toward fewer overt pronouns since the youngest speakers are more likely to omit them. This finding further weakens the English contact hypothesis (Silva-Corvalán 1994a;

Otheguy & Zentella 2012) predicting that there will be a shift toward more overt SPs in Spanish-English contact situations given the (near) obligatory status of overt SPs in English. If there is a change in progress in the sense that as time progresses for the Roswell Mexican community, fewer pronouns are used, we are left with little evidence of transfer of the English pronoun structure to the Spanish grammar. This does not negate, however, the possibility that bilingualism could be playing a role in influencing pronoun usage in Roswell, an issue addressed in the following section.

Moreover, the question arises of, given that similar age patterns were observed in Mexico City (Lastra & Martín Butragueño 2015), whether this effect is specific to Mexico City speakers in the current data. An additional analysis of *Age* was carried out, solely for the eight speakers that are from Mexico City, to determine if the age of these speakers was also a significant factor conditioning pronoun use. The results, which are shown in Table 6.5 below, revealed that *Age* was in fact significant with this particular subset of speakers ($p = 0.0205$). Generally speaking, we see that the age effect is very similar in that younger speakers produce fewer overt SPs than older speakers (compare with Table 6.6).

Table 6.5: Constraint hierarchy for *Age* (Mexico City speakers only [N=8]

Factor Group	**Factor Weight**	**% Overt**	**Total N tokens**	***p*-value**
Age				0.0205
35-49	.55	27%	562	
50+	.49	26%	551	
20-34	.46	22%	921	
RANGE 9				

Table 6.6: Constraint hierarchy for *Age* (all speakers [N=20])

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

Nonetheless, numerous differences are apparent and are important to point out. First, the overall effect is less dramatic than it was when all 20 speakers included. In particular, whereas the initial analysis that included non-Mexico City speakers demonstrated a 10% difference between older (34%) and younger (24%) speakers, the Mexico City data indicate only a 5% difference (27% vs. 22%). It could be the case, then, that exposure to other Mexican dialects has strengthened the age effect, with a wider distinction between older and younger speakers' pronoun use. More specifically, it is the older speakers (50+) that appear to drive this change, as evidenced by the increase in pronoun rate from 26% with exclusively Mexico City speakers to 34% when all speakers are included. We do not observe an increase of this magnitude for younger speakers (20-34), as the difference is marginal (22% with Mexico City speakers vs. 24% with all speakers).

The second primary difference is that there was a distinction made in pronoun rate between speakers who were 50 and older and those that were 35-49 years old when all 20 speakers were included. This distinction, however, is neutralized when only considering Mexico City speakers, as shown in Table 6.5 by the virtually identical overt pronoun rates for these two

age categories (27% for 50+ and 26% for 35-49). This finding diverges from the monolingual Mexico City data examined by Lastra and Martín Butragueño (2015:49) since these researchers found a significant difference between similar age categories (30% for 55+ vs. 20% for ages 35-54). In this regard, the current data with regard to *Age* actually look *less* like the monolingual Mexico City data when analyzing exclusively Mexico City speakers. Simultaneously (and counterintuitively), the Roswell data look *more* like the monolingual Mexico City data when the non-Mexico City speakers are included, suggesting that this age effect is not unique to Mexico City Spanish.

Aside from the analysis of *Age*, seven additional social variables were explored in the present study; however, none of them demonstrated a significant role in the speakers' variable pronoun use. These included *Gender, Length of Residency, Age of Arrival, Socioeconomic Status, Level of Education, Preferred Media Language*, and *English Proficiency*.

6.4 Research Question #4: The role of contact-induced change

What role, if any, does English contact play in the conditioning of SP rates in Roswell Spanish? That is, do the overall rate of overt SPs or the patterning of constraints on subject expression in comparison with non-contact Spanish reflect externally-motivated change due to English or do they reflect a lack of such change?

To address this question, it will be necessary to discuss both a comparison of pronoun rates and a comparison of the constraint patterning between Roswell Mexican Spanish and a Mexican Spanish variety that is not in contact with English. First, in terms of overall pronoun rates, we did observe a higher rate for Roswell Mexicans (27%) than what was reported for Mexico City Spanish (22%, Lastra & Martín Butragueño 2015). Is this sufficient evidence to

conclude that Mexicans in Roswell have been influenced by English in terms of pronoun rate? From the perspective of this investigation, an increase in overall pronoun rate alone is not sufficient to make a determination regarding contact-induced change. As addressed in Chapter 3, using overall frequencies to assess language contact has been called into question by some researchers (e.g. Poplack & Levey 2010; Torres Cacoullos & Travis 2010), and a similar perspective is adopted here. Specifically, it has been widely observed that frequencies of overt SPs differ substantially across dialects of Spanish (e.g. Poplack & Levey 2010; Torres Cacoullos & Travis 2010; Otheguy & Zentella 2012; Alfaraz 2015; Orozco 2015), and it cannot be ruled out that the increased usage observed in the current data is not simply a reflection of regional variation.

Additionally, and perhaps more importantly, there have been cases documented of Spanish in contact with pro-drop languages (as opposed to non-pro-drop languages like English) in which there was also an increase in overt SP usage (e.g. Prada Pérez 2015; Michnowicz 2015). For example, this was observed by Michnowicz (2015) in his analysis of Spanish in contact with Yucatec Maya. For this study, Michnowicz reported that Spanish monolinguals used overt SPs 16% of the time while Spanish-Maya bilinguals showed a pronoun rate of 24%, a finding that was interpreted as a simplification strategy for bilinguals whereby they lighten the cognitive load associated by simplifying pragmatic constraints. Studies such as these cast doubt on the idea that, in contexts of Spanish-English contact, transfer has taken place from English to Spanish regarding the overt pronoun structure. Further, they raise the question of whether bilinguals experience increases in pronoun rates regardless of the language with which they are in contact. For example, it could be the case that a process of simplification is taking place and that the influence in increased overt frequencies (as well as different constraint patterning) is due to

bilingualism itself rather than the contact language (e.g. English) in particular (Sorace 2012; Prada Pérez 2015; Michnowicz 2015). I will return to the issue of simplification in the latter part of this section.

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

Nevertheless, once we begin looking beyond mere pronoun rates and focus on particular constraints on SP variation, we see that bilingualism in general seems to be playing a role in the pronoun behavior of Roswell speakers. As highlighted above, in comparing variable hierarchies between Roswell Mexican Spanish and Mexico City Spanish, speakers in Roswell demonstrated several distinctions regarding the strength of some of the constraints, which was evidenced by relative position in the hierarchy and by their range values. First, the variables of *morphological ambiguity*, *person/number*, and *verb class* were stronger in Roswell. In fact, *verb class* was an emergent constraint not found to be significant at all in Mexico City. Regarding *morphological ambiguity*, the range value for the present results was 20 (ranking third overall) compared to 13 (ranking sixth overall) in the results reported by Lastra and Martín Butragueño (2015) for Mexico City. For *person/number*, a range comparison relative to *switch reference* in both varieties revealed that the gap was much wider in Roswell (range value difference of 28) than in

Mexico City (range value difference of 1), showing that *person/number* is a much more important factor than *switch reference* for Roswell speakers whereas the predictive power of the two variables is about the same for speakers in Mexico City.

Secondly, and on the contrary, a weakening/loss of certain constraints was observed in the Roswell data. Returning to the comparison of *person/number* and *switch reference* in the preceding paragraph, while *person/number* was more important in Roswell, *switch reference* was substantially weaker, a finding not seen in Mexico City since the two variables were on par with each other concerning their ranges (i.e. speakers attributed about the same importance to *switch reference* as they did *person/number*). Such a weakening in *switch reference* has been repeatedly observed in other contact situations (e.g. Flores-Ferrán 2002; Shin & Otheguy 2009; Otheguy & Zentella 2012; Michnowicz 2015). In fact, Otheguy and Zentella (2012) observed the same differential relationship between *switch reference* and *person/number* as was found in the current study, namely that the former showed a weaker effect and the latter, a stronger effect, particularly when comparing New York City born/raised bilinguals to newly arrived immigrants to the city. Moreover, Michnowicz (2015:111) found that Spanish-Maya bilinguals did not rely on *switch reference* as much as Spanish monolinguals did, concluding that "Maya-speakers are not attributing the same importance to co-reference [switch reference] as do Spanish-speakers, suggesting that these bilinguals and monolinguals may not share all of the pragmatic constraints on SPE [subject pronoun expression]."

Another constraint that exhibited a weaker effect in Roswell than in Mexico City was *polarity*. This was shown by the range value of 10 in the current results, which is more than twice as weak as the strength observed in Mexico City (range = 23). Finally, a loss of some

constraints was observed: neither *TMA* nor *verbal mood* was significant in the Roswell data, but both were significant for Mexico City speakers in Lastra and Martín Butragueño's (2015) study.

What do these divergences suggest in terms of the potential linguistic processes at work for bilinguals in the current data? As reviewed in Chapter 3, certain processes such as *simplification*--the weakening of discourse or morphosyntactic constraints to lighten the cognitive load (Silva-Corvalán 1994b; Sorace 2004)--as well as *complexification* (the emergence of constraints) (Shin 2014) come into play as an influence of language contact and bilingualism. Regarding the present analysis, then, the abovementioned weakening and loss of constraints (*switch reference, polarity, TMA, verbal mood*) compared to Mexico City Spanish demonstrates evidence of *simplification* of the variable pronoun grammar for Roswell Mexicans. In agreement with other researchers, I maintain that this simplification process is influenced by bilingualism in general (Michnowicz 2015). Returning to the proposal above that increased pronoun rates could generally be due to bilingualism and not necessarily influenced by the specific contact language (i.e. English), this explanation can also be applied to constraint differences (Sorace 2012; Michnowicz 2015). For instance, Michnowicz (2015:113) posited that the decreased sensitivity to *switch reference* among Maya-Spanish bilinguals (summarized above) could constitute evidence of contact-induced simplification, particularly in the sense of an influence of bilingualism rather than a direct influence of the Maya language specifically. In the same way, I argue that the primary mechanism of change for the simplification of *switch reference, polarity, TMA,* and *verbal mood* observed in the Roswell data is that of Spanish-English bilingualism, and not necessarily English in particular. In terms of the processing costs of using both languages, the Roswell speakers may be lightening the cognitive load associated with bilingualism (Silva-Corvalán 1994b; Sorace 2004) of these discourse-pragmatic constraints on SP variation.

Simplification understood in this sense follows the predictions of Sorace (2012), namely that bilinguals will experience a weakening of pragmatic constraints (i.e, *Switch Reference*) due to the processing costs of bilingualism and that this occurs regardless of the particular language pair involved.

In tandem with the process of simplification, we can also see evidence of a process of *complexification*. Specifically, the emergence of the *verb class* constraint in Roswell (not significant in Mexico City) as well as the strengthening of *morphological ambiguity* point toward a more complex variable grammar (see Shin 2014). Shin (2014) explains what constitutes a more complex variable grammar upon making a contrast with the notion of a simpler grammar (i.e. simplification) in the following way: "[t]he loss of a linguistic factor that constrains linguistic choice is a type of simplification, while the *emergence of a new factor* is a type of complexification" (p. 305, my emphasis). In the present study, I extend *complexification* to not only refer to the *emergence* of a new factor, but also to the strengthening (increased sensitivity) of a given factor, as is the case with *Morphological Ambiguity* in the Roswell data. The observation of this simultaneous process of simplification and complexification in the present study lends support to Shin's (2014) hypothesis that the increased sensitivity to some factors occurs as a way of compensating for the decreased sensitivity (simplification) to other factors. In the current analysis, this type of compensation is extended to refer to the strengthening of certain constraints already operative in the variable grammar of monolinguals, and not solely to the new constraints emerging in the variable grammar that were not present in monolingual varieties.

Overall, although a *direct* effect of English does not seem to be the case, bilingualism in general seems to be playing a role in shaping the variable system of subject expression of speakers in Roswell, showing an *indirect* effect of English (Silva-Corvalán 1994b). The above

discussion of differential patterning does not, however, intend to suggest that there is no continuity in the variable pronoun system between monolingual Mexico City speakers and bilinguals in Roswell. As presented in the previous chapter, it was found that the first two variables in the hierarchy were the same and in the same order according to importance: *Person/number* ranked 1st and *switch reference* ranked 2nd in both studies; *style* was ranked last, showing the least amount of importance for both varieties; and the two varieties shared the lack of influence of *specificity* on SP variation, shown by its absence in the variable hierarchies. Further, upon comparing the constraint hierarchies between the two varieties (the direction of effect and the constraint ordering within each variable), it was demonstrated that Roswell and Mexico City speakers share the same patterns for each significant variable in terms of which contexts favor the use of overt SPs, and regarding the hierarchical ordering of these contexts, with the exception of *person/number*. Therefore, taking all the above comparisons into consideration, we have a case of both linguistic change and linguistic continuity in the Mexican Spanish of Roswell, Georgia.

6.5 Research Question #5: Roswell Spanish as an emergent variety and comparisons to more stable varieties

What do the rates and/or constraints of subject expression tell us about the emergent nature of Roswell Spanish? How does subject expression in this variety compare to that of more stable varieties?

According to Moreno-Fernández (2009), an *emergent* variety is one that adopts characteristics, primarily in contexts of language or dialect contact, that are not part of the original linguistic configuration of the non-contact variety:

> [c]uando una variedad adopta unas características que no le son propias por la incidencia de causas externas, entre las cuales los contactos con otras lenguas o variedades dialectales son las más potentes, llevando a una configuración lingüística progresivamente distante de la original, puede hablarse de variedades o dialectos *emergentes*." (p. 217, emphasis mine)[66]

In light of the above discussion (Research Question #5) regarding *simplification* and *complexification* processes due to bilingualism, the differential constraint patterns on subject expression seen in the Spanish of Roswell suggest that this is an emergent variety of Spanish in the U.S. This is the case since what we see occurring is that Mexican Spanish speakers in Roswell are developing their own variable grammar for subject expression that is distinct from monolingual Mexican Spanish in terms of the emergence/strengthening of some constraints and weakening/loss of others. As argued above, such differential constraint patterns appear to be an influence of bilingualism.

With regard to the comparison of SP variation in the current data to that of more stable U.S. Spanish varieties, the data show us that Roswell Spanish is beginning to move toward more stable contact varieties, particularly in terms of the shifts in constraints (e.g. Silva-Corvalán 1994a; Flores Ferrán 2002; Shin & Otheguy 2009; Otheguy & Zentella 2012; Shin 2014). As highlighted above, it is well documented that U.S. born Spanish speakers have experienced a weakening/loss of constraints on subject expression. These include the weakening of *switch reference* in New York (Flores Ferrán 2002, Shin & Otheguy 2009; Otheguy & Zentella 2012)

[66] "When a variety adopts characteristics that are not its own due to the influence of external causes, among which contact with other languages or dialectal varieties are the most powerful, leading to a progressively distant linguistic configuration from the original, one can speak of emergent varieties or dialects."

and Los Angeles (Silva-Corvalán 1994a) as well as the loss of *Morphological Ambiguity* in Los Angeles (Silva-Corvalán 1994a). Simultaneously, an emergence of constraints such as *TMA* among New York bilinguals has been reported (Shin 2014). Such shifts in constraints between bilinguals and monolinguals are representative of the nature of long-standing U.S. Spanish varieties, and the findings of the present investigation demonstrate that such differential constraints also characterize Roswell Mexican Spanish. Thus, despite being only in the initial stages of language contact (first generation immigrants) and not yet being a well-established variety (i.e. second or third generation immigrant Spanish), we can already observe signs of an emergent dialect in Roswell with regard to variable subject pronoun expression in that we are seeing evidence of processes of simplification and complexification upon comparing Roswell Spanish to monolingual Spanish. Now that general answers have been offered to this research question, it is essential to expand upon the issue of emergent varieties.

Reflecting on the quotation above from Moreno-Fernández (2009) concerning the notion of an *emergent* variety, we see evidence that the variety of Mexican immigrant Spanish spoken in Roswell is adopting characteristics that are not its own, meaning that it has a unique and distant linguistic configuration from the original, monolingual Mexican Spanish. I argue that these distinct characteristics are due to the influence of contact with another language (English), and more generally, to bilingualism itself. Furthermore, aside from being distinct from that of monolingual Mexican Spanish speakers, Spanish in Roswell is also distinct from less established U.S. varieties, such as those of newly arrived immigrants, or what Otheguy and Zentella (2012) would call *newcomers*. These researchers, among others, commonly approximate the pronoun usage of newcomers with monolingual speakers, given that the former have spent very minimal time in the U.S. and that their Spanish reflects that of their home varieties (Otheguy, Zentella, &

Livert 2007; Otheguy & Zentella 2012; Shin & Erker 2015). For instance, in terms of rates for SP use, if we compare Roswell speakers to first-generation Mexican speakers in NYC reported by Shin & Erker (2015), we see a 6% increase in overt pronoun rate (21% in NYC vs. 27% in current study). A similar change in rate is observed when compared to Otheguy and Zentella's (2012) finding for Mexican newcomers in NYC (19% vs. 27% in current study). We also see a neutralization in the distinction of preterit and imperfect contexts when compared to first-generation Mexican speakers in NYC (Shin & Erker 2015). Such differences, especially those regarding *TMA,* can be taken as evidence that the Roswell Mexican variety is adopting a distant linguistic configuration from the original, and therefore constitutes an emergent variety.

Additionally, as addressed in Research Question #4 above (and first presented in Chapter 5), in comparing variable hierarchies by means of range comparisons and order in hierarchy, the evidence of the linguistic processes of simplification and complexification in the current data when compared to non-contact Mexican Spanish suggests that Roswell Mexican Spanish is an emergent variety. The differential patterns observed for the factors of *TMA, Switch Reference, Verbal Mood,* and *Polarity* on one hand (simplification), and *Morphological Ambiguity* and *Verb Class* on the other (complexification) show us that the grammar of Roswell Mexicans displays a distant linguistic configuration (though with several similarities) from the original or benchmark variety. Such configuration in constraints is distinct from monolingual Mexican Spanish varieties, thus showing that unique characteristics are emerging that are not present in non-contact (or *newcomer*) varieties.

Moreover, previous research shows that more long-standing U.S. Spanish varieties such as those of California and New York typically exhibit different configurations among U.S-born speakers (second or third immigrant generation) (e.g. Silva-Corvalán 1994a; Flores-Ferrán 2004;

Shin & Otheguy 2009; Otheguy & Zentella 2012). The current data—foreign born, first generation immigrants—show similar characteristics to these more established varieties, suggesting that at relatively initial stages of language contact speakers are beginning to develop linguistic characteristics that differ from their homeland varieties. At the same time, such characteristics are beginning to fall in line with some of the contact-induced changes observed among second and third generation speakers in other U.S. regions.

In sum, we find evidence in the patterns of SP usage of a newly emerging variety of Mexican Spanish as a result of contact and bilingualism, and thus are justified in characterizing Roswell Mexican Spanish as an *emergent* variety or *new* variety (in the sense of Lamanna 2012). In fact, one could describe Spanish in Roswell as an *incipient-stage* emergent dialect given that the observed shifts in constraints, while evident, are not to the same extent as more long-standing Latino communities in the U.S. For instance, although there were changes observed in the variable hierarchy (ranking and range values), there were essentially no changes in constraint ranking within each variable when compared to Mexico City, as discussed above in Section 6.4. This would constitute a more substantial change in the variable grammar and is one that would be expected in more established contact varieties (e.g. Silva-Corvalán 1994a; Otheguy & Zentella 2012). Thus, it appears that the effects of bilingualism and language contact are beginning to surface in Roswell but are not (yet) part of the more advanced stages of language shift seen, for example, in Spanish in New York or Los Angeles. The above findings reveal an insight, then, into the linguistic mechanisms at work in the initial stages of language contact (i.e. Spanish in Georgia as a site of recent emergence of Spanish-speaking populations), which is something obscured in the sociolinguistic SP research on Spanish in the U.S. that primarily examines more established varieties of Spanish. Moving forward to the final chapter of this

dissertation, I will present an overall summary of the present investigation, general conclusions, the primary contributions of the current study to the field, and avenues for future research.

CHAPTER 7

CONCLUSION

7.1 Summary and general conclusions

This dissertation has explored variable subject expression in a Southeastern U.S. Mexican immigrant community, that of Roswell, Georgia. One of the major goals of the study was to further our knowledge of the linguistic behavior of emerging bilingual speech communities, particularly with regard to incipient stages of language contact. As a city whose significant demographic shifts reflect the broader picture of recent and ever-increasing immigration from Latin America to the Southeast, it was argued that Roswell is an ideal location for a study that explores new bilingual speech communities, their linguistic contact, variation, and change.

Through a variationist sociolinguistic analysis of subject expression among 20 first-generation Mexican immigrants in Roswell and a comparison to other varieties of Spanish, I have been able to draw several conclusions. Regarding the overall overt pronoun rate, Roswell Mexicans exhibited differential behavior of pronoun usage in comparison to most non-contact varieties of Mexican Spanish, particularly an increased rate of pronoun usage. This sets Roswell Mexican Spanish apart from other Mexican varieties (Yucatan, Mexico; Valladolid, Mexico; NYC Mexican newcomers; Mexico City; Washington/Montana Mexican monolinguals), but it was argued that the influence of this change is not due to language contact and bilingualism. With regard to the variables that work together to constrain the variation in overt and null SP

usage, it was shown that six language internal variables and one language external variable played a significant role for this Spanish variety. With respect to linguistic factors, the strongest predictor of SP variation was *person/number*, followed by *switch reference, morphological ambiguity, verb class, polarity,* and *speech style.* Although explored in the present investigation as well, the variables of *specificity*, *TMA*, and *verbal mood* did not play a significant role. As for language external factors, the age of the speaker was the one variable that played a significant role in the variation of pronominal subjects, with the oldest speakers being the highest pronoun users. The additional language external variables studied in the present analysis did not show a significant effect on pronoun usage. Moreover, some of the language internal variables significantly interacted with one another to regulate SP variation, namely *polarity, verb class, person/number, switch reference,* and *speech style*. Thus, rather than exhibiting independent effects on pronoun variation, they jointly constrain SP usage by depending on each other. This finding sheds light on types of interactions not explored in previous subject expression research while at the same time confirming and corroborating findings from other researchers. Thus, the significant interactions observed in the present study reveal a more nuanced view and further our understanding of how the conditioning factors of SP variation work together while also lending support to previous scholarship.

Regarding the potential effects of English contact and bilingualism, I conclude that in terms of overall pronoun rate, there is not sufficient evidence to assume that the behavior of SPs in Roswell Spanish is attributable to contact with English. Despite being higher overt pronoun users in general, no direct correlations were found between overt frequencies and factors such as length of residency, age of arrival, English proficiency, or preferred media language, factors which we would predict to influence pronoun use in a case of contact-induced change (Lapidus

& Otheguy 2005; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella 2012; Abreu 2012; Shin & Montes-Alcalá 2014). Additionally, it is unclear to what extent increases in overall pronoun rates would be an indication of transfer of the overt pronoun structure from English to Spanish as there is evidence of such rate increases in Spanish varieties in contact with languages other than English, particularly other pro-drop languages such as Yucatec Maya (Michnowicz 2015). Thus, some argue that it is bilingualism itself, and not influence of the contact language specifically, that serves as the point of influence (e.g. Sorace 2012; de Prada Pérez 2015; Michnowicz 2015).

Furthermore, as some researchers argue, there is sufficient evidence for contact-induced change only if there are differences in constraints (grammatical patterning) between the varieties under study but rates alone do not substantiate change (Poplack & Levey 2010; Torres Cacoullos & Travis 2010). Therefore, in terms of our question of the role of language contact, we do see evidence in favor of an effect of bilingualism in general given the differences in constraints observed in the current data. While subject expression in Roswell is governed by many of the same constraints discussed in previous literature for both contact and non-contact Spanish (*Person/Number* [Cameron 1993; Silva-Corvalán 1994a; Flores-Ferrán 2002; Orozco & Guy 2008; Shin 2012; Lastra & Martín Butragueño 2015]; *Switch Reference*, [Bentivoglio 1987; Cameron 1994; Bayley & Pease-Alvarez 1997; Torres Cacoullos & Travis 2010; Otheguy & Zentella 2012; Michnowicz 2015; Orozco 2015]; *Morphological Ambiguity* [Hochberg 1986; Travis 2007; Prada Pérez 2009; Erker & Guy 2012; Lastra & Martín Butragueño 2015; Michnowicz 2015]; and *Verb Class* [Bentivoglio 1987; Silva-Corvalán 1994a; Travis 2007; Otheguy & Zentella 2012; Orozco 2015), when compared to monolingual Mexican Spanish there are more differences than similarities. In particular, Roswell Spanish exhibits differential

predictive power for certain conditioning factors (*morphological ambiguity*, *polarity*), the emergence of some (*verb class*), and the complete loss of others (*TMA, verbal mood*). These differences in constraints suggest processes of both simplification (Sorace 2012; Michnowicz 2015) and complexification (Shin 2014), which are processes documented in other varieties of contact Spanish, such as Spanish-English varieties (Silva-Corvalán 1994a; Shin & Otheguy 2009; Shin 2014) as well as Spanish-Maya varieties (Michnowicz 2015). While we cannot see evidence of English transfer *per se*, as stated above, these processes do suggest that bilingualism plays a role. Specifically, the effect of English seems to be **indirect** (Silva-Corvalán 1994b): English happens to be the contact language, but most evidence points toward an effect of bilingualism itself rather than of English in particular.[67]

Finally, if we compare Roswell Spanish to other, more established contact varieties, we see that the shifts in constraints and the behavior that indicates the above-mentioned processes of grammatical simplification and complexification make Spanish in Roswell look more like stable contact varieties, particularly varieties of Spanish in New York City and Los Angeles (e.g. Flores Ferrán 2002, Shin & Otheguy 2009, Silva-Corvalán 1994a; Shin 2014). Such differential linguistic configuration in terms of variable subject expression in the Spanish of Roswell Mexicans suggests that this is an *emergent* variety (Moreno-Fernández 2009) of Mexican Spanish in the U.S., one that is distinct from both its monolingual counterparts and the more established varieties of U.S. born bilinguals. That is, we see substantial divergences from monolingual Mexican Spanish, but also evidence that Roswell Spanish is in incipient stages of language shift that are not to the same extent as the more advanced stages of shift in long-

[67] See Silva-Corvalán (1994b) for a similar discussion regarding usage of the subjunctive and indicative moods in Los Angeles Spanish.

standing Spanish varieties in the U.S. The Mexican Spanish dialect of Roswell, Georgia, then, is emerging as a new variety.

7.2 Contributions and future directions

At the beginning of this dissertation, I underscored the fact that Spanish in the Southeastern U.S. is generally understudied in comparison to other U.S. regions. Carrying out the present study has contributed to the sociolinguistic literature by further filling in this regional gap, by dialoguing with previous studies on Spanish subject expression, and by expanding on issues of SP variation previously explored. Moreover, the current investigation has afforded the timely opportunity to document the language of Spanish-speakers in a relatively recent contact situation in the U.S. with the purpose of not only understanding initial stages of language contact and language change, but also providing findings that can serve as a baseline and point of comparison for future studies that examine language contact at more advanced stages in the Southeastern U.S.

More specifically to scholarship on subject expression, this dissertation has made the following contributions: First, it has explored and answered questions about linguistic and social variables that have received little attention in previous research such as polarity, speech style, and preferred media language. Regarding polarity and speech style, the present findings corroborate the results of the very few studies that have employed these factors (Ávila-Jiménez 1996; Lastra & Martín Butragueño 2015; Geeslin & Gudmestad 2016). Additionally, interactions were found between polarity, speech style, and other variables that further explain and provide a more nuanced view of SP variation. For example, it was revealed that the effect of speech style applies only to first-person singular verbs and not to all persons/numbers. It was also found that the polarity effect is constrained by other factors like verb class, person/number, and switch

reference. With regard to preferred media language, this variable was previously not taken into account in any SP study to my knowledge. The present study found a lack of such influence on subject expression, suggesting that exposure to television or radio in English does not affect pronoun usage in particular. Secondly, it was revealed in carrying out this study that the TMA effect is not always operative cross-dialectally since it was not found to be significant in the multivariate analysis. This is an exceptional case and requires further research to determine the particular motivation for this pattern as well as its relation to morphological ambiguity.

Finally, this dissertation has contributed to the study of lexical subjects through a preliminary qualitative analysis, going beyond the exclusive study of pronominal forms seen in most of the variationist literature. This lexical subjects analysis showed that lexical noun phrases are often repeated soon after their first mention, lending support to Dumont's (2006) analysis which underscores that lexical noun phrases are used beyond simply introducing a new referent. Such pattern observed in the current data also raised questions concerning the possible effect of priming on lexical subject use. Methodologically speaking, a variable context for lexical subjects was established in this dissertation (see Chapter 4) that had not been acknowledged in previous studies.

Moving forward, it would be beneficial for future research to investigate several issues. First, it would be useful to explore a wider range of independent linguistic variables than the ones included in the present investigation. The analysis of factors such as priming, lexical frequency, and reflexivity with respect to their influence on subject expression would provide us with more explanatory value concerning how SP variation is regulated language-internally. Secondly, an expansion on the interactions between the variables would uncover further complexities with regard to how multiple variables intersect to regulate subject expression

patterns. In addition, as stated in Chapter 5, a more thorough examination of lexical subjects that involves quantitative analysis of variation relative to both overt and null subjects is warranted. This would help us to further explain what motivates speakers to produce, for example, a lexical subject when they could have produced an overt SP (or vice versa). Third, with regard to the selection of participants, future investigations should include a more balanced speaker sample, particularly regarding levels of education. For example, including more speakers with a university education would be more representative of the Spanish-speaking community.

Finally, it is crucial to explore issues that will help us gain more insight into the role of language contact. For instance, cross-generational data would be very useful in revealing how the Spanish of second-generation and/or third-generation immigrants within the same community compares with that of the first-generation speakers employed in the present study. Do we see evidence of contact-induced change that was absent from the present analysis, such as correlations between pronoun use and English proficiency? Would we observe more dramatic constraint shifts that reveal a more advanced stage of simplification or complexification? Having data from other immigrant generations would help to answer these questions and reveal in more detail how Spanish is changing in Georgia and in the Southeastern U.S. more generally. A second issue that would be interesting to examine is dialect contact, that is, the linguistic behavior of Spanish speakers who are in contact with speakers of different varieties of Spanish. For example, how is Mexican Spanish influenced by Colombian Spanish or Venezuelan Spanish in the Southeast? How do these patterns compare to the dialect contact effects documented in other regions, such as New York City (e.g. Otheguy & Zentella 2012)? Aside from the opportunity to study dialect contact, employing speakers of other national origins apart from

Mexico would portray a more diverse and representative picture of Spanish and Spanish-speakers in Roswell and in the Southeastern U.S.

REFERENCES

Abreu, Laurel. 2012. Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, ed. by Kimberly Geeslin and Manuel Díaz-Campos, 1-8. Somerville, MA: Cascadilla Proceedings Project.

Alfaraz, Gabriela G. 2015. Variation of overt and null subject pronouns in the Spanish of Santo Domingo. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 3-16. Georgetown: Georgetown University Press.

Amaral, Patrícia Matos and Scott A. Schwenter. 2005. Contrast and the (non-) occurrence of subject pronouns. In *Selected proceedings of the 7th Hispanic linguistics symposium*, ed. by David Eddington, 116-127. Somerville, MA: Cascadilla Proceedings Project.

Ariel, Mira, 1994. Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. *Journal of Linguistics* 30.3-42.

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

Barrenechea, Ana María and Alicia Alonso. 1977. Los pronombres personales sujetos en el español hablado en Buenos Aires. In *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; Kristen Greer; and Cory C. Holland. 2013. Lexical frequency and syntactic variation: A test of a linguistic hypothesis. In *University of Pennsylvania Working Papers in Linguistics* 19(2).21-30.

Bayley, Robert and Lucinda Pease-Alvarez. 1996. Null and expressed pronoun variation in Mexican-descent children's Spanish. In *Sociolinguistic variation*: *Data, theory, and analysis*, ed. by Jennifer Arnold, Renee Blake and Brad Davidson, 85-99. Stanford: Center for the Study of Language and Information.

Bayley, Robert and Lucinda Pease-Alvarez. 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9.349–71.

Beck, Jr., Robin A.; David G. Moore; and Christopher B. Rodning. 2006. Identifying Fort San Juan: A sixteenth-century Spanish occupation at the Berry site, North Carolina. *Southeastern Archaeology* 25(1).65-77.

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

Bentivoglio, Paola; Luis A. Ortiz; and Carmen Silva-Corvalán. 2011. La variable expresión del sujeto pronominal. Guía de codificación. In *PRESEEA – Grupo de análisis de expresión de sujetos*. Retrieved from *http://preseea.linguas.net/portals/0/metodologia /guia_codificacion _ sujetos_julio_2011.pdf*

Bessett, Ryan M. 2018. Testing English influence on first person singular "yo" subject pronoun expression in Sonoran Spanish. In *Contemporary trends in Hispanic and Lusophone linguistics: Selected papers from the Hispanic Linguistic Symposium 2015*, ed. by Jonathan E. MacDonald, 355-372. Amsterdam: John Benjamins.

Blackwell, Sarah E. 1998. Constraints on Spanish NP anaphora: The syntactic versus the pragmatic domain. *Hispania* 81(3).607-618.

Blackwell, Sarah E. 2001. Testing the neo-Gricean pragmatic theory: The influence of consistency constraints on implicatures of coreference. *Journal of Pragmatics* 33.901-41.

Blackwell, Sarah. E. 2003. *Implicatures in discourse: The case of Spanish NP anaphora.* Amsterdam: John Benjamins.

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. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, ed. by Kimberly Geeslin and Manuel Díaz-Campos, 142-64. Somerville, MA: Cascadilla Proceedings Project.

Blanco Canales, Ana. 1999. Presencia/ausencia de sujeto pronominal de primera persona en español. *Español Actual* 72.31-40.

Bosch, Peter. 1983. *Agreement and anaphora: A study of the role of pronouns in syntax and discourse*. London: Academic Press.

Bosque, Ignacio and Violeta Demonte. 1999. *Gramática descriptiva de la lengua española*. Madrid: Espasa.

Butt, John and Carmen Benjamin. 2004. *A new reference grammar of modern Spanish*. London: Arnold.

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

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. In *Papers from the regional meeting of the Chicago Linguistic Society*: Parasession on variation in linguistic theory 30.27–45.

Carter, Phillip M. 2005. Quantifying rhythmic differences between Spanish, English, and Hispanic English. In *Theoretical and experimental approaches to Romance linguistics: Selected papers from the 34th Linguistic Symposium on Romance Languages*, ed. by Randall Gess, 63-75. Amsterdam: John Benjamins.

Carter, Phillip M. 2014. National narratives, institutional ideologies, and local talk: The discursive production of Spanish in a 'new' US Latino community. *Language in Society* 43(2).209-40.

Carvalho, Ana M. and Ryan M. Bessett. 2015. Subject pronoun expression in Spanish in contact with Portuguese. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 143-65. Georgetown: Georgetown University Press.

Carvalho, Ana M. and Michael Child. 2011. Subject pronoun expression in a variety of Spanish in contact with Portuguese. In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, ed. by Jim Michnowicz and Robin Dodsworth, 14-25. Somerville, MA: Cascadilla Proceedings Project.

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.

Dumont, Jenny. 2006. Full NPs as subjects. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. by Nuria Sagarra and Almeida Jacqueline Toribio, 286-96. Somerville, MA: Cascadilla Proceedings Project.

Eckert, Penelope and Sally McConnell-Ginet. 2003. *Language and gender*. Cambridge: Cambridge University Press.

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

Erker, Daniel and Gregory R. Guy. 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88.526-57.

Escobar, Anna María and Kim Potowski. 2015. *El español de los Estados Unidos*. Cambridge: Cambridge University Press.

Flores, Antonio. 2017. How the U.S. Hispanic population is changing. Retrieved from http://www.pewresearch.org/fact-tank/2017/09/18/how-the-u-s-hispanic-population-is-changing/

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

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

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. 2007. Los mexicanos in New Jersey: Pronominal expression and ethnolinguistic aspects. In *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.

Geeslin, Kimberly and Aarnes Gudmestad. 2016. Subject expression in Spanish: Contrasts between native and non-native speakers for first and second-person singular referents. *Spanish in Context* 13(1).53-79.

Geeslin, Kimberly; Bret Linford; and Stephen Fafulas. 2015. Variable subject expression in second language Spanish: Uncovering the developmental sequence and predictive linguistic factors. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 191-210. Georgetown: Georgetown University Press.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In *Topic continuity in discourse: A quantitative cross-language study*, ed. by Talmy Givón, 1-41. Amsterdam: John Benjamins.

Gundel, Jeanette K. 1999. On different kinds of focus. In *Focus: Linguistic, cognitive, and computational perspectives*, ed. by Peter Bosch and Rob van der Sandt, 293-305. Cambridge: Cambridge University Press.

Gundel, Jeanette K.; Nancy Hedberg; and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69.274-307.

Gutiérrez-Rivas, Carolina. 2007. *Variación pragmática de los cubanos y cubanoamericanos en Miami. El efecto de género y generación en el uso de estructuras discursivas*. PhD dissertation, University of Florida.

Hidalgo Downing, Raquel. 2003. *La tematización en el español hablado*. Madrid: Gredos.

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

Hoffman, Michol. 2013. Sociolinguistic interviews. In *Research methods in sociolinguistics: A practical guide*, ed. by Janet Holmes and Kirk Hazen, 25-41. Chichester, England: Wiley-Blackwell.

Hurtado, Aída and Luis A. Vega. 2004. Shift happens: Spanish and English transmission between parents and their children. *Journal of Social Issues* 60(1).137-55.

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.

King, Larry D., and Margarita Suñer. 2004. *Gramática española: Análisis y práctica*. New York, NY: McGraw-Hill.

Klee, Carol A. and Andrew Lynch. 2009. *El español en contacto con otras lenguas*. Washington, D.C.: Georgetown University Press.

Kretzschmar, 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.

Labov, William and Joshua Waletzky. 1967. Oral narratives of personal experience. Obtained from http://www.ling.upenn.edu/~wlabov/Papers/FebOralNarPE.pdf

Labov, William. 1966. *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

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

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

Lapidus, Naomi and Ricardo Otheguy. 2005. Overt nonspecific *ellos* in Spanish in New York. *Spanish in Context* 2(2).157–74.

Lastra, Yolanda and Pedro Martín Butragueño. 2015. 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, 39-57. Georgetown: Georgetown University Press.

Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, Massachusetts.: MIT Press.

Limerick, Philip P. 2017. Language contact in the US Southeast: The case of Spanish subject expression in an emerging bilingual community in Georgia. *Spanish in Context* 14(1).53-77.

Limerick, Philip P. 2018. The discursive distribution of subject pronouns in Spanish spoken in Georgia: A weakening of pragmatic constraints? *Studies in Hispanic and Lusophone Linguistics*. doi: https://doi.org/10.1515/shll-2018-0013.

Lipski, John M. 2005. Code-switching or borrowing? No sé *so* no puedo decir, *you know*. In *Selected proceedings of the Second Workshop on Spanish Sociolinguistics*, ed. by Lofti Sayahi and Maurice Westmoreland, 1-15. Somerville, MA: Cascadilla Proceedings Project.

Lipski, John M. 2008. *Varieties of Spanish in the United States*. Washington, D.C.: Georgetown University Press.

Llombart, Alberto G. 2003. Do you habla Spanglish? *Especulo* 23(1).1–6.

Lozano, Cristóbal. 2009. Selective deficits at the syntax-pragmatics interface: Evidence from the CEDEL2 corpus. In *Representational deficits in second language acquisition*, ed. by Neal Snape, Yan-kit Ingrid Leung, and Michael Sharwood Smith, 127-66. Amsterdam: John Benjamins.

Lynch, Andrew. 1999. *The subjunctive in Miami Cuban Spanish: Bilingualism, contact, and language variability*. PhD dissertation, University of Minnesota, Minneapolis.

Mayol, Laia. 2010. Contrastive pronouns in null-subject Romance languages. *Lingua* 120.2497-2514.

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

Meyerhoff, Miriam. 2009. Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change* 21.297-317.

Michnowicz, Jim. 2015. Subject pronoun expression in contact with Maya in Yucatan Spanish. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 101-119. Georgetown: Georgetown University Press.

Milroy, Lesley and James Milroy. 1992. Social network and social class: Toward an integrated sociolinguistic model. *Language in Society* 21.1–26.

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

Montrul, Silvina A. and Celeste Rodríguez Louro. 2006. Beyond the syntax of the null subject parameter: A look at the discourse-pragmatic distribution of null and overt subjects by L2 learners of Spanish. In *The acquisition of syntax in Romance languages*, ed. by Vincent Torrens and Linda Escobar, 401-18. Amsterdam: John Benjamins.

Morales, Amparo. 1986. *Gramáticas en contacto: Análisis sintácticos sobre el español de Puerto Rico*. Madrid: Playor.

Morales, Amparo. 1997. La hipótesis funcional y la aparición del sujeto no nominal: El español de Puerto Rico. *Hispania* 80.153–65.

Moreno Fernández, Francisco. 2009. Dialectología hispánica de los Estados Unidos. In *Enciclopedia del español en los Estados Unidos*, coord. by Humberto López Morales, 200-221. Madrid: Instituto Cervantes/Santillana.

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

Olloqui de Montenegro, Liliana de. 1987. Un aspecto de la sintaxis: Los pronombres personales sujeto en el habla estudiantil santiaguera. In *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 Gregory Guy. 2008. El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana? In *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*, ed. by Maurice Westmoreland and Juan Antonio Thomas, 70-80. Somerville, MA: Cascadilla Proceedings Project.

Orozco, Rafael. 2015. Pronominal variation in Colombian Costeño Spanish. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 17-37. Georgetown: Georgetown University Press.

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

Otheguy, Ricardo; Ofelia García; and Mariela Fernández. 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; Ana Celia Zentella; and David Livert. 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 83.770-802.

Otheguy, Ricardo and Ana Celia Zentella. 2012. *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. New York: Oxford University Press.

Pastor, Manuel, Jr. 1998. Interdependence, inequality, and identity: Linking Latinos and Latin Americans. In *Borderless borders: US Latinos, Latin Americans, and the paradox of interdependence*, ed. by Frank Bonilla, Edwin Meléndez, Rebecca Morales, María de los Angeles Torres, 17-33. Philadelphia, PA: Temple University Press.

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 Stephen Levey. 2010. Contact induced grammatical change: A cautionary tale. In *Language and space: An international handbook of linguistic variation*, ed. by Peter Auer and Jürgen Erich Schmidt, 391-419. Berlin: Mouton de Gruyter.

Poplack, Shana and Sali Tagliamonte. 1989. There's no tense like the present: Verbal -s inflection in early Black English. *Language Variation and Change* 1(1) 47-84.

Posio, Pekka. 2011. Spanish subject pronoun usage and verb semantics revisited: First and second person singular subject pronouns and focusing of attention in spoken Peninsular Spanish. *Journal of Pragmatics* 43(3).777-98.

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

Prada Pérez, Ana de. 2015. First person singular subject pronoun expression in Spanish in contact with Catalan In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 121-142. Georgetown: Georgetown University Press.

Quesada, Margaret L. 2015. *The L2 acquisition of Spanish subjects: Multiple perspectives*. Berlin/New York: Mouton de Gruyter.

Quesada, Margaret Lubbers and Sarah Blackwell. 2009. The L2 acquisition of null and overt Spanish subject pronouns: A pragmatic approach. In *Selected proceedings of the 11th Hispanic linguistics symposium*, ed. by Joseph Collentine, Maryellen García, Barbara Lafford and Francisco Marcos Marín, 117-30. 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.

Rivera-Mills, Susana V., and Daniel J. Villa. 2010. Introduction. Spanish of the U.S. Southwest: A language in transition. In *Spanish of the U.S. Southwest: A language in transition*, ed. by Susana V. Rivera-Mills and Daniel J. Villa, 11-14. Madrid: Iberoamericana/Vervuert.

Roever, Carsten and Donald E. Powers. 2005. Effects of language of administration on a self-assessment of language skills. Princeton, NJ: Educational Testing Service.

Romaine, Suzanne. 2003. Variation in language and gender. In *The handbook of language and gender*, ed. by Janet Holmes and Miriam Meyerhoff, 98-118. Oxford: Blackwell.

Rothman, Jason and Amy Beth Rell. 2005. A linguistic analysis of Spanglish: Relating language to identity. *Linguistics and Human Sciences*. 1(3).515–36.

Said-Mohand, Aixa. 2008. Aproximación sociolingüística al uso de *entonces* en el habla de jóvenes bilingües estadounidenses. *Sociolinguistic Studies* 2.97-130.

Sánchez, Miriam Sancho. 2001. Algunas reflexiones de un viaje por el spanglish y el español de Latinoamérica. *Cuadernos Cervantes de la Lengua Española* 32(7).10–11.

Sankoff, David. 1988. Sociolinguistics and syntactic variation. In *Linguistics: The Cambridge survey*, ed. by Frederick J. Newmeyer, 140-61. Cambridge: Cambridge University Press.

Sankoff, Gillian. 2006. Age: Apparent time and real time. In *Encyclopedia of Language and Linguistics*, ed. by Keith Brown, 110-116. Amsterdam: Elsevier.

Schwenter, Scott A. 2002. The notion of contrast in the study of Spanish subject personal pronouns. Paper given at the *Workshop on subject pronoun expression in Spanish*, CUNY Graduate Center.

Schwenter, Scott A. 2017. Null objects with and without bilingualism in the Portuguese- and Spanish-speaking world. In *Multidisciplinary approaches to bilingualism in the Hispanic and Lusophone world*, ed. by Kate Bellamy, Michael W. Child, Paz González, Antje Muntendam, and M. Carmen Parafita Couto, 95-119. Amsterdam: John Benjamins.

Scrivner, Olga and Manuel Díaz-Campos. 2016. Language Variation Suite: A theoretical and methodological contribution for linguistic data analysis. In *Proceedings of the Linguistic Society of America Vol. 1, Article 29,* 1-15.

Shin, Naomi Lapidus. 2010. Efficiency in lexical borrowing in New York Spanish. *International Journal of the Sociology of Language* 2010(203).45-60.

Shin, Naomi Lapidus. 2012. Variable use of Spanish subject pronouns by monolingual children in Mexico. In *Selected proceedings of the 14th Hispanic linguistics symposium*, ed. by Kimberly Geeslin and Manuel Díaz-Campos, 130-141. Somerville, MA: Cascadilla Proceedings Project.

Shin, Naomi Lapidus. 2014. Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change* 26.303-30.

Shin, Naomi and Daniel Erker. 2015. The emergence of structured variability in morphosyntax: Childhood acquisition of Spanish subject pronouns. In *Subject Pronoun Expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 169-190. Georgetown: Georgetown University Press.

Shin, Naomi Lapidus and Cecilia Montes-Alcalá. 2014. El uso contextual del pronombre sujeto como factor predictivo de la influencia del inglés en el español en Nueva York. *Sociolinguistic Studies* 8(1).85-110.

Shin, Naomi Lapidus and Ricardo Otheguy. 2009. Shifting sensitivity to continuity of reference: Subject pronoun use in Spanish in New York City. In *Español en Estados Unidos y otros contextos: Sociolingüística, ideología y pedagogía*, ed. by Manel LaCorte and Jennifer Leeman, 111-36. Madrid: Iberoamericana.

Shin, Naomi Lapidus and Ricardo Otheguy. 2013. Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. *Language in Society* 42(4). 429-52.

Shin, Naomi Lapidus and Jackelyn Van Buren. 2016. Maintenance of Spanish subject pronoun expression patterns among bilingual children of farmworkers in Washington/Montana. *Spanish in Context* 13(2).173-94.

Silva-Corvalán, Carmen. 1982. Subject expression and placement in Mexican-American Spanish. In *Spanish in the United States: Sociolinguistic aspects*, ed. by Jon Amastae and Lucía Elías-Olivares, 93–120. Cambridge: Cambridge University Press.

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

Silva-Corvalán, Carmen. 1994b. The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change*. 6(3).255-72.

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

Silva-Corvalán, Carmen. 2008. The limits of convergence in language contact. *Journal of Language Contact* THEMA 2.213-24.

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

Solé, Yolanda. 1977. Continuidad / descontinuidad idiomática en el español tejano. *The Bilingual Review / La Revista Bilingüe* 4.189-99.

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

Sorace, Antonella. 2004. Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition* 7(2).143-45.

Sorace, Antonella. 2012. Pinning down the concept of interface in bilingual development: A reply to peer commentaries. *Linguistic Approaches to Bilingualism* 2(2).209–17.

Sperling, Bert. 2005. Sperling's best places. Accessed February 27, 2018. http://www.bestplaces.net/city/georgia/roswell

Steinmetz, Sara and Francisco Moreno-Fernández. 2015. Mapa hispano de los Estados Unidos. *Informes del Observatorio / Observatorio Reports* 14.1-85.

Stewart, Miranda. 2003. 'Pragmatic weight' and face: Pronominal presence and the case of the Spanish second person singular subject pronoun *tú*. *Journal of Pragmatics* 35(2).191-206.

Sturtevant, William C. 1962. Spanish-Indian relations in Southeastern North America. *Ethnohistory* 9(1).41-94.

Tagliamonte, Sali A. 2002. Comparative sociolinguistics. In *The handbook of language variation and change*, ed. by J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes, 729-763. Malden and Oxford: Blackwell.

Tagliamonte, Sali A. 2003-2006. Linguistic changes in Canada entering the 21st century. Research Grant. Social Sciences and Humanities Research Council of Canada (SSHRC). #410-2003-0005.

Tagliamonte, Sali A. 2006. *Analysing sociolinguistic variation*. Cambridge, NY: Cambridge University Press.

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

Tagliamonte, Sali A., and R. Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2).135-78.

Thompson, Sandra A. and Anthony Mulac. 1991. A quantitative perspective on grammaticalization of epistemic parentheticals in English. In *Approaches to grammaticalization, Vol. 2*, ed. by Elizabeth Traugott and Bernd Heine, 313-29. Amsterdam: John Benjamins.

Torres Cacoullos, Rena and Catherine E. Travis. 2010. Variable yo expression in New Mexico: English influence? In *Spanish of the U.S. Southwest: A language in transition*, ed. by Susana Rivera-Mills and Daniel Villa, 185-206. Madrid: Iberoamericana/Vervuert.

Torres Cacoullos, Rena and Catherine E. Travis. 2011. Testing convergence via codeswitching: Priming and the structure of variable subject expression. *International Journal of Bilingualism* 15(3).241–67.

Torres Cacoullos, Rena and Catherine E. Travis. 2015. Foundations for the study of subject pronoun expression in Spanish in contact with English: Assessing interlinguistic (dis)similarity via intralinguistic variability. In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, ed. by Ana M. Carvalho, Rafael Orozco and Naomi Lapidus Shin, 81-100. Georgetown: Georgetown University Press.

Torres, Lourdes. 2002. Bilingual discourse markers in Puerto Rican Spanish. *Language in Society* 31.65-83.

Travis, Catherine E. 2005. The yo-yo effect: Priming in subject expression in Colombian Spanish. In *Selected Papers from the 34th Linguistic Symposium on Romance Languages (LSRL)*, ed. by Randall Gess and Edward. J. Rubin, 329–49. Amsterdam: John Benjamins.

Travis, Catherine E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19.101-35.

Travis, Catherine E. and Rena Torres Cacoullos. 2012. What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics* 23(4).711-48.

Travis, Catherine E. and Rena Torres Cacoullos. 2013. Making voices count: Corpus compilation in bilingual communities. *Australian Journal of Linguistics* 33(2).170-194.

U.S. Census Bureau. 2010. American Fact Finder. Retrieved from *http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml*.

Veltman, Calvin. 2000. The American linguistic mosaic: Understanding language shift in the United States. In *New immigrants in the United States*, ed. by Sandra McKay and Sau-Ling Cynthia Wong, 58-95. Cambridge: Cambridge University Press.

Waltermire, Mark. 2014. The influence of English on U.S. Spanish: Introduction. *Sociolinguistic Studies* 8(1).1-21.

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

Winders, Jamie. 2005. Changing politics of race and region: Latino migration to the US South. *Progress in Human Geography* 29(6).683–99.

Wolfram, Walt; Phillip Carter; and Becky Moriello. 2004. Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8.339-58.

Wolfram, Walt; Mary E. Kohn; and Erin Callahan-Price. 2011. Southern-Bred Hispanic English: An emerging socioethnic variety. In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, ed. by Jim Michnowicz and Robin Dodsworth, 1-13. Somerville, MA: Cascadilla Proceedings Project.

Yáñez, Rosa H. 1990. The complimenting speech act among Chicano women. In *Spanish in the United States: Sociolinguistic issues*, ed by John J. Bergen, 79-85. Washington, DC: Georgetown University Press.

Zentella, Ana Celia. 1997. *Growing up bilingual*. Malden, MA: Blackwell Publishers.

APPENDIX

SAMPLE INTERVIEW

Speaker: F26Mex

Date of interview: June 30, 2015

Place of interview: Roswell Latin Plaza

I: Pues estamos aquí en Latin Roswell, Ro- Roswell Latin Plaza y... y yo soy Philip Limerick, es el treinta de junio de dos mil quince, y su ¿y tu nombre es?

R: María ((pseudonym))

I: María, OK um... ¿cuál es tu fecha de nacimiento?

R: Junio.

I: ¿Perdón?

R: junio

I: ¿Junio? ¿De qué día?

R: X

I: X y ¿qué año?

R: X

I: OK bien, em puedes hablar un poco sobre tu educación escolar?

R: Sí, yo estudié, en México,

I: Uh-huh.

R: y me gradué, aquí, en Estados Unidos, de la preparatoria.

I: Oh OK. ¿Cuántos años asiste aquí entonces?

R: Aquí cuatro, a la high school.

I: Oh OK, muy bien, um, ¿me puedes hablar un poco sobre um tu trabajo actual y cualquier otros trabajos anteriores?

R: Mis trabajos anteriores em, todo el tiempo he tenido trabajos de, cashier, he tenido contacto siempre con la gente como como... customer service,

I: OK.

R: eh, prácticamente ha sido siempre mi puesto de trabajo.

I: OK OK, muy bien, y ¿vives aquí en Roswell o?

R: Mhm, sí.

I: Sí, mhm. ¿En qué ciudad naciste?

R: Yo nací en la ciudad de Monterrey.

I: ¿De…?

R: Nuevo León, XMonterrenolX.

I: ¿Cuándo viniste aquí a Estados Unidos?

R: Vine cuando tenía 14 años.

I: Oh OK um... y ¿viniste directamente a Georgia o también viviste en otros lugares?

R: Viví en Florida.

I: en Florida también... y luego ¿cuándo viniste a-

R: Después, después de cinco años, allá, vine para acá,

I: Oh OK.

R: Ahorita ya voy hacer casi, doce años...

I: Casi doce años, acá

R: Mhm.

I: Muy bien, gracias, ¿me puedes hablar un poco de tu familia, de tus padres, abuelos, hermanos?

R: Mi familia entera... está... en México, por parte de mi papá, eh algo- tengo familiares aquí, y tengo familiares en Florida.

I: ¿Y en México tienes uh... tienes familia extendida?

R: Sí, la de mis papás, la familia de mis papás están todos allá, pero mis papás están aquí.

T: 2.48

I: Sí OK, muy bien, um... ¿puedes hablar un poco de tus contacte- contactos sociales en la comunidad um, compañeros de trabajo, amigos, pues son los con que con quienes estás en contacto más o menos regularmente?

R: Pues, amigos casi no tengo, me dedico más a trabajar, todo el tiempo estoy trabajando, trabajo desde muy temprano hasta muy noche,

I: OK.

R: y este, pues mis compañeros de trabajo todos son... eh... pues se llevan bien conmigo yo soy una persona muy sociable, y trato de llevarme siempre, bien con mis compañeros de trabajo para que no haya problemas,

I: Sí, sí sí.

R: Amigos, pocos, pero casi no los veo.

I: OK. Y ¿sólo trabajas acá o tienes otro-

R: No, soy también secretaria de una base de taxis.

I: Oh OK, ¿y dirías que la mayoría de la gente con que estás en contacto son hispanohablantes o anglohablantes?

R: Eh, de los dos, porque en mi otro trabajo hay mucho... mucha persona que habla, inglés,

I: Oh OK.

R: pero no son americanos son como de Haití, o de otros lugares,

I: Oh sí.

R: No sé exactamente de dónde pero son de otros lugares y no hablan español.

I: ¿Entonces tanto los compañeros y los clientes?

R: Sí, pues, en mi trabajo de la mañana, hay muchos clientes americanos,

I: OK.

R: O, sí, americanos personas, que ah- me refiero a personas que hablan inglés, pueden ser y... de, de la India de, muchos lugares no sé de dónde.

I: ¿Entonces hablas mayormente inglés en este trabajo?

R: Sí.

I: OK, muy bien, um, ¿hablas otros idiomas a parte de español y inglés?

R: No, sólo español e inglés.

I: X

R: @

I: he aprend- he aprendido un poco de portugués, pero, X

R: Soy mala para aprender idiomas @

I: es difícil…OK um dirías que tienes una lengua preferida, entre los dos?

R: mm... se me complica mucho el inglés,

I: Mhm.

R: ah me X T:5:12 una leng- una lengua, a mí en lo personal ¿no? hay personas que la agarran muy rápido para mí en lo personal se me hace muy difícil,

I: Uh-huh.

R: um, sí lo hablo, porque estudié aquí y aprendí a escribir y todo eso pero, no es que me guste mucho, X T:5:28 me gusta mucho mi idioma.

I: Sí sí sí, también para medios de comunicación, música, periódicos, prefieres español?

R: Sí, mayormente en español, pero por ejemplo hay cosas que no puedes leer en español, porque a veces las traducciones no son buenas,

I: Sí.

R: entonces es preferible mejor leerlo en inglés que es más entendible que, el español.

T: 5.54

I: Muy bien, um si tuvieras que autoevaluarte como tu nivel de inglés como en una escala de uno a cinco, donde pondrías como uno está muy bien y cinco muy mal, qué dirías?

R: no, yo creo que sí estoy bastante bien pero no un cien por ciento no, digamos que un, cuatro,

I: Sí.

R: No es que que cien por ciento, no.

I: OK , ¿y español?

R: No al español sí @.

I: ¿Cinco?

R: Cinco sí.

I: OK, y ¿aprendiste algo de inglés en México o empezaste cuando viniste?

R: Según eran clases de inglés, pero no te enseñan el inglés que, realmente es, sólo te enseñan algunas palabras como ¿cómo estás? o ¿cómo te llamas? pero nada, serio como para entablar una conversación no, el inglés lo aprendí acá.

I: Oh OK, muy bien, um... OK ¿sabes algo del papel que juegan las iglesias acá en la comunidad de Roswell, si tienen un papel importante en las vidas de la gente?

R: La verdad no sabría decirte no... no es que no crea en Dios, yo creo en Dios, eh fui muchos años a una iglesia cristiana, pero aquí, aquí aquí no sé realmente cómo esté eso la verdad no,

I: OK.

R: No me he interesado ir a otra iglesia que no sea la que y- q- en la donde yo me crié, so entonces, no sé mucho cómo.

I: Está bien, um... ¿por casualidad sabes algo de las bodas aquí, como se celebran o cuales son las tradiciones?

R: Tampoco, le- en la cuestión americana, ¿no? … los Americanos como las celebran o?

I: Los hispanos también.

R: Los hispanos pues claro ¿no? el, la... como el típico como lo normal que siempre se ha vivido, claro que la gente cristiana no lo celebra igual, eh pero sí lo normal que la ceremonia del civil en la iglesia,

I: Uh-huh.

R: que hacen sus votos y no sé qué hacen, yo por ejemplo no soy católica y no sé qué tanto trámite tengan que hacer… nada, lo normal comida baile,

I: Sí.

R: y trago @

I: @ OK vamos a hablar un poco ahora de la vida cotidiana aquí, ¿me puedes describir como es para ti un día típico?

R: Mi vida es muy... (2) es como una rutina, todos los días, es lo mismo, la única diferencia que hace es cuando yo estoy con mi, con mi pareja, y él me al- me alegra un poco el día ¿no? verlo,

I: Sí.

R: y hacer cosas diferentes pero, es aquí, yo siento que no hay mucha... mucha distracción mucha vida acá es, trabajar, y a la casa, trabajar y a la casa, y el único momento en el que puedes disfrutar es los fines de semana y si, trabajas los fines de semana pues no puedes.

T: 9.15

I: Uh-huh, OK, um... vamos a hablar de cuando eras niña en México, qué tipos de... um ¿tenías quehaceres que hacer en casa?

R: Claro la crianza en México es totalmente diferente, allá lo enseñan a uno desde muy pequeño a que tiene responsabilidades, la vida es totalmente diferente allá, allá no había lujos de PlayStation tablets computadoras todo eso, no existía, allá, cada quien tenía que tender su cama,

recoger su cuarto, limpiar la casa, colaborar, y no tenías, que, pensar que era como obligado, que tú estabas pagando algún castigo o algo, simplemente eran rutinas de la casa, era lo que te correspondía hacer, pero, la infancia aquí es muy, diferente a la de allá, allá lo enseñan a uno a ser, más trabajador más acomedido,

I: Uh-huh.

R: aquí yo no veo eso… allá te enseñan a respetar a tus mayores,

I: Claro.

R: ah cualquier cosa y ahí sí te @, te daban @, pero, no es que tengan que pegarle a los niños porque no es así Xse puede hablarX ¿no? pero, sí es diferente.

I: Y a parte como de actividades, qué tipos de actividades hacías de niña?

R: Pues por ejemplo, si, mi mamá estaba limpiando la casa yo tenía que ayudar, yo lavaba mi ropa, allá no había lavadoras no había secadoras, ay a mí me tocaba lavar mi ropa, hasta que ya tenía una cierta edad ¿no? yo doblaba mi ropa la guardaba, recogía mi habitación, colaboraba también en la cocina todo eso pero, no era la ama de casa sólo ayudaba, pero no lo hacía por castigo o no lo veía yo como que si era un castigo lo veía como algo, normal, porque ahora los niños no piensan eso, piensan que, porque los mandas a tender la cama los estás castigando, y piensan que X T:11:30 trabajo y no es así.

I: ¿Y había actividades que que hacías para divertirte?

R: Claro en las noches o las tardes cuando ya todos salían de las escuelas, que hacías tu tarea y todo eso ya tus papás te dejaban salir, y pues eran, juegos en la calle… físicos, salías a correr por horas, a jugar sí, y este pero nada de…de electrónicos ni nada de eso.

T: 12.02

I: Sí, muy bien, pues siempre es divertido... está X T: 12.08 está de juegos y cosas así.

R: Sí se divierte uno más porque tiene contacto físico con las personas.

I: Sí, exactamente, OK um que tal ¿qué tipo de comida solías comer cuando eras niña?

R: Pues, en mi... en mi lugar de origen, de donde yo soy, es muy típico la co- las tortillas de harina,

I: OK.

R: porque soy del norte, y en exclusivamente en ese estado, es muy particular las tortillas de harina, entonces, eso era, algo que no tenía que faltar en mi casa,

I: Uh-huh.

R: las tortillas de harina ¿no? o sea, pero el resto no era comida extravagante así como, cara ni nada de eso, comida normal,

I: Sí.

R: arroz frijoles, huevos ¿si si yo explico?, nada más, nada fuera de lo... de lo básico

I: Sí, sí sí bien OK, la comida yo estuve en DF de vacaciones, y la comida estaba riquisima @

R: Sí eso es- es algo muy bueno de allá, allá- aquí te X T:13:26 comida orgánica, allá todo es orgánico.

I: Sí.

R: Por eso yo creo que uno estaba más… no crecía tanto, crecía así, no así, @

I: @ Sí, y ¿qué tal la comida que sueles comer aquí, en Roswell?

R: Me aburre… hoy en día, yo tengo hambre, y no sé qué comer, porque no se me antoja nada, siempre pan, siempre, comida, de a mentiras, diría yo, comida que no es... comida de mentiras porque, tú puedes pedir una ensalada, y nada de lo que está la ensalada es real,

I: Sí.

R: Aunque según tú estás comiendo sano… nada es real ya,

I: es cierto X… bueno, ah , ¿puedes describir tu casa cuando eras niña?

R: Mi casa era una casa pequeña,

I: OK.

R: De dos cuartos, compartía con mi hermano, era una casa con... un, una cochera…para un carro, y un patio pequeño en la parte de atrás donde... donde uno lavaba ropa y todo eso, no era una casa muy grande, era una casa pequeña.

T: 14.58

I: Um ahora vamos a hablar un poco de las escuelas, ¿puedes hablar de tu experiencia... de la escuela en México?

R: En cuestión de la escuela, am, cuando yo estaba en la escuela, me parecía que era muy buena… cuando yo vine para acá… ah, yo según tenía que entrar, a nueve, había escuchado que te regresan un año, pero a mí me dejaron en el noveno, pero cuando yo entré a la escuela, en cuestión de matemáticas, me decían que yo sabía más, de lo que debería saber para ese grado,

I: Sí, muy bien.

R: o sea que en México, la matemáticas era, muy buena,

I: Sí.

R: que lo que aquí me estaban enseñando en diez en once yo, ya lo sabía,

I: muy bien @

R: Sí… o sea ya me lo habían enseñado apenas se estaban poniendo al, parejo… de- a mí me parecía muy buena la i- eh la... sí la escuela ya, en mis tiempos, ahora escuchaban, otras cosas, ahora en día otras cosas pero en- cuando yo estaba, estaba bien.

I: OK, ¿había como... había como cliques en la escuela en Mexico?

R: ¿Como qué?

I: ¿Cliques como... grupos exclusivos, este grupo hace deportes este grupo hace tal cosa.

R: Eh había grupos de... por ejemplo la escolta, había grupos de, el bailable de típico de México, mi familia es bailarina de, de baile folklórico, y a mí me gustaba también, y yo como por ejemplo yo, estaba en ese grupo de baile, eh, pero no como clubs que hacen aquí no,

I: OK.

R: Que de ciencia y de eso, no, de eso no.

I: Perdón umm

T: 17:06 ((PAUSE IN RECORDING))

I: ok bueno, Vamos a hablar algo de... bueno, de la escuela, ahora vamos un poco de Atlanta, secciones de Atlanta mucho pero me puedes hablar de las diferencias notables entre Atlanta y Roswell?

R: ¿Atlanta y Roswell?

I: Sí.

R: Eh, la verdad nunca he vivido en Atlanta, en la ciudad de Atlanta nunca he vivido

I: OK

R: y las veces que he ido siempre está cerrado @, el que voy en las noches a caminar, eh, me parece una ciudad bonita, bastante grande, pero he visto bastante gente como homeless,

I: Sí.

R: he visto demasiado de eso y creo que eso no le da un buen aspecto a la... a la ciudad, no se ve en Roswell que es una ciudad más pequeña, porque se debería de ver en Atlanta que es la ciudad grande, ehm, no sé, no sé mucho de ese lugar.

T: 18.18

I: Está bien, um... crees que la la gente que vive en Roswell sienten que... que son de Roswell o son de Atlanta?

R: Okay, eh buena pregunta, eh…pues uno cuando dice "de dónde vienes" dice uno "vengo de Atlanta" por ser la m- la ciudad, principal ¿no?

I: Sí.

R: Pero el Roswe- Roswell también es una ciudad muy conocida ¿sabes? ah si tú les dices "¿de qué parte?" o te dicen, tú dices "Roswell" y la gente sabe que es Roswell, porque, es muy nombrada,

I: Sí.

R: También, es muy nombrada así como Alpharetta, pero, yo en lo personal digo "Roswell"

I: OK.

R: siempre.

I: OK bien, um crees crees que para para la gente recién llegada que vienen aquí a vivir, ¿qué tan fácil es que se..

R: ¿Adapten?

I: ¿se integren y se adapten a la vida?

R: Es, dependiendo la edad…porque cuando eres un niño, luego luego es- encuentras con quien hablar porque te meten a la escuela, y haces amigos, pero para una persona adulta como de... cuarenta, es difícil, tuve una experiencia con una persona muy allegada a nosotros…llegó, y lloraba todos los días, porque decía que no conocía a nadie, que eh nos- no conocía la ciudad, que para moverse necesitaba carro, por fuerza,

I: Uh-huh.

R: y lloraba todos los días, y era... estresante, para mí, verla llorar todo el tiempo,

I: Sí.

R: Una persona mayor, que ella yo creo que no podría, cambiar de lugar, donde vivir no se puede, si aún los adolescentes, como X T:20:25 escala porque, como tienen sus amigos ya…imagínate una persona mayor, yo creo que es difícil,

I: Sí, sería muy dificil.

R: es muy difícil, pero, hay personas que tienen la vida hecha o sea, llegan aquí, conociendo algún familiar, y llegan a dónde vivir, luego luego consiguen trabajo pero cuando llegas y no conoces a nadie, es peor aún,

I: Uh-huh.

R: Que no conoces a nadie no conoces... la ciudad, no sabes dónde vas a vivir no conoces nada,

I: Sí.

R: pero son los que más valoran, los que no batallan no valoran,

I: Ya

R: porque dicen “ah yo sufro” pero, tenías dónde llegar, tuviste techo tuviste comida tuviste todo, no te quejes,

I: Sí, claro.

R: Sí, bueno yo creo eso,

I: OK.

R: a mí no- yo no sufrí @

I: Sí.

R: No no no soy de muchos amigos, entonces llegué aquí y, soy una persona muy... sola, o sea, no tenía así como... muchas amistades, como era cristiana, de la iglesia pues no todos pensaban como yo,

I: Ya

R: Entonces, yo era la que se vestía diferente, la que hablaba diferente entonces, no tenía muchos amigos, ahora ya no @

I: @

R: ahora ya no @

T: 21.57

I: @, OK y bueno um... ¿cómo has visto que ha que ha cambiado Roswell desde que llegaste tú?

R: Eh, pues, no tengo mucho tiempo aquí, tengo seis, que seis años… no... he visto algunas construcciones nuevas, Roswell es una ciudad vieja, pero he visto por ejemplo aquí atrás, el lugar es muy lindo, lo que tumbaron ahí me parece muy bien, y me parece que hay algunos, eh, condominios o departamentos que también deberían de... de quitarlos y hacer cosas nuevas porque,

I: Sí.

R: porque así se vería mejor,

I: Uh-huh.

R: Los Riverwood, son viejos,

I: Uh-huh.

R: Y son feos, los departamentos cerca del cuatrocientos… yo creo que son los más... más malos, y eso no le da buen aspecto, a la ciudad como no,

I: Uh-huh.

R: X T: 23:00 así como los de acá atrás, eso está bien.

I: Uhm, ¿hay ciertas actividades? ¿Hay ciertas organizaciones acá que ayudan a la gente a X?

R: ¿Como económica?

I: Sí o como cualquier organización que ayude con trabajo, con el idioma, cualquier cosa.

R: Sí, hace poco estuvo unas personas aquí dando clases de inglés.

I: OK.

R: Aquí en, en la plaza.

I: OK.

R: Eh, he escuchado, económicamente, las personas que tienen hijos acá, les ayudan mucho el FCC, que le dicen, la ayuda del gobierno, yo he escuchado que mucha gente va ahí, y más porque mi trabajo, en mi otro trabajo yo veo que mucha gente, pide ir para allá.

I: OK.

R: Eh, en las iglesias creo que también les dan ayuda, les dan como despensas, o algo así.

I: Qué bien.

R: Sí yo creo que hay bastante ayuda, como para, decir que no tienes qué comer, siempre hay, siempre.

I: OK, uhm, ¿hay un cierto uhm, cantidad de tiempo que hay que vivir en Roswell, crees, para considerarse de Roswell? O tiene que ver con algo diferente..

R: mm… yo no s-, yo no lo diría, si vives aquí eres, de aquí ¿no? De Roswell o sea no que naciste aquí pero digamos, ¿dónde vives?, bueno "vengo de Roswell".

I: Sí, sí.

R: Sí, yo creo que en el momento que tú, empiezas a vivir aquí ya, si viajas ya no eres de otro lugar, sino que de ahí, o sea, no eres de ahí, por tus raíces ¿verdad?, pero vives ahí.

I: Sí.

R: Eres parte de, de esa ciudad, colaboras ahí, tus impuestos todo, tu trabajo todo está ahí.

I: Sí.

R: Yo creo que sí, una vez que c- que llegas acá.

I: OK, muy bien. Uhm, pues, ¿sabes algo de los uhm, los edificios históricos aquí?

R: No.

I: No?

R: No.

I: Uhm, ¿conoces alguna celebración o evento que tienen cada año aquí en Roswell?

R: Pues tampoco lo no-, lo normal, que es el 4 de julio, ¿no?, la independencia.

I: Sí.

R: Pero, yo no no, no estoy muy involucrada con las, cul- con la cultura americana me imagino que eso es más de americanos.

I: OK, sí.

R: Creo que lo que yo veo que celebran mucho es el 5 de mayo y eso es de México ¿no?, @.

I: Sí.

R: Pero, este, no, no estoy muy involucrada con la cultura americana.

I: Aja.

R: No mucho o sea, no sé de eventos, no sé, de lo que pasa afuera.

I: OK.

R: No veo ni el sol @. Desde aquí no lo veo @.

I: Uhm, OK, pues, ¿qué tal las compras? Uhm, ¿dónde vas de compras y como son las tiendas de aquí?

R: Eh, yo compro la mayoría del tiempo en Walmart… pero sí me he dado cuenta que Publix tiene mejor calidad, de verduras… pero las tiendas hispanas, dan más barato.

I: Sí.

R: Ehm, los productos de verdura, de vegetales y todo eso, los enlatados y todo eso no porque como esos son este traídos de, son importados o algo así, entonces eso te lo venden más caro, pero, los productos americanos, que no son verduras y vegetales y eso, en Walmart siempre los consigo…sí.

I: Uhm, ¿me puedes hablar un poco de cómo es diferente la vida aquí de la vida en Mexico?

R: Sí, es totalmente distinta.

I: Sí.

R: Allá, se batalla, para vivir, un salario mínimo, yo creo que, hace muchos años no estoy allá pero, en aquel tiempo el salario de mi mamá eran 600 pesos a la semana, me imagino que, pongamos que ahorita sacan 1.000, a la semana, lo que sacaría una persona aquí en dos, o con un salario mínimo, ganan allá, en dólares, 100, 100 pe-, 100 dólares, semanales, allá,…es muy poco, 100 dólares semanales en México, sin contar el tipo de cambio, porque si lo pones en el cambio y

eso es que te lo estén dando a 10, ahorita está a 15, o sea que serían como ¿80 dólares, a la semana? Es muy difícil, es di-, yo soy de México, y yo no conozco México, yo solo conozco el lugar de donde yo soy, pero yo no conozco México, nunca salí de vacaciones, porque, el, el, la economía allá no- o el trabajo que tengan tus papás, o el grado de estudios que llegues a tener y todo eso, cuentan mucho, si no tienen nada, es difícil.

I: Sí.

R: Pocas veces, te digo, salimos, pero no fue como unas vacaciones, para tener unas vacaciones hay que ahorrar, mucho tiempo, para comer mucha gente batalla, pero siempre hay algo, allá no, no se mueren de hambre, y cualquier cosa y las mamás inventan, comida, aquí no, aquí todo es fácil, la comida sobra en cantidades, sobra, o si te hace falta de comer vendes algo que tienes por ahí, y ya, con 100 dólares yo creo que, para dos personas compras, algo de comer y te dura por ahí una semana, hasta dos, pero no- es totalmente distinto, aquí la vida es más fácil, cualquiera puede tener un carro nuevo, un carro del año un carro de lujo, cualquiera puede tener un computador muy caro, un teléfono, el más caro, todos pueden, es fácil conseguir, el dinero, o llegar al precio del dinero, del, del producto, perdón, sí es fácil, cualquiera puede, aunque algunos los deben ¿verdad?

I: Ajá.

R: Pero, todos lo pueden tener, allá no.

I: OK.

R: Allá no, es difícil, muy difícil.

I: Uhm, ¿me puede decir X T: 30:17?

R: ¿Cómo?

I: ¿X?

R: Eh… pues, ¿qué te diré? Pues del estudio ya te lo, ya te dije.

I: Ajá.

R: En, cuestión, hay muchas clases también, gente que vive muy mal, con las casas de, de cartón, de lámina, construyen en donde no son sus tierras, eh, los, policías se toman demasiadas atribuciones hay mucha corrupción, probablemente aquí la haya también.

I: Sí.

R: Pero aquí no se ve, o sea, no está al alcance de los ojos de uno, o, yo nunca lo he visto, nunca, pues nunca he visto que alguien, diga, "ah soborné a un policía americano" o me di, me pararon, he, he visto, policía es buena gente, que le dan una oportunidad a la persona, pero, o sea no como, no se ve la corrupción, si es que hay, sí me supongo ¿no?, en todos los lugares hay pero, no se ve, allá sí son como más descarados como, la gente tampoco es que se una para, hacer algún cambio, o algo y es que ahora ya se calmó la violencia y todo eso, porque antes estaba,

terrible, pero sí también depende del, del círculo de personas con las que te, con las que te juntas ¿no?, si te juntas con maleantes, te va a ir mal.

I: Sí.

R: Sí, pero, sí hay mucha injusticia, mucha injusticia allá sí, acá es más, más, fácil la vida, los niños acá tienen la vida hecha, no batallan no sufren no- por dinero por comida por, carro por, celulares, no sufren, hasta un niño de ocho años, trae ya teléfono.

I: Es increíble, sí.

R: Sí, aquí no, aquí no sufren, allá sí, ¿qué niño va a traer un teléfono a los ocho años allá? No ¿y en mis tiempos? Olvídalo, olvídate eso no iba a suceder jamás, a menos de que fueras de dinero o una clase alta.

I: Ya.

R: Yo no tenía celular.

I: Uhm, ¿qué dirías que es el reto más grande de venir a Roswell a vivir?

R: ¿Como mi meta?

I: ¿Perdón?

R: ¿Como mi meta más grande? ¿O qué fue lo más difícil para mí?

I: O para ti o en general como, X, el reto más grande.

R: El idioma.

I: OK.

R: Totalmente, el idioma, si no hablas inglés no puedes conseguir tan fácil un buen trabajo.

I: Ajá.

R: Si no sabes inglés, el estudio también se complica, se complica también porque no entiendes nada no sabes nada o sea no vas a entender nada, es primordial yo creo llegar y aprender no que lo hablen todo el tiempo porque, es ilógico que yo me tope con una mexicana como yo y me empiece a hablar en inglés, si cuando las dos sabemos que, tenemos la vida hecha allá en México y llegar acá y que me hables en inglés pues no ¿verdad? Pero yo creo que primordial es el, el idioma, es el reto que todo el mundo se tiene que poner y cuando eres adulto, es muy complicado, muy complicado, no se te quedan las cosas o tu lengua está tan, tan, tan acostumbrada a tu idioma que para tratar de pronunciar otras palabras es, difícil… es difícil, es otro, como que, diferentes pronunciaciones y todo eso, es difícil. Yo creo que eso sería lo, el reto como que, más grande.

I: OK.

R: Porque, ¿qué otra cosa? Sí ya no hay más… estudiar.

I: Sí.

R: Y ya.

I: ¿Qué consejos le darías a una persona que viniera aquí a vivir, qué consejos como para adaptarse mejor?

R: Paciencia, paciencia porque si la pierdes, y dices, "no es que no puedo" o "no es como yo pensaba", sí puede ser que sí, pero tienes que tener paciencia como todo.

I: Sí.

R: O sea no puedes, llegar y decir, "ya tengo que ganar bien" o por dos días que vayas a la escuela, ya vas a decir, "ya hablo inglés", porque no, o sea tienes que tener paciencia y no desesperarte, ni la nostalgia es buena, la nostalgia tampoco es buena porque uno se deprime, empieza a llorar, "¿para qué me venía?" "¿Por qué estoy aquí?" "Estoy solo" no, eso no ayuda.

I: Sí.

R: Eso no ayuda… mucha paciencia, mucha paciencia, sí, bueno yo creo.

I: Sí.

R: Sí, y tampoco decaer, salir, si no conoces el lugar salte, camina, conoce la ciudad, conoce el lugar donde estás viviendo, porque ¿qué haces?, "no conozco a nadie y por eso no salgo", pues si no sales, menos vas a conocer, menos, así puedes salir y hablar con las personas y preguntar por un trabajo, o "mira yo acabo de llegar", esto tal, haces amigos o haces conocidos más que nada, ellos mismos te pueden decir, "oh pues yo, conozco tal persona"…y ya, saliendo y hablando s-, hablando se entiende la gente.

I: Sí.

R: Sí, consigues trabajo, eh, no sé alguien que te rente un cuarto, de todo, pero sí mucha paciencia y más para los que somos más grandes o los que son, adultos.

I: Sí.

R: Sí. Que no se desesperen…no es el sueño americano, eso es una mentira, no hay un sueño americano el sueño americano no, no existe, ese es el día a día, y eso es todo yo creo.

I: OK, Bueno, muchas gracias X, gracias.

END