A Tale of Two Coasts: The Impact of Language Contact on Subject Pronoun Expression and Number Marking in Two Varieties of Spanish in Nicaragua

by

Madeline Critchfield

(Under the Direction of Chad Howe)

Abstract

This dissertation investigates two morphosyntactic features in Mosquito Coast Spanish (MCS) and Nicaraguan Spanish (NS): subject pronoun expression (SPE) and variable number marking. Both NS and MCS are spoken in Nicaragua; however, NS is a monolingual, non-contact variety and MCS is a contact variety spoken by first language Miskitu speakers.

The analysis of SPE examines the distribution of explicit and implicit third person subject pronouns, while the analysis of variable number marking examines variation in third person plural subject-verb agreement. The study involved the statistical analysis of oral data, which was subjected to a mixed-effects logistic regression to determine what linguistic and social factors motivated the variation of each feature. The participants were 10 NS speakers and 10 MCS speakers for SPE, and 20 MCS speakers for variable number marking.

The results for SPE show NS speakers align with previous research: singular referents, less distinctive TMA, contexts of switch reference, and non-reflexive verbs all motivated the production of explicit pronouns. In the MCS data, less distinctive TMA and contexts of switch reference motivated the production of explicit pronouns, along with participants over the age of 30 and females. Additionally, MCS speakers showed a low rate of explicit pronouns overall (16.5%, compared to 23.9% in NS). The results for variable number marking showed lack of agreement occurred in 23.87% of the MCS data and was motivated by low phonic salience verbs, subject position (distantly preceding the verb and post posed), animate referents, participants over the age of 30, and males. For both features, MCS speakers used salience and L1 influence as strategies to com-

pensate for differences that exist between Spanish and Miskitu regarding SPE and number marking.

The findings provide evidence for incomplete acquisition of Spanish by MCS speakers. The low rate of explicit subject pronouns and variable subject-verb agreement are likely remnants of the rapid language shift to Spanish that took place in the past but have now become fixed features of MCS. Younger speakers, however, are approximating to distributions found in NS, which suggests that these features are currently changing in MCS.

INDEX WORDS: [Mosquito Coast Spanish, Nicaraguan Spanish, subject pronoun expression, variable number marking, Miskitu, morphosyntactic variation, language contact]

A Tale of Two Coasts: The Impact of Language Contact on Subject Pronoun Expression and Number Marking in Two Varieties of Spanish in Nicaragua

by

Madeline Critchfield

B.A., DePaul University, 2011
M.A., University of Georgia, 2017

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

2021

A Tale of Two Coasts: The Impact of Language Contact on Subject Pronoun Expression and Number Marking in Two Varieties of Spanish in Nicaragua

by

Madeline Critchfield

Major Professor:	Chad Howe
Committee:	Pilar Chamorro
	Timothy Gupton

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
August 2021

Acknowledgments

I would like to thank a variety of people who have contributed to the successful completion of this dissertation. First, I would like to extend my gratitude to my committee: Dr. Chad Howe, Dr. Pilar Chamorro, and Dr. Timothy Gupton. Your feedback and support during not only this process but throughout my graduate education at the University of Georgia has been invaluable.

I would also like to thank my colleagues, friends, and Ling Ladies who have encouraged me, guided me, and inspired me. I would not have gotten to this point without each of you, and I can truly say that I am a better person, both professionally and personally, because of your friendship.

Finally, I would like to thank my parents, sister, and husband for the many years of love, patience, and encouragement they have given me. I would not have been able to achieve my goals without their support, and that of countless other family and friends.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation will analyze two morphosyntactic features in Mosquito Coast Spanish (MCS): subject pronoun expression (SPE) and variable number marking. MCS is an understudied contact variety spoken along the Caribbean coast of Nicaragua by first-language Miskitu speakers. The Miskitu community experienced rapid language shift to Spanish, beginning in the middle of the twentieth century when Spanish-only education was implemented into the school system throughout the country (García, 1996, p. 99). Today, many Miskitu people are bilingual speakers of Miskitu and MCS, particularly in the coastal city of Puerto Cabezas, where the data collection for this study took place (Wilson Withe et al., 2012, pp. 12-14). There, Spanish is utilized in a number of public spaces, including schools and businesses; however, Miskitu continues to be passed down from generation to generation as an integral part of the community's ethnic identity (Wilson Withe et al., 2012, pp. 12-14).

By examining linguistic features in this contact situation, the findings will provide a more robust understanding of how morphosyntax is affected during natural language acquisition and group language shift. The findings will also highlight what strategies speakers utilize under the contact circumstances described for MCS, both past and present. The results will show how these strategies that originated from initial contact and acquisition are changing as speakers have increased exposure to standard language norms. In this study, MCS will be the primary variety under investigation and Nicaraguan Spanish (NS) will be used as a point of comparison as it is the variety spoken on the Pacific coast of the country by monolingual speakers of Spanish who have no contact with other languages. While NS is not necessarily the target language for MCS speakers, the analysis of these varieties together provides a foundation for understanding the linguistic characteristics that differentiate the two.

This dissertation will take a variationist sociolinguistic approach. The primary assumption of this perspective is that variation is a fundamental part of language and that speakers alternate between two or more linguistic choices depending on a variety of linguistic and social factors (Labov, 1969; Tagliamonte, 2012). This approach involves both a descriptive and interpretive component where the variation observed is described in detail and later interpreted based on the context. A statistical analysis of variation forms the foundation of variationist socioliguistics (Tagliamonte, 2012, pp. 2-3). The current dissertation will therefore both describe SPE and variable number marking, as well as provide a statistical analysis of their distribution in NS and MCS.

1.1 Subject Pronoun Expression

This dissertation will analyze SPE, which is one of the most widely studied topics in sociolinguistics literature in Spanish (Bentivoglio, 1987; Cameron, 1992; Silva-Corvalán, 1994, Solomon, 1998; Lapidus and Otheguy, 2005; Barnes, 2010; Otheguy and Zentella, 2012; Michnowicz, 2015, among others). This variation is analyzed via the rate of explicit versus implicit forms and is a morphosyntactic feature of language. Spanish is considered a null subject language, making both explicit and implicit forms possible. The rate at which speakers use explicit subjects varies greatly by region and is based on a complex set of discourse-pragmatic factors. Past studies have shown that the highest rate of explicit SPE occurs in lowland varieties of Spanish (de Olloqui de Montenegro, 1987; Cameron, 1993; Orozco and Guy 2008) and in varieties in contact with a non-Romance language (for English see Lapidus and Otheguy, 2005; Otheguy, Zentella, and Livert, 2007; Otheguy and Zentella, 2012; for Yucatec-Maya see Michnowicz, 2015). No published studies have been done on SPE in NS or MCS; however, this topic has been analyzed extensively in other contact and non-contact varieties.

This section of the dissertation will add to the body of work that has been done on this topic and will provide additional information about how SPE is affected when Spanish comes into contact with an indigenous language. The analysis for SPE will be a variationist statistical analysis of linguistic and social factors that predict the production of explicit subject pronouns. The results will indicate which variables are significant predictors of SPE in both varieties. The findings will also show how the use of subject pronouns in MCS compares to both NS and other contact varieties of Spanish.

1.2 Variable Number Marking

Variable number marking on verb morphology is a phenomenon that has not been attested in many varieties of Spanish, with the exception of Caribbean Spanish (Hochberg, 1986a,b; Cameron, 1993, 1996; Alba, 2004; Foote and Bock, 2012; Guy, 2017), Spanish in the U.S. (Lipski, 2008c), and Afro-Bolivian Spanish (Lipski, 2008b; Sessarego, 2011, 2012). The current dissertation will add to this body of literature by analyzing variable number marking in MCS, which has been observed in the context of 3pl. referents. This variation in number marking results in lack of agreement between subject and verb (1.1 b), compared to (1.1 a) where agreement occurs.

(1.1) (a) *En cuanto al estilo de vida, ellos viv-ían*
in regard to style of life they live-PST.IPFV.3PL
diferente
differently
'In regard to lifestyle, they lived differently'

(b) *Nuestros ancestros antes viv-ía de la*
our ancestors before live-PST.IPFV.3SG from the
naturaleza
nature
'Our ancestors used to live from nature'

Variable number marking in MCS has only been observed with third person verb forms, specifically with plural referents; therefore, this dissertation will only include an analysis of this context. Previous research on this phenomenon has primarily involved second person final /s/ deletion (see Hochberg, 1986a,b; Cameron, 1993, 1996 for Puerto Rican Spanish and Foote and Bock, 2012 for Dominican Spanish). Research that looks at third person forms involve final /n/ deletion (see Foote and Bock, 2012 for Dominican Spanish) or the use of 3sg. for various grammatical persons/numbers (see Lipski, 1993, 2008c for Spanish in the U.S. and Lipski 2008a,b; Sessarego, 2011, 2012 for Afro-Bolivian Spanish). The phenomenon that I will examine in MCS is distinct because it does not involve phonological weakening but the use of the 3sg. verb form in place of the 3pl. verb form (1.2and 1.3).

(1.2) *Ellos le dij-eron que no les llev-ara*
they him tell-PST.3PL that no them take-PST.SBJV.3SG
'They told him not to take them'

(1.3) *Entonces los pajarito-s dij-o que*
then the.PL bird-PL say-PST.3SG that
hab-ia un techo
there-was-PST.3SG a roof
'Then the birds said that there was a roof'

Variable number marking has been studied in depth in Brazilian Portuguese (BP) (Scherre, 1998; Naro and Scherre, 2000; Scherre and Naro 2001, 2014; Guy, 2005; Lucchesi et al., 2009; Brandão and Vieira, 2012; Mendes and Oushiro, 2015), where it is pervasive across all social statuses. Due to breadth of literature that exists for BP, the current analysis will draw heavily on this research, particularly in regard to the methodology.

This dissertation will only analyze variable number marking in MCS, as NS does not display any evidence of this phenomenon. The variationist analysis will include a variety of linguistic and social factors, with the results indicating which are significant predictors in the non-agreement between 3pl. subjects and their corresponding verbs. The findings will then be compared to BP and general trends observed in varieties of Spanish where similar number marking phenomena are found.

1.3 Incomplete Language Acquisition

This dissertation will investigate how contact and language acquisition have affected SPE and number marking in MCS. This variety offers important information about the effects of contact on morphosyntactic features in Spanish, as Miskitu speakers first acquired Spanish due to political and social pressure (García, 1996; Escobar, 2013; Wilson Withe et al., 2012), which often encourages language shift (Winford, 2003, pp. 257-258). In the case of the Miskitu community, this shift took place fairly quickly and has had a profound impact on the development of MCS. In addition, Miskitu speakers were geographically isolated from NS-speaking populations located on the Pacific coast, leading to the incomplete acquisition of certain linguistic features during initial language shift. The distribution of subject pronouns and the presence of variable number marking found in MCS today are evidence of this past incomplete acquisition, and are now characteristics of this variety of Spanish. The typological differences that exist in Miskitu and Spanish regarding SPE and number marking on verbs made them susceptible to processing difficulties for learners, particularly under the contact circumstances described for MCS.

The final part of the dissertation will include a discussion of how the results of the data analyses support the claim of past incomplete language acquisition

in MCS, a feature present in many bilingual grammars due to transfer or fossilization (Montrul, 2006, p. 336). Learners utilize a variety of strategies and processes, including simplification and L1 transfer, to compensate for aspects of the target language not acquired in their entirety (Winford, 2003, p. 224). These learner strategies involved in the acquisition process will be examined in order to identify a common explanation for the findings of both linguistic topics under investigation.

1.4 Dissertation Contributions

This dissertation will examine two morphosyntactic features, SPE being a high-frequency variable with a large body of previous research, and variable number marking being a low-frequency variable with little research in varieties of Spanish. These topics will be discussed for two varieties of Spanish spoken within the country of Nicaragua. MCS is the primary variety under investigation and is considered to be a contact variety, where speakers are bilingual in Miskitu and Spanish. NS will be used as a point of comparison in order to examine the effects of acquisition and language contact, as it is a monolingual, non-contact variety of Spanish.

Very little literature exists about linguistic variation in Central American varieties of Spanish, particularly those found in Nicaragua. The previous studies that do exist on the Spanish spoken in Nicaragua (Lipski, 1984, 1985, Lipski, 1996; Chappell, 2013, 2014, 2015a, 2015b, 2016) provide a description of only a few notable features, primarily phonological phenomena. The current dissertation seeks not only to compensate for this lack of representation in the Hispanic Linguistics literature by providing a description of additional features found in these two Nicaraguan varieties, but more broadly aims to address the effects of contact and bilingualism on morphosyntactic elements in Spanish. This analysis will then lead to a discussion about what explanations best account for the results for both of these features in the two varieties in question and how this relates to sociohistorical context and language contact in Nicaragua.

The individual analyses presented for MCS and NS will provide additional insight into what motivates SPE and variable number marking. Their analysis together will reveal what strategies speakers used in the face of incomplete language acquisition in a contact situation that involves two typologically different languages. While research on linguistic features found in contact varieties of Spanish is growing, studies that analyze bilingual speakers of Spanish and an indigenous language are limited. MCS is particularly understudied (Chappell, 2016; Critchfield and Lívio, forthcoming). The findings will provide impor-

tant insight into the effects of language contact between typologically different languages as well as the results of rapid language shift and incomplete language acquisition.

1.5 Outline

In this introduction, I presented preliminary information on the two morphosyntactic features I will analyze: variable number marking and subject pronoun expression. I also gave a brief introduction to the two varieties of Spanish I will be investigating. Finally, I discussed the purpose of my study and what the results will contribute in regard to the field of contact linguistics and the study of language contact, bilingualism, and language acquisition.

In chapter 2, I will provide the historical background of MCS and NS as it pertains to contact. I will also include an overview of the structure of Miskitu, specifically number marking on verb morphology and pronominal subjects.

In chapter 3, I will give an overview of the relevant previous literature on SPE and variation in number marking on verbs in Spanish. I will include a discussion of the findings for contact versus non-contact varieties in regard to both topics. I will also include a section on variable number marking in BP since there is a large body of variationist literature about the phenomenon in this language. Additionally, I will give an overview of the explanations I plan to discuss related to speaker strategies amid incomplete language acquisition.

In chapter 4, I will describe my participants and methodology in detail, including the tasks and how I coded the data.

In chapter 5, I will provide the results for the quantitative analysis of subject pronoun expression for both NS and MCS and in chapter 6, I will provide the results for the quantitative analysis of variable number marking for both varieties.

In Chapter 7, I will discuss explanations regarding contact and incomplete language acquisition that support the results of both statistical analyses.

In chapter 8, I will give a conclusion of the findings for the two analyses and how they contribute to the broader study of Spanish linguistics and language contact.

Chapter 2

Sociohistorical Context of Mosquito Coast Spanish

The unique contact situation of MCS, both past and present, provides important information about the effects of language contact and bilingualism on morphosyntactic features in Spanish. MCS is a contact variety of Spanish that resulted from rapid group language shift during the second half of the twentieth century (García 1996; Meringer 2014). Speakers were geographically isolated from NS speakers but were required to learn Spanish in school. Both the rapid shift and lack of widespread input from target language speakers resulted in a contact situation where the Miskitu community had to acquire a language with limited input, which is reflected in the distribution of morphosyntactic features discussed in this dissertation.

2.1 Linguistic Background

In this section, I will discuss the background of Mosquito Coast Spanish, including the impact of natural geography on language distribution in Nicaragua, as well as the current linguistic demography of the country.

2.1.1 Geographic and Linguistic Divisions

Nicaragua's geography has had an important impact on the languages spoken in different regions of the country. A central mountain range, two coastlines, and an abundance of natural resources has resulted in contact between English, Spanish, and various indigenous languages. This contact has had particularly

interesting results along the Caribbean coast of the country, where the contact variety of MCS is spoken by first language Miskitu speakers.

The whole of Central America contains a natural division via a mountain range that stretches from southern Mexico to Costa Rica (Floyd, 1967, pp. 1-16). These mountains create a geographic boundary separating the north and the south. Along the south, the climate allows for successful agriculture, which throughout history has resulted in more development, with indigenous groups initially settling in areas with greater agricultural potential.

Figure 2.1: Map of geographic separation of Nicaragua's coasts (INETER, 2020)

In the case of Nicaragua, this mountain range prevented a high level of movement between the Pacific and Atlantic coasts (Figure 2.1). During the initial migration of indigenous groups, between 1000 and 1300, the Uto-Aztecans came down from the north and settled to the south of the mountains along the Pacific Coast. Around the same time, Sumu indians from the Chibchan family came up from the south, specifically from Colombia and Panama, to settle to the north along the Atlantic coast (Floyd, 1967, pp. 3-4). Due to the separation of the coasts by this natural geographic barrier, a unique political and linguistic history developed on each side. The people along the Pacific were generally peaceful and acclimated to the arrival of different indigenous groups, as well as the arrival of the Spaniards. The Sumus on the other hand, were known for being warriors and resisted European contact (Floyd, 1967, p. 5). The most powerful of the Sumu tribes, the Sambo-Miskitos, settled along what is now the coast of Nicaragua and Honduras where they established what has been referred to as a kingdom, consisting of over 10,000 subjects (Floyd, 1967, p. 5). The descendants of this group are called Miskitu and consist of both monolingual speakers of the Miskitu language, as well as bilingual speakers of Miskitu and Mosquito Coast Spanish.

Figure 2.2: Map of the departments in Nicaragua (INETER, 2020)

2.1.2 Linguistic Demographics of Nicaragua

Today Spanish is considered to be the official language of Nicaragua and is used as the majority language by the people living on the Pacific coast of the country (Catter, 2011, p. 721). In contrast, the Caribbean coast is much more linguistically diverse. Apart from Spanish, an English-based creole, as well as various indigenous languages, including Miskitu, are spoken throughout the two departments that comprise the region: Región Autónoma Atlántico Norte (RAAN) and Región Autónoma Atlántico Sur (RAAS) (Figure 2.2 and Figure 2.3).

The Miskitu language is classified as part of the Misumalpan language family and contains three different dialects that are divided by geographic region. *Tawira* is the dialect spoken in the city of Puerto Cabezas, while *Wangki* is the dialect spoken along the Coco River, on the border of Nicaragua and Honduras. Finally, *Mam* is the dialect spoken by the Miskitu people in Honduras. The differences between the three dialects are mainly related to vocabulary with the grammar being the same throughout all Miskitu speaking areas (Heath, 1927, p. 3).

In Nicaragua, the Miskitu-speaking population numbers around 120,000 people[1], with 91.22% of speakers located on the Caribbean Coast and 85.09% in the RAAN (Wilson Withe et al., 2012, p. 12). Within the RAAN, 32.7% of the people identify as Miskitu, particularly in the northern and coastal municipalities of Waspam, Puerto Cabezas, and Prinzapolka, where 93.3%, 70.4%, and 54.8% of the population is Miskitu respectively (Wilson Withe et al., 2012, pp. 12-13).

[1] While Spanish is also spoken within the Miskitu community, the number of MCS speakers is unknown.

Figure 2.3: Map of Nicaragua language families (Ethnologue, 2020)

The linguistic diversity found along the Caribbean coast stems from both historic and more recent contact, the earliest of which was with speakers of English, first British and later Americans. Spanish has also had an impact on the region's linguistic make-up. These cases of contact have had a lasting effect on the linguistic identity of this region in Nicaragua, both in policy and language practice, to be discussed more in detail in the following section. The geography of Nicaragua has played an important role in determining what kind of contact took place between different groups in the region, ultimately, influencing the languages spoken.

2.2 Language Contact and Group Language Shift

The sociohistorical background of MCS is crucial to consider when distinguishing this variety from monolingual Nicaraguan Spanish (NS). The history of contact, along with the current demographics of its speakers make MCS unique within the context of Nicaragua. The contact situation where MCS is spoken, both in the past and present, provides important information on the effects of contact on morphosyntactic features in Spanish.

2.2.1 Early Contact

Nicaragua is a country bordered by two coasts, the Pacific and the Atlantic, reflected in the blue bands that are found on the national flag (Figure 2.4). Not only did the central mountain range impact the development of two cultural landscapes in Nicaragua, the possibility of access via two oceans resulted in migration of many outside groups to the country throughout its history.

Figure 2.4: Nicaraguan national flag (MINED, 2019)

The Spanish first arrived on the Pacific coast of Nicaragua in the sixteenth century. The explorers Gil González Dávila and Andrés Niño landed in present-day Nicaragua in 1523. They made contact with the Nicarindos living on the south end of the coast and later founded the city of Granada in 1524 and the city of León in 1526 (Floyd, 1967, p. 5). The conquest of western Nicaragua, and Central America as a whole, was rapid and met with little resistance.

The history of contact on the Caribbean coast was quite different. Historically the Miskitu community has had significant influence from English speakers, unlike the indian groups on the Pacific. The first, of British origin, arrived on the Atlantic coast of Nicaragua in the 1630s. English language contact continued to grow due to trade between Puritan settlers of Old Providence, an island off the coast founded by British settlers in 1629, and the Atlantic coast (Rowland, 1935, p. 298). The connection between Old Providence is considered to be the source of the first African slaves to come to Nicaragua. Slaves from this territory fled to the mainland when the Spanish conquered the Puritans in 1641 (Michaelis et al., 2013, p. 115). The first slaves that arrived were shipwrecked on the Atlantic coast and settled in the British-occupied territory of Nicaragua. Another ethnic group, today known as the Miskitu people, also inhabited this area of the Atlantic coast. It is believed that the present-day Miskitu ethnic group originated from the intermarriage of British settlers, the shipwrecked African slaves, and Sumu indians, an indigenous group already present in Nicaragua (Craig, 1985, p. 381).

Consistent contact between the Caribbean coast and English-speakers occurred through the end of the seventeenth century between the Miskitus and British pirates (Helms, 1971, p. 14). The two groups exchanged merchandise and formed an alliance in order to fight against Spaniard colonizers in the western part of Nicaragua. Due to this relationship, the Miskitu people became the most politically powerful indigenous group on the Caribbean coast. In addition, the Miskitus benefited from an agreement with the English where they would capture members of other indigenous groups and sell them into

the British slave trade (Dennis, 2004, pp. 24-25). This original contact, mutually beneficial for both parties involved, established the foundation for a long history of diplomatic relations with English-speaking groups. Due to this initially positive interaction, particularly in the economic sense, groups along the Caribbean coast developed strong and lasting ties to the English language and its speakers. The relationship between the two groups stems from a political alliance where both parties seemed to benefit from the interactions.

The British offensive against the Spanish in Nicaragua continued up until the end of the eighteenth century, when finally the English were driven from the Caribbean coast. They were weakened by their participation in the Napoleonic Wars taking place in Europe and as a result lost some of their control over land in the Caribbean (Dozier, 1985, p. 28). The United States benefited from their absence and took advantage of greater trade in the region.

2.2.2 British-U.S. Rivalry

The beginning half of the nineteenth century was a conflicting time in regard to diplomatic relations in Latin America, with many countries pushing for revolution and independence from their colonizers (Spain in the case of Nicaragua). While the British had previously been at odds with the Spanish, they were now allies fighting against the Napoleon dynasty in Europe. The British therefore were in a difficult position. They wanted to maintain their trade with Latin America and feared the United States would edge them out; however, they could not openly defy the Spanish, even though they did not agree with their rigid trade policies (Dozier, 1985, p. 31). While neither the U.S. nor England officially supported the wars for independence at first, they continued to trade and establish commercial relations along the Caribbean coast of Nicaragua.

Starting in the early part of the nineteenth century, the British began to regain influence and control over ports along the Caribbean coast but eventually their presence dissipated with the implementation of the Monroe Doctrine and the United State's staunch policy against European interventionism in the Western Hemisphere (Dozier, 1985, p. 57, 69). Slowly, the United States began to take a greater economic and political interest in the region, contradictory to their policy in the Doctrine, which was only directed to outside nations.

2.2.3 Growing U.S. Influence

Nicaragua is a land rich in agricultural exports and natural resources including coffee, beef, gold, sugar, peanuts, shrimp, lobster, and tobacco (CIA, 2019). This fact has not gone unnoticed by outsiders. Toward the end of the nineteenth

century and into the twentieth century, English language contact continued with the arrival of people from the United States looking to establish mills and mining companies on the Caribbean coast. Helms (1971, pp. 27-32) terms this part of the region's history as the "company period". After officially recognizing the new Latin American republics as legitimate governments in 1822, the U.S. sought to expand its control and influence in the region, distancing itself from other European powers and exploiting the natural resources the region had to offer (Dozier, 1985, p. 40). While the "company period" brought about the exploitation of natural resources found on the Caribbean coast, the jobs in the mills and mines were well paying jobs, and as a result, Americans were looked upon with favor in the region (Dennis, 2004, p. 32).

In addition to commercial development on the Caribbean coast, the prospect of religious proselytizing brought missionaries from the United States to this region in Nicaragua. These English-speaking missionaries from the Moravian church also arrived around the same time as the merchants and developers. While not officially affiliated with the U.S. government, these American missionaries contributed to the relationship developing between Nicaragua and the United States at the time. The Moravians, English-speaking missionaries, also played an important role in the Miskitu language, arriving on the Atlantic coast of Nicaragua in the mid-nineteenth century and having an enormous impact on the religious demography of the region. The reception of the Moravian church was quite positive, particularly because the missionaries learned Miskitu and chose to communicate to the community in its first language. The incorporation of Miskitu into predication contributed to the strong relationship between the church and the community, to the point where church-goers actually preferred to hear the religious teachings in English due to the social power it held (García, 1996, p. 78).

In regard to language, the Moravians, most notably the Reverend G. R. Heath, played a particularly integral role in the Miskitu language, creating its first writing system and also publishing the first grammar books about the language in the early twentieth century (Dennis, 2004, p. 29). The participation of English-speakers from the United States in the documentation of the Miskitu language has many consequences, with biases being incorporated into the language system. Dennis (2004, p. 29) notes that a problematic aspect of the initiative to document the Miskitu language and its grammar system by missionaries was that they often tried to fit Miskitu, a typologically very different language from English or Spanish, into a Latin-based paradigm, which is not linguistically accurate. However, the contact the church had along the coast has

strengthened the positive association with English speakers along Nicaragua's Atlantic coast.

2.2.4 Post-Revolution Spanish Contact

Greater political involvement in the Caribbean coast on the part of the United States began during the Cold War period, when many Latin American nations were experiencing communist or Marxist based revolutions. Nicaragua was no exception to this. The Sandinistas overthrew the national government in the late 1970s and took control of the country, much to the dismay of the United States. As a result, the U.S. became involved in funding a counterrevolution in Honduras and northern Nicaragua to overthrow the newly established regime. With many Miskitu speakers participating in this counter offensive, the participation and support by the Americans only strengthened the region's ties to the English language and reaffirmed an old alliance between the Caribbean coast and English-speaking nations (Dennis, 2004, pp. 33-36).

Due to the constant contact with foreigners, the Miskitu community has demonstrated the ability to adapt to other cultures. According to Helms (1971, pp. 3-4), this adaptation has had a considerable impact on the Miskitu. Although Spanish-speaking culture is currently closer geographically speaking, the English-speaking culture has held more influence and prestige historically (Helms, 1971, p. 221). This changed in the middle of the twentieth century with the speech community transitioning to bilingual in Miskitu and Spanish (García, 1996, p. 99). Starting in the 1950s, a UNESCO sponsored initiative, called the Río Coco Pilot Project for Basic Education, was enacted with the goal of hispanicizing the Miskitu people and quelling ethnic heterogeneity (Meringer, 2014, pp. 205-207). As part of this project, Spanish-only education was introduced in Miskitu-speaking communities along the Caribbean coast (García, 1996, p. 99; Meringer, 2014, p. 208). When the Sandinistas came to power after the Revolution against the national government in the 1970s, a more constant and robust relationship developed between the Miskitu people on the Atlantic coast of Nicaragua and the Spanish speakers from the Pacific. After the Sandinistas took control, a national literacy program was established that promoted the use of the Spanish language (Escobar, 2013, p. 732). The Sandinistas wanted to unify the country under a collective identity (García, 1996, p. 104). Prior to this campaign, there was little interaction between the inhabitants of the Pacific coast and the Miskitu community, with the contact between indigenous groups and the national government being "historically, minimal" (*The Miskitos in Nicaragua, 1981-1984.* 1984, p. 3). The Miskitu community grew to distrust the mestizo people living on the opposite side of the country and referred to

them as "Spaniards", or in Spanish, *colonos* (*The Miskitos in Nicaragua, 1981-1984*. 1984, p. 3).

One result of this forced linguistic integration was the overall community shift toward Spanish as the dominant language in all regions of the country. Today along the Caribbean coast, Miskitu and Spanish are spoken on an individual and community level; however, first language Miskitu speakers have slowly begun to shift to Spanish. According to Wilson Withe et al. (2012, pp. 12-14), although there is still a strong intergenerational transmission of Miskitu, Spanish is becoming the majority language among the Miskitu community, particularly in the city of Puerto Cabezas, the capital of the RAAN and the epicenter of the MCS speakers in Nicaragua.

This political relation with the Spanish-speaking population from the Pacific coast is in striking contrast with British and American interactions. The Miskitu people already had ill-feelings toward the Spanish colonizers from early colonial contact, which was exacerbated by a form of linguistic colonization attempted by the Sandinistas.

2.2.5 A Reflection of English Language Contact

A look at the Miskitu language provides intriguing evidence of the positive diplomatic and commercial relations this indigenous group has experienced with speakers of English. The early and constant contact with English-speaking populations, both British and American, has had a lasting effect on the Miskitu language. This can be seen in the countless lexical borrowings from English, including the frequent use of English nouns, as well as auxiliary verb constructions. These lexical borrowings fall into two different categories according to Winford (2003, p. 45): load words and loan shifts. The lexical classes where English is often found include nouns related to clothing, housing, and religious terms, among others (Table 2.1).

Table 2.1: English Borrowings in Miskitu

Miskitu Word	English Origin
windar 'window'	window
glas 'mirror'	glass
tibil 'table'	table
traus 'pants'	trousers
sus 'shoes'	shoes
holi 'holy'	holy
Jisus 'Jesus'	Jesus
Baibil 'Bible'	Bible

In many of the cases, the English word has been adapted to the Miskitu phonological system, which includes fewer vowels. In Miskitu there are only three vowel sounds: /i/, /u/, and /a/ (Salamanca, 1988, p. 42). The majority of the nouns that come from English can be characterized as loanwords according to Winford's classification, mainly "pure" loans where there is a slight degree of phonemic substitution according to the available vowel sounds and orthographic rules. For example, the Miskitu writing system does not include the letter 'y' so this sound is often replaced by 'i'. However, words like *glas* represent cases of loanshifts where the semantic meaning has been extended from the material 'glass' in English to mean 'something made of glass' in Miskitu, in this case a mirror.

These same adaptations can be seen in English verb borrowings. However, in Miskitu, the English is incorporated in a compound verb structure where the English lexical borrowing is placed before one of three Miskitu verbs: *kaia* ('to be'), *munaia* ('to do'), or *takaia* ('to become') to form one action (Table 2.2).

Table 2.2: English Verb Borrowings in Miskitu

Miskitu Verb	English Verb Origin
laik kaia 'to like'	like
nu kaia 'to know'	know
shar munaia 'to share'	share
alp munaia 'to help'	help
stadi takaia 'to study'	study

In the borrowings above, the verb examples seem to represent a case of derivational blends where a foreign root is used with native inflexion. However,

the borrowing from English, the first word in each compound structure, is an example of a "pure" loan, similar to the nouns discussed above.

The examples presented here represent a variety of lexical borrowings found in Miskitu. Based on the historical contact in political, religious, and commercial settings, words that would have been used in these contexts were borrowed from English. In most cases, the Miskitu word has made small changes to accommodate the different phonological inventory and writing system, resulting in a "pure" loan. In other cases the semantic meaning of the English word has been extended or a compound verb structure combines a foreign root with native inflexional morphology. In both cases, the contact has led to a long lasting linguistic influence in the region, arguably due to the profound, and positively received contact between the Miskitu community and English-speaking nations throughout history.

2.2.6 Spanish Language Contact and Bilingualism

Unlike British and U.S. contact, interactions between Spanish speakers and populations on the Caribbean coast were more recent and unwanted on the part of the Miskitu community (Helms, 1971; García, 1996; Dennis, 2004). While extensive relations with English speakers has led to significant linguistic influence, the case of Spanish is quite different. Instead of being integrated into local languages via lexical borrowing, Spanish is a completely separate entity that while spoken by members of the Miskitu community, has not been integrated into the language.

In Latin America, one of the most common types of language contact is that which occurs between Spanish and indigenous languages (Escobar, 2013). First language speakers of indigenous languages often learn Spanish as a second language and are therefore bilingual, as is the case with Miskitu speakers on the Caribbean coast of Nicaragua (Wilson Withe et al., 2012) . The result of this contact between Miskitu and Spanish is diglossia within the community.

Diglossia is a phenomenon that occurs when two languages or language varieties are used "in complementary distribution across different situations" (Winford, 2003, p. 112). Winford (2003, p. 112) states that the low language is the language usually learned as a first language and is utilized in more private or informal domains. In contrast, the high language is the language usually learned via formal education and is employed in official or more public domains. In the case of Latin America, the indigenous language is usually considered the low language and Spanish the high language.

Diglossia can be divided into situations of stable diglossia and situations of unstable diglossia. Stable situations are those where the low language and

the high language remain used in their respective domains. Unstable diglossia occurs when the division between domains breaks down and the low language begins to permeate public or official domains.

As mentioned above in Latin America, language contact occurs between Spanish and indigenous languages in many countries. In Nicaragua, this type of contact can be found on the Caribbean coast between Spanish and Miskitu. In the RAAN, these two languages are spoken on an individual and community level. According to a manual published by UNESCO, in the urban centers of this department Spanish is used in government institutions and national media sources (Wilson Withe et al., 2012, pp. 12-14). Also, in Puerto Cabezas, the capital city of the RAAN, many businesses are run by monolingual Spanish speakers, which increases the use of Spanish in business transactions. Government and business are both considered official domains where the high language would be used, in this case Spanish.

Spanish is used in many official domains in the RAAN, providing evidence for stable diglossia. However, Miskitu and Spanish on the Caribbean coast more accurately represent a case of unstable diglossia due to the presence of Miskitu in traditionally high domains. One domain where diglossia has broken down is in education, where there is a bilingual education program in place within the public education system. This is a clear example where the low language has permeated into the high domain.

In addition to the domain of education, politics offers interesting insight into the cases of diglossia along the Caribbean coast. In the RAAN there are several radio stations in Miskitu and a local Miskitu political party that promotes the use of the language and the rights of the people in the community (Wilson Withe et al., 2012, pp. 12-14). Due to the history of political conflict with the monolingual Spanish speaking community, this political party, called YATAMA, has gained immense following in the RAAN and is the source of extreme pro-Miskitu ideology (Butler, 2000, pp. 37-40). This may be due in part to the size of Miskitu speaking community in the RAAN. While Spanish is considered the official language of Nicaragua, in the capital city of Puerto Cabezas, 70.4% of the population consider themselves Miskitu (Wilson Withe et al., 2012, pp. 12-13).

The state of bilingualism in the RAAN between Miskitu and Spanish provides an interesting case of diglossia. Due to a history of strong cultural pride, Miskitu, the low language, has been able to spread into high domains, such as education and government. When compared to the case of English, the result of Spanish contact has been quite different. Whereas, English was incorporated into Miskitu language and community identity, Spanish has been intentionally

blocked. Instead of a melding of both languages, speakers chose to use one or the other, often in different social contexts. This reflects the conflicting diplomatic relations with Spanish speakers on the Caribbean coast of Nicaragua.

Apart from diglossia, another interesting result of the forced linguistic integration attempted by the Sandinistas after the Revolution is the overall community shift toward Spanish as the dominant language in the region. As mentioned previously, Miskitu and Spanish are spoken on an individual and community level in Puerto Cabezas; with an overall shift toward Spanish. In this region of Nicaragua, while the majority of the community identifies as Miskitu, Spanish has a strong presence in public spheres. Children learn and use Spanish as a second language in the domain of education as well as professional life, with the majority of business transactions being conducted in Spanish (Wilson Withe et al., 2012, pp. 12-14). It is interesting to note, that while English was incorporated into the language, it never replaced Miskitu completely, while Spanish, as a result of diglossia, seems to be doing this through contact with speakers from the Caribbean coast in a much shorter amount of time.

MCS represents a situation, in which adult speakers acquired a second language due to social or political pressure (García, 1996; Wilson Withe et al., 2012; Escobar, 2013), a common social factor that encourages language shift (Winford, 2003, pp. 257-258). This nativization process, in which the speech community transitioned from being monolingual Miskitu speakers to bilingual Miskitu-Spanish speakers in the second half of the twentieth century via group language shift, has played an important role in the development of MCS, due to the fact that speakers were isolated to some degree from the target language in question. Prior to the Counterrevolution, Puerto Cabezas (the city where the data was collected for this dissertation) communicated with the capital only via telegraph as there were no television or telephone services. In addition, the first highway connecting the Pacific and Atlantic coasts was finished in 1982 (*The Miskitos in Nicaragua, 1981-1984*. 1984, p. 4). These barriers, along with the sociohistorical differences between MCS speakers and NS speakers, is important to keep in mind when analyzing the two linguistic features discussed in this dissertation.

2.3 Miskitu Language Structure

In this section I will present relevant linguistic information that describes various aspects of the Miskitu language, including pronominal subjects, verb inflection morphology, and subject expression.

2.3.1 General Information on Miskitu Phonology, Syntax, and Morphology

Miskitu contains 14 consonants and three vowels: /a/, /i/, and /u/, all of which can be short or long. The language permits a variety of syllable structures (2.3).

Table 2.3: Miskitu Syllable Structure

Syllable	Example
CCVCC	*praks* 'close'
CCVC	*kruskaia* 'to fold'
CCV	*plikaia* 'to search'
CV	*nakra* 'eye'
CVC	*nasma* 'honey'
CVCC	*kiks* 'laugh'
VCC	*arkbaia* 'to gnaw'
VC	*asmala* 'nail'
V	*apaia* 'to lay eggs'

(adapted from Badlato, 2000; Salamanca, 1988)

Miskitu is a head-final, SOV language (2.1). Like Spanish, Miskitu would be classified as a synthetic, fusional morphological language based on the presence of multiple morphemes in each word, including inflexional and derivational morphemes (Salamanca, 1988, p. 104).

(2.1) *yang* *Maria* *ra* *Juan* *buk-ka* *yab-ri*
I Maria to Juan book-POSS.3SG give-PST.3SG
'I gave Juan's book to Maria'

2.3.2 Pronominal System and Verbal Inflections

Looking at the grammatical properties of the Miskitu language in regard to number marking and subject pronouns is important in order to determine structural differences that exist between Miskitu and Spanish. The pronominal system of Miskitu, in subject position, consists of first, second, and third person forms in singular and plural (Table 2.4). In order to form plural subject pronouns, the plural marker *nani* is added after the singular pronoun. This differs slightly from Spanish, which has different forms for all persons and numbers that are morphologically unrelated, with the exception of third person (i.e *él/ellos* 'he/they' and *ella/ellas* 'she/they').

Table 2.4: Nominal pronouns and present indicative verb paradigm of *aiwanaia* ('to sing') in Miskitu

Person/Number	Singular	Plural	Present Indicative
First person	*yang*	*yang nani*	*aiwanisna* 'I/we sing'
Second person	*man*	*man nani*	*aiwanisma* 'you/you all sing'
Third person	*witin*	*witin nani*	*aiwanisa* 'he/she/they sing'

(adapted from Salamanca, 1988, p. 115, p. 250)

In regard to the verb systems in Miskitu and Spanish, both mark person in the morphology of the verb; however, an important difference between the two languages involves the marking of number. In Miskitu, number is not marked in the verb morphology, unlike in Spanish (Table 2.5) (Salamanca, 1988, p. 252).

Table 2.5: Nominal pronouns and present indicative verb paradigm of *cantar* ('to sing') in Spanish

Person/Number	Singular	Present Indicative	Plural	Present Indicative
First person	*yo*	*canto* 'I sing'	*nosotros/nosotras*	*cantamos* 'we sing'
Second person	*tú/vos*	*cantas/cantás* 'you sing'	*vosotros/vosotras*	*cantaís* 'you all sing'
Third person	*él/ella*	*canta* 'he/she sings'	*ellos/ellas*	*cantan* 'they sing'

2.3.3 Subject Expression

While both Miskitu and Spanish allow null and explicit referential subjects, subjects are always interpreted as singular in Miskitu due to the ambiguity that arises via the lack of number marking (CIDCA, 1985, p. 129) (2.2).

(2.2) *pro Maria ra kaikri*
pro Maria to see-PST.1SG/PL
'I saw Maria'

In contrast with Miskitu, the use of null and overt subjects in Spanish is governed by discourse-pragmatic elements, with null subjects permitted for both singular and plural referents (Quesada, 2015). Due to this distribution

of implicit and explicit subjects, Spanish is considered to be a full null subject language. While Miskitu also allows null subjects, they are only permitted with singular referents. Based on this distribution, Miskitu would be classified as a partial null subject language. In a partial null subject language, the use of implicit subjects can be restricted by a variety of dimensions, including person, tense, or in the case of Miskitu, number (Camacho, 2013, p. 34).

The only way to avoid the use of plural pronouns in Miskitu is through another verb construction that does not require explicit pronouns. This construction consists of a main verb in the form of a present participle along with the auxiliary verb *banghwaia* conjugated for person (CIDCA, 1985, p. 130). The distinctions between the two different systems are shown below in (2.3 c) and (2.4 c).

(2.3) (a) *yang nani aiwanisna*
we PL sing-PRS.1
'we sing'

(b) *man nani aiwanisma*
you PL sing-PRS.2
'you all sing'

(c) *witin nani aiwanisa*
they PL sing-PRS.3
'they sing'

(2.4) (a) *aiwani banghwisna*
sing-PP 1PL-aux-PRES
'we sing'

(b) *aiwani banghwisma*
sing-PP 2PL-aux-PRES
'you all sing'

(c) *aiwani bangwisa*
sing-PP 3PL-aux-PRES
'they sing'

Due to the structural differences in Miskitu and Spanish, it seems likely that MCS speakers would differ from their NS-speaking counterparts in regard to number marking on verb morphology and the use of null and overt subject pronouns. One would expect that given the contact situation between these two languages, MCS speakers may not mark number in the verb morphology as consistently and may use more null subjects with singular referents.

2.4 Summary

In this chapter, I presented information about the sociohistorical background of MCS, including an explanation of the language shift Miskitu speakers underwent in the twentieth century, and how the geography of Nicaragua has led to little contact between NS speakers on the Pacific coast and Miskitu/MCS speakers on the Caribbean coast. I also provided an overview of the Miskitu language, as it relates to SPE and number marking on verb morphology. Subsequent chapters will address the following questions: what affect has the history of language contact between Miskitu and Spanish had on morphosyntactic features in MCS?, and do the linguistic differences that exist in these two languages, regarding SPE and number marking, lead to processing difficulties for MCS speakers?

Chapter 3

Previous Findings on SPE and Number Marking

This dissertation explores two morphosyntactic features in Mosquito Coast Spanish: subject pronoun expression and variable number marking. The former is a well-documented high frequency variable and the latter a lesser-documented low frequency variable. In this chapter, I will give an overview of previous research on both topics. In addition, I will present information on how these two variables have been studied together in the past, as well as the explanations that I will use to describe the results of my MCS data in regard to both features.

3.1 Subject Pronoun Expression in Spanish

In this section, I will discuss the findings for subject pronoun expression (SPE) across varieties of Spanish, focusing on situations of language contact as these are the most relevant to the current study.

3.1.1 Distribution of Overt and Null Subjects

Spanish is considered to be a null subject language that permits both explicit and implicit referential subjects (3.1).

(3.1) *Me vin-e con Pepe hoy a la oficina; Pepe/él/Ø*
I come-PST.1SG with Pepe today to the office; he
viv-e cerca de mi casa
live-PRS.3SG close of my house
'I came to the office with Pepe today; he lives close to my house.'
(adapted from Silva-Corvalán, 2001, p. 154)

The type of variation in (3.1) is considered morphosyntactic in nature. The fact that Spanish is a null subject, or pro drop, language makes all of the subject choices in (3.1) available to the speaker (*Pepe/él/Ø*); however, many linguistic factors contribute to whether an overt subject is realized or not, including person/number of the referent (Cameron, 1993; Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008; Michnowicz, 2015), the variety of Spanish (Otheguy, Zentella, and Livert, 2007), the type of information expressed by the subject (topic, focus, new, previously activated, contrastive, etc.) (Silva-Corvalán, 1994; Davidson, 1996), the ambiguity of the verb form (Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008; Michnowicz, 2015), and the coreferentiality of the current subject with the previous subject (Cameron, 1992; Silva-Corvalán, 1994; Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008; Michnowicz, 2015).

The use of a null subject in Spanish is permissible due to the rich verb morphology and agreement between subjects and verbs in the language (Quesada, 2015, p. 21); however, the decision to choose one form over the other is not random. Silva-Corvalán (2001, p. 154) states that what has been called "optional" subject expression in Spanish is really only valid in a limited number of discourse contexts. Although it is not obligatory to use an explicit subject, they can be used to avoid ambiguity between first and third person forms (*yo hablaba* 'I spoke' vs. *él/ella hablaba* 'he/she spoke') or to mark a contrast (Bruyne and Pountain, 1995, p. 143). In addition, explicit subjects are used to indicate focus or in cases of switch reference. In contrast, null subjects communicate that the same referent continues to be the focus of attention (Quesada, 2015, p. 32). To summarize, if the information transmitted by the subject is new or focal, an explicit form will be used. On the other hand, if the information expressed by the subject is recoverable by the hearer, a variety of pragmatic factors come into play in order to determine if the subject should be expressed as a noun phrase, an overt pronoun, or a null pronoun.

While the use of explicit vs. implicit subjects forms follow the general guidelines described above, certain contexts result in more flexibility than others, for example, coreferentiality (3.2). According to Silva-Corvalán (2001, p. 154), an explicit subject is realized 25% of the time when the current and previous subject are coreferential and 50% of the time when the current subject indicates a change in referent.

(3.2)	*Ø/yo*	*habl-o*		*bien*	*español,*	*pero*	*el*	*frances*	*Ø/yo*	*lo*
	I	speak-PRS.1SG		well	Spanish,	but	the	French	I	it
	habl-o		*muy*	*mal*						
	speak-PRS.1SG		very	badly						

'I speak Spanish well but French, I speak very badly.'
(adapted from Silva-Corvalán, 2001, p. 154)

In contrast with the flexibility of coreferentiality, contexts where an explicit subject is almost obligatory or expected include contrast (3.3), focal information (new or contrastive) (3.4), and referent clarification (3.5).

(3.3) (a) *Mi señora habla bien inglés pero yo lo hablo muy quebrado.*
'My wife speaks English well put I speak it very brokenly.'
(b) **Mi señora habla bien inglés pero Ø lo hablo muy quebrado.*
*'My wife speaks English well put I speak it very brokenly.'
(adapted from Silva-Corvalán, 2001, pp. 154-155)

(3.4) (a) A: *¿Quién trajo este diario? B: Yo lo traje.*
(b) A: *¿Quién trajo este diario? B: Ø lo traje.*
(adapted from Silva-Corvalán, 2001, p. 155)

(3.5) S: *Pues, se me grabó tanto en la mente que cuando la sepultaron, yo de noche miraba visiones, pero era la realidad. Porque yo despertaba gritando y mi hermana tenía que levantarse a verme.*
C: *Ah fíjate.*
S: (a) *Y ella iba a mi lado* (b) *y yo estaba temblando, que hasta los dientes se oían que pegaban.*
(adapted from Silva-Corvalán, 2001, p. 155)

From a variationist perspective, subject pronoun expression is studied by analyzing the rate of explicit vs. implicit forms. This dissertation will follow this methodological approach by reporting overall rates of usage, as well as what language-internal and language-external factors motivate speakers' choices of overt vs. null subject pronoun production. Findings on Spanish subject expression have indicated that certain factors motivate the explicit realization of pronouns across varieties: grammatical person/number, TMA, and coreferentiality (Cameron, 1993; Orozco and Guy, 2008; Otheguy and Zentella, 2012; Michnowicz, 2015 among others), with research showing that the highest rate of explicit SPE occurs in contexts of first person singular, contrast, new information, focus, or in cases of switch reference (Silva-Corvalán, 2001).

No published studies have been done on subject pronoun expression in NS or MCS; however, this topic has been analyzed extensively in other contact

and non-contact varieties of Spanish. Regarding monolingual dialects, previous studies have shown that the highest rate of explicit subject expression occurs in Caribbean varieties (Olloqui de Montenegro, 1987; Cameron, 1993; Orozco and Guy, 2008), in contrast with low rates in Mexican and Peninsular varieties (Ranson, 1991; Solomon, 1999; Silva-Corvalán, 2001, p. 157; Lastra and Butragueño, 2015; Michnowicz, 2015). An overview of SPE across monolingual varieties suggests that highland varieties have lower rates of explicit subject pronoun use, compared to lowland varieties (Table 3.1).

Table 3.1: Rates of subject pronoun expression in monolingual varieties of Spanish

Variety	Pronoun Rate	Source
Yucatan, Mexico	16%	Michnowicz, 2015
Valladolid, Mexico	21%	Solomon, 1999
Mexico City, Mexico	22%	Lastra and Butragueño, 2015
Puente Genil, Spain	24%	Ranson, 1991
New York Newcomers	30%	Otheguy et al., 2007
Barranquilla, Colombia	36%	Orozco and Guy, 2008
Buenos Aires, Argentina	36%	Barrenechea and Alonso, 1973
Santiago, Dominican Republic	39%	Olloqui de Montenegro, 1987
San Juan, Puerto Rico	45%	Cameron, 1993

(adapted from Limerick, 2017)

Based on Table 3.1, we can see that Caribbean varieties tend to produce high rates of subject pronouns when compared to other regional dialects. In these varieties, not only are overt pronouns used in a redundant manner (3.6), they are also used for inanimate nouns (see 3.7 for the use of *él* used to refer to a telephone), as well as subjects of infinitival verbs (3.8) (Lipski, 1996, pp. 366-367).

(3.6) *Cuando tú acab-e tú me avis-a*
when you finish-SBJV.PRS.3SG you me inform-PRS.3SG
'When you finish, let me know'
(adapted from Jiménez Sabater, 1977, p. 14)

(3.7) *Él son-ó y son-ó*
He ring-PST.3SG and ring-PST.3SG
'It rang and rang'
(adapted from Jiménez Sabater, 1977, p. 14)

(3.8) *para yo poder venir*
for I able.INF come.INF
'for me to be able to come'
(adapted from Quesada, 2015, p. 43)

The information on subject pronoun expression in Caribbean Spanish is relevant for the current dissertation because Nicaraguan Spanish tends to share several characteristics with these varieties, including lenition of coda consonants and variable number agreement in the noun phrase (Guy, 2014, p. 446). In terms of dialect regions of Spanish, Central American Spanish varieties therefore approximate to Caribbean Spanish which would lead us to predict that the Spanish varieties in Nicaragua should exhibit similar rates of SPE.

3.1.2 Contact Varieties

The findings for non-contact varieties of Spanish have been fairly consistent in regard to the highland/lowland split, while contact varieties do not follow the same pattern in terms of their rates of explicit subject pronoun production. Instead, rates seem to be based more on the type of contact language involved: contact with unrelated (non-Romance) languages, such as English or Yucatec Maya, results in increased rates of overt subject pronouns (Lapidus and Otheguy, 2005; Otheguy et al., 2007; Otheguy and Zentella, 2012; Michnowicz, 2015), while contact with related languages, such as Portuguese or Catalan, does not (Carvalho and Bessett, 2015; Prada Pérez, 2015). There are cases that diverge from this basic pattern, such as Veneto-Spanish contact in Mexico, in which contact with another Romance language results in higher rates of explicit pronouns; however, this seems to be an exception in the literature (Barnes, 2010).

Presumably when Spanish is in contact with another Romance language, the constraints on subject expression are similar, therefore not resulting in a significant change in the rate of explicit pronouns produced. When considering contact between Spanish and non-related languages, typological differences are expected, making it easier to use explicit pronouns as the default, a process referred to as a form of simplification. Since overt subjects are permissible in Spanish and do not cause ambiguity, speakers of contact varieties may opt to use explicit forms in order to avoid complex discourse-pragmatic constraints (Michnowicz, 2015, p. 102).

Non-Romance Languages in Contact with Spanish

The literature on SPE in contact varieties, where the other language is considered non-Romance, can be divided into two categories according to whether or

not the language in question is also a null subject language like Spanish. The majority of the research on non-null subject languages in contact with Spanish involves English in the United States. Lapidus and Otheguy (2005) look at the use of explicit subject pronouns in a very specific context. They find that among bilingual speakers of Spanish and English in New York City, there is an increasing use of non-specific *ellos* (3.9). In cases of non-specific reference, an implicit subject would be expected because the referent is unidentifiable; however, Lapidus and Otheguy (2005) find that speakers born and raised in the United States as opposed to being born outside the U.S. and then immigrating, produce more cases of explicit *ellos* in non-specific contexts. These results suggest a weakening of pragmatic restrictions on the use of non-specific *ellos*.

(3.9) **Ellos me rob-aron el auto*
they me rob-PST.3PL the car
'They robbed my car' (without knowing who robbed the car)
(adapted from Lapidus and Otheguy, 2005, p. 67)

Otheguy et al., 2007 also investigate the use of explicit vs. implicit subject expression in New York City but in a more general context. They find that bilingual speakers born and raised in the U.S. use overt pronouns at a higher rate than those that immigrate to the country. These findings suggests cross-linguistic influence or transfer from English, where explicit subject expression is required. Due to the fact that in Spanish, both implicit and explicit subjects are permitted, the results found in these studies represent a change in the strength of a governing restriction in the language.

In contrast with the studies on Spanish-English contact, studies of non-Romance null-subject languages in contact with Spanish provide different information in regard to SPE because they eliminate the possibility of transfer (due to pro-drop status). Studies on this type of contact involve a variety of languages, including indigenous languages.

A study on SPE where a non-Romance null subject language is in contact with Spanish is that of Michnowicz, 2015, in which overall rates and grammatical constraints on subject pronoun expression are analyzed in Yucatan Spanish, comparing Maya-Spanish bilinguals and Spanish monolinguals. Findings show that bilingual speakers produce much higher rates of overt subject pronouns, compared to their monolingual counterparts. In addition, among speakers of Yucatan Spanish overall, overt pronouns are used more often with singular verbs, less-distinctive tense/moods, and in contexts of switch reference, all of which align with other research on SPE in Spanish.

Romance Languages in Contact with Spanish

In terms of Spanish in contact with other Romance languages, de Prada Perez (2015) provides a case of two null-subject languages in contact: Spanish and Catalan. She investigates first person singular subject pronoun expression in Spanish dominant bilinguals and Catalan dominant bilinguals, in comparison with Spanish and Catalan. de Prada Perez (2015) finds that bilingual speakers do not produce more overt subject pronouns than their monolingual counterparts. Her results do not show evidence of convergence seeing as the groups had the same results in their constraint rankings, nor does her data suggest simplification in the overall pronoun rate. She does find some evidence of simplification in weakening or fortifying of certain constraints in the bilingual speakers when compared to monolingual results but only in lower ranked variables (i.e. less categorical distributions).

Carvalho and Besset (2015) investigate Spanish and Portuguese in contact along the border of Brazil and Paraguay, another case of two typologically similar Romance languages. They predict that contact may result in the convergence of grammatical features based on the bilingual context. Overt Spanish subject pronoun expression is based on discourse continuity (switch reference), type and number of grammatical person (singular referent), and TMA (imperfect), while subject pronoun expression in Brazilian Portuguese is also based on person, discourse continuity, and TMA, it is now considered to be a partial pro-drop language (possibly due to reduced agreement morphology). Carvalho and Besset (2015) use data from bilingual speakers in both Spanish and Portuguese. Their results suggest no convergence (25% overt pronouns in Spanish and 46% overt pronouns in Portuguese). They also find the same factor group results for both languages: grammatical person, TMA, and coreferentiality. These results suggest that there is not strong evidence overall for convergence between the two language systems.

While contact with other Romance languages seems to show similar SPE production rates for monolinguals and bilinguals, Barnes' (2010) study is an exception. She analyzes the use of explicit and implicit subject pronouns in bilingual Veneto-Spanish speakers in Mexico. While both Spanish and Veneto are considered to be null subject languages, the bilingual speakers in Barnes' study produce explicit pronouns at a higher rate than monolingual Mexican Spanish speakers. In addition, the bilinguals use overt subject pronouns where a null subject is expected for pragmatic reasons. Despite these findings, the patterning is in-line with monolingual/non-contact varieties, with more overt pronouns being used with first person singular imperfect and in cases of topic

shift and contrast. Barnes does not attribute the results to transfer but to cognitive economy.

Findings for Spanish SPE in Contact Situations

Studies that look at this topic from a variationist perspective generally use data from either sociolinguistic interviews or corpora[2] to determine the rate of explicit subject pronoun expression and the constraints that motivate the realization of overt forms. In regard to linguistic predictors, findings on subject expression in contact varieties of Spanish have indicated that certain factors motivate the explicit realization of pronouns across studies: grammatical person, TMA, and coreferentiality (Otheguy and Zentella, 2012; Michnowicz, 2015 among others). In terms of social factors, other than those related to bilingual competence, language-external predictors do not seem to be significant. This reflects what Silva-Corvalán (2001, p. 163) claims regarding extralinguistic variables, stating that variable subject expression is motivated by the discourse context and not social factors.

[2] *City University of New York (CUNY) Project on the Spanish of New York* used by Lapidus and Otheguy (2005)

Previous studies on Spanish SPE show higher rates of overt pronouns for contact varieties involving a non-Romance language (Lapidus and Otheguy, 2005; Otheguy et al., 2007; Michnowicz, 2015), with the exception of Veneto-Spanish bilinguals also producing higher rates of explicit pronouns. Production rates of SPE among bilingual speakers in these contact varieties of Spanish have been explained by various processes. The results are described as either a simplification strategy to lighten the cognitive load on the speaker or as convergence between two different language systems. Varieties that show evidence of simplification display an increase in overt subject pronoun expression or a weakening of pragmatic constraints (for example less sensitivity to contexts of coreferentiality), while varieties that demonstrate convergence produce rates of subject pronouns that fall between the contact languages (Michnowicz, 2015, p. 102). Based on these descriptions, the increase in explicit subject pronoun expression in Spanish in contact with English could be due to either simplification or convergence, while the increase found in Yucatan Spanish and Veneto-Spanish bilinguals would be due to simplification.

The language in contact with Spanish appears to play a role in the subject expression rate. First of all, there is consistent evidence of transfer when the language in contact with Spanish is a non-null subject language. In studies of Spanish in contact with English (Lapidus and Otheguy, 2005; Otheguy et al., 2007; Otheguy and Zentella, 2012), bilingual speakers with more experience in an English-speaking environment show increased rates of overt subject pronoun expression. The contact language also seems to play a role among the

studies between Spanish and other null subject languages. When looking at related languages, such as Portuguese and Catalan, there is no evidence of transfer (Caravalho and Bessett, 2015; de Prada Pérez, 2015); however, when looking at cases of a non-related language in contact with Spanish, there is (Michnowicz, 2015). Based on these results, it can concluded that languages in a contact situation that have more in common (whether that involves their status as a null subject language or a Romance language) will not show as much evidence of transfer in regard to the expression of subject pronouns.

A Note on Heritage Speaker Literature

"A language qualifies as a heritage language if it is a language spoken at home or otherwise readily available to young children, and crucially this language is not a dominant language of the larger (national) society" (Rothman, 2009). In the context of the United States, a heritage speaker is someone who uses either a non-English indigenous or immigrant language (Valdés, 2001). In the context of Nicaragua, Miskitu would be considered a heritage language but not Mosquito Coast Spanish. While MCS is not classified as a heritage language based on the definitions above, the literature on heritage speaker Spanish and SPE is relevant to the current dissertation topic because both involve simultaneous bilingual speakers of Spanish and a non-Romance language.

The study of heritage speakers of Spanish generally refers to bilingual speakers of English and Spanish in the United States, with the level of bilingualism varying depending on the individual. The topic of subject expression has been examined extensively for this population, including the studies detailed above from Lapidus and Otheguy (2005) and Otheguy, Zentella, and Livert (2007). Other findings suggest that heritage speakers of Spanish possess an in-depth knowledge of implicit and explicit subjects but differ from monolingual speakers in the pragmatic circumstances that they use each (Montrul, 2004). Her findings suggest possible evidence for structural convergence with English. While monolingual Spanish speakers and bilingual English-Spanish speakers do not differ greatly in type of subject (lexical noun phrase vs. pronominal subject) or word order (pre or post verbal subject placement), they do differ in the contexts in which they use overt vs. null subjects. Heritage speakers produce higher rates of redundant overt subjects and illicit null subjects, compared to monolingual Spanish speakers. Lipski (2008c, p.61) confirms this notion as he states that one common characteristic of transitional bilinguals in the United States is the redundant use of subject pronouns.

The findings on subject pronoun expression by heritage speakers of Spanish in the United States can be used to inform my predictions for speakers of MCS

based on the similarities in the type of languages in contact, and due to the state of bilingualism and process of language acquisition that occurs in both contexts. Previous studies (Silva-Corvalán, 1994; Lipski, 1996; Montrul, 2002, 2004) show that where syntax interfaces with semantics and pragmatics, as is the case with subject pronoun expression, language loss or attrition, or incomplete acquisition may occur.

3.2 Variation in Morphological Number Marking and Agreement

Variation in number marking on verbs occurs when the expected number morphology is not always realized. In the context of this dissertation, this occurs specifically with 3pl. forms that lack the necessary plural morpheme that distinguishes the verb from its singular counterpart. The plural marked form is considered to be the standard, while the singular marked form is non-standard based on the fact that it results in a lack of agreement between subject and verb. This type of variation, or something similar, occurs in both Brazilian Portuguese and a few varieties of Spanish.

3.2.1 Variable Number Marking in Brazilian Portuguese

The topic of variable number marking has been observed and discussed widely for Brazilian Portuguese (BP) (Guy, 1981, 2005; Scherre, 1998; Naro and Scherre, 2000; Scherre and Naro, 2001, 2010, 2014; Lucchesi et al., 2009; Brandão and Vieira, 2012; Scherre and Naro, 2014; Mendes and Oushiro, 2015) and manifests itself in a variety of contexts, including in the noun phrase (3.10 a), the predicate phrase (3.11 b), and most relevant to the current dissertation, between subjects and verbs (3.12 b).

(3.10) (a) *O-s* *fregues-es*
the-PL customers-PL
'The customers'

(b) *A-s* *codorna-Ø*
the-PL gamehen-SG
'The gamehens'
(adapted from Naro and Scherre, 2000, p. 235)

(3.11) (a) *A-s* *coisa-s* *(es)t-ão* *muito* *cara-s.*
the-PL thing-PL be-PRS.3PL very expensive-PL
'Things are very expensive.'

(b) *A-s* *coisa-Ø* *(es)tá* *cara-Ø.*
the-PL thing-PL be-PRS.SG expensive-SG
'Things are very expensive.'
(adapted from Naro and Scherre, 2000, p. 235)

(3.12) (a) *Eles* *ganh-am* *demais.*
they earn-PRS.3PL too much
'They get too much.'

(b) *Eles* *ganh-a* *demais.*
they earn-PRS.3SG. too much
They get too much.
(adapted from Naro and Scherre, 2000, p. 235)

The importance of the previous research on subject-verb number agreement in BP lies in the breadth and depth of studies over the past 40 years, providing us with a clear picture of how this variable interacts with social and linguistic factors. In addition, this topic has been studied from a variationist framework, which makes it suitable for comparison with the methodological perspective of this dissertation.

Variable subject-verb number marking in BP has been observed with both 1pl. (*nós fazemos/faz* 'we do') and 3pl. (*eles fazem/faz* 'they do) forms, with 3pl. cases making up the majority of this research (Mendes and Oushiro, 2015, p. 362). While several linguistic predictors have been identified in the research on 3pl. subject-verb agreement in BP, phonic salience, preceding markers, and animacy of the subject are the three most prominent factors that are shown to favor singular marked forms (Mendes and Oushiro, 2015, p. 362).

In the context of 3pl., phonic salience relates to the number morpheme marked on the verb, referred to as the singular/plural opposition. When there is "a less perceptible difference in form between the singular and plural", we see higher rates of non-agreement between subjects and verbs (Naro and Scherre, 2000, p. 243). This has been connected to tense/mood, where present and imperfect forms are considered less salient (*ganha/ganham* 'he/she wins/they win'), where the difference in forms is based on the nasalization of the final vowel. In contrast, the most salient forms are found in the preterite (*fez/fizeram* 'he/she did/they did'), where the difference in forms is found in both the root vowel and in the number of syllables (Scherre and Naro, 2010, p. 165).

Naro (1981, pp. 75-76) presents a hierarchy of salience for 3pl. forms which is based on whether or not the singular and plural morpheme on a verb is stressed, with level 1 encompassing lower salience forms and level 2 encompassing higher

salience forms (Table 3.2). Within each of the two levels, several morphological classes are ranked based on a variety of factors such as presence/absence of nasalization (level 1.a. and 2.d.), the presence of an additional morpheme (level 1.c. and level 2.e.), distinct morpheme stressed (level 1.b and level 2.f.), two completely distinct forms (level 2.g.), and change in stress position (level 2.h.).

Table 3.2: Hierarchy of Salience for 3pl.

Level 1: Unstressed	
a. *come/comen* [-i/-ĩ]	'they eat'
b. *fala/falam* [-a/-ũ]	'they speak'
c. *faz/fazem* [-Ø/-ĩ]	'they do'
Level 2: Stressed	
d. *dá/dão* [-a/- ãw]	'they give'
e. *comeu/comeu* [- ew/- er ũ, - iw/- irũ, - oj/- or ũ]	'they ate'
f. *falou/falaram* [- o/- ar ũ]	'they spoke'
g. *é/são*	'is/are'
h. *disse/disseram* [-Ø/- rũ]	'they said'

(adapted from Mendes and Oushiro, 2015, p. 362; Naro, 1981, pp. 75-76)

The hierarchy and accompanying salience hypothesis proposed by Naro (1981) suggest that the smaller the difference in phonetic realization of plural vs. singular forms, the more lack of agreement that occurs. This theory was validated by Scherre and Naro (2010), who looked at whether salience or tense/mood was a stronger predictor for singular marked forms. They find that while both are significant factors, salience is the stronger predictor.

In addition to salience, as previously mentioned, preceding markers and animacy are also significant factors in predicting singular marked forms in BP. In regard to preceding markers, the research shows that both plural morphemes in the noun phrase as well as plural referents in previous sentences, result in higher rates of subject-verb agreement (*os meninos foram/foi* 'the boys went' vs. *os menino-Ø foram/foi* 'the boys went') (Scherre and Naro, 1992; Pereira, 2004; Rubio and Gonçalves, 2012; Mendes and Oushiro, 2015; Oushiro, 2015). Animacy as a factor refers to the subject NP. Past results have shown that [+animate] and [+human] subjects are more likely to result in agreement between subject and verb (Monguilhott et al., 2009; Brandão and Vieira, 2012; Pereira, 2004; Rubio and Gonçalves, 2012; Mendes and Oushiro, 2015; Oushiro, 2015).

Additional language-internal factors include subject position and the presence of intervening material between the subject and the verb. According to previous research (Guy, 1981; Brandão and Vieira, 2012; Rubio and Gonçalves,

2012; Mendes and Oushiro, 2015; Oushiro, 2015), cases where the subject directly precedes the verb strongly favor agreement (*eles fazem/faz* 'they do'). The presence of intervening material between subject and verb slightly disfavors agreement when the word order is SV (*ós ainda não casamos/casou* 'we haven't yet gotten married') (Mendes and Oushiro, 2015, p. 363). Finally, postponed subjects strongly disfavor agreement (*chegaram/chegou dois caras* 'two guys arrived') (Mendes and Oushiro, 2015, p. 363).

The social predictors of variation in subject-verb agreement include level of education and speaker sex. This feature is highly stigmatized and socially stratified in BP, which speaks to the claim that social stratification is more common for features that are widespread (Lafford, 1986). While rates of non-standard agreement, or lack of number agreement between subjects and verbs, differ greatly depending on the speaker, it is a feature present in all levels of the BP speech community (Mendes and Oushiro, 2015). Higher rates of singular marked forms are found among speakers with less formal education and in rural areas compared to educated urban speakers (Brandão and Vieira, 2012, Naro and Scherre, 2013, Mendes and Oushiro, 2015; Oushiro, 2015). According to Mendes and Oushiro (2015, p. 362), 3pl. subject-verb agreement varies greatly among communities in Brazil, ranging from 17% to 94%. Evidence of this variability can be found in the various studies done within state of São Paulo. Research shows that speakers in the rural zones produce rates of 3pl. agreement as low as 24% (see Pereira, 2004), while speakers in urban zones, such as São José do Rio Preto and the city of São Paulo, produce rates of 3pl. agreement 73% and 87% respectively (see Rubio and Gonçalves, 2012 and Oushiro, 2015).

Speaker sex is another significant language-external factor predicting subject-verb agreement in BP, with men in urban communities producing more cases of singular marked forms. In contrast, in rural communities, women produce higher rates of non-agreement (Guy, 1981; Pereira, 2004; Rubio and Gonçalves, 2012; Mendes and Oushiro, 2015). Mendes and Oushiro (2015, p. 364) explain that this difference is due to professional roles in each context. Findings for urban settings represent expected sex stratification while findings for rural settings are based on men interfacing with outside communities due to work opportunities available outside the community (Labov, 2001).

Due to the lack of literature on subject-verb number agreement in Spanish, this dissertation will draw heavily from the research of this phenomenon in BP to guide its methodology, specifically the analysis of linguistic and social factors that predict non-standard singular marked forms. This comparison is justified also due to the similar backgrounds of adult second language acquisition that occurred in both Brazil and Nicaragua with MCS speakers. Critchfield and

Lívio (*forthcoming*) discuss the connection between BP and MCS, and while these two groups of speakers differ in the details of their historical language shifts, they share similarities in internal linguistic factors conditioning non-agreement.

3.2.2 Variable Number Marking in Spanish

Variation in number marking has been widely studied in Spanish in regard to final /s/ deletion in the noun phrase, attested in Caribbean, Central American, coastal South American, and Rioplatense dialects according to Guy (2014, p. 446) (3.13); however, this topic is not widely attested in the context of verb morphology.

(3.13) *La-s* *niña* *bonita*
the-PL girl.SG pretty.SG
'The pretty girls'
(Guy, 2014, p. 445)

Of the few observed cases of variation in number marking on verbs, Caribbean Spanish varieties have been the most widely published on, particularly in regard to final /s/ deletion on 2sg. verb forms. Both Puerto Rican Spanish (3.14) and Dominican Spanish (3.15) show high rates of /s/ deletion in this context.

(3.14) *Tú* *com-ía*
you eat-PST.IPFV.3SG
'You ate'
(Hochberg, 1986b, p. 611)

(3.15) *Tú* *cant-a*
you sing-PRS.3SG
'You sing'
(Alba, 2004, p. 125)

In Dominican Spanish, variation in number marking occurs not only in 2sg. but in 3pl. forms as well, via final /s/ and /n/ deletion respectively. While final /s/ deletion is a common feature in many varieties of Spanish, final /n/ deletion is not, and according to Lipski (1996, p. 364), is specific to Dominican Spanish. This weakening of coda consonants results in apparent subject-verb non-agreement and what appears to be a less richly inflected verb system, where the 3sg. verb form can be used with three different subjects: 2sg., 3sg., and 3pl. (3.16).

(3.16) *Com-e*
eat-PRS.3SG
'you eat'; 'he/she/it eats'; 'they eat'
(Foote and Bock, 2012, p. 432)

While final /n/ elision could be attributed to the phenomenon discussed in this dissertation for MCS, clear evidence can be provided that counters this idea, specifically cases of past tense conjugations. Imperfect verb forms may appear to exhibit final consonant deletion (3.17 a); however, this is not true for preterite examples. If final /n/ deletion were occurring in this context, speakers would produce forms such as *dijero-Ø* (instead of *dijeron* 'they said'), which is not the case. Examples of verbs in the preterite like (3.17 b) demonstrate the clear use of 3sg. verb morphology. Based on this distinction from Dominican Spanish, this variety is not ideal for comparison with MCS. The lack of compatibility in regard to this feature in MCS and Caribbean Spanish dialects in general is also based on the fact that in Puerto Rican Spanish and Dominican Spanish, the phenomenon appears to be phonological in nature, while in MCS, it appears to be morphological.

(3.17) (a) *Ellos nunca aguant-aba de nada.*
they never tolerate-PST.IPFV.3SG of nothing
'They never tolerated anything.'

(b) *Los miskito-s dij-o*
the-PL miskito-PL say-PST.3SG
'The Miskitus said'

Another case of subject-verb non-agreement has been observed in Afro-Bolivian Spanish (also referred to as Yungueño Spanish), where speakers use 3sg. morphology as an invariant verb form for all persons and numbers (3.18 b) (Lipski, 2008a,b).

(3.18) (a) *Nojotro tien-e jrutita*
we have-PRS.1SG fruit
'We have fruit'

(b) *Yo no conoc-ió hacienda*
I no know-PST.1SG ranch
'I never knew the haciendas'
(Sessarego, 2011, p. 128)

While the result of this process is apparent non-agreement, unlike the cases described for Caribbean Spanish and MCS, there is no variation in this variety.

While not necessarily suitable for comparison with MCS in regard to constraints on variability, the historical contact between speakers of African languages and speakers of Spanish that occurred in the Yungueño region of Bolivia is important to discuss when reviewing possible explanations for variable number marking and non-agreement.

Lipski (2008a) attests that Afro-Bolivian Spanish displays semi-Creole tendencies due to the reduction of verb morphology to one form for all subjects. In this hypothesis, Lipski suggests that this phenomenon reflects a previous creole stage in the language caused by contact with African substrate languages. In contrast, Sessarego (2011, 2012) suggests that this feature of Afro-Bolivian Spanish is the "result of imperfect processes of second language acquisition which were able to crystallize in an environment far from the pressure posed by the linguistic norm and language standardization" (Sessarego, 2011, p. 126). He claims that this seems more plausible based on the fact that other second language varieties and nonstandard varieties of Spanish have similar characteristics but do not have history of contact with creole languages. Sessarego's theory attributes the simplified verb morphology system in Afro-Bolivian Spanish to processing difficulties that L2 learners experience at the syntax/semantics interface. Regardless of explanation, the use of one invariant verb form in Afro-Bolivian Spanish does not reflect what has been observed for MCS, where variation in subject-verb non-agreement only occurs in third-person contexts.

An additional case of variation in number marking between subjects and verbs is found in the speech of bilingual speakers of English and Spanish in the United States (Lipski, 1993, 2008c). Lipski (2008c, p. 56) labels speakers with highly developed passive linguistic skills but less developed active production of Spanish, as transitional bilinguals. Transitional bilinguals are often people who speak Spanish as children but who later become English-dominant as they get older (Lipski, 2008c, p. 56). A feature of the Spanish spoken by transitional bilinguals in the U.S. is the production of third person verb forms with first and second person referents; this includes the use of 3sg. verb morphology (3.19) in place of 3pl. (Lipski, 2008c, p. 59). Lipski (2008c) attributes the non-agreement produced by transitional bilinguals to insufficient input of Spanish in an English-speaking environment or to the shift toward English as the dominant language.

(3.19) *Vien-e* *mis* *tíos*
come-PRS.3SG my aunt and uncle
'My aunt and uncle come'
(Lipski, 2008c, p. 59)

Previously, Critchfield and Lívio (*forthcoming*) analyzed variable number marking on verb morphology in MCS, the only other study on the topic in this variety. Drawing from the literature on BP, they look at four factors that have proven to be significant in predicting lack of number agreement between subjects and verbs. Their findings indicate that VS syntactic order, presence of intervening material between subjects and verbs, and high phonic salience of the verb morphology all favor lack of agreement, occurring overall in 20.3% of the data set. This dissertation will expand on the study Critchfield and Lívio (*forthcoming*) did by increasing the number of participants and also examining additional internal and external factors.

Explanations for the number variation observed in the Spanish dialects discussed above vary for each of the aforementioned cases. One possible account attributes the lack of subject-verb agreement in Afro-Bolivian Spanish to simplification, where Lipski (2008a) suggests that this dialect displays semi-Creole tendencies due to the use of one verb form for all subjects. Dominican Spanish also represents a case of simplification due to the 3sg. verb form being utilized for several persons and numbers.

Number variation in Spanish has also been attributed to acquisition. In the case of transitional bilinguals in the United States, it has been claimed that number variation is due to residual competence in Spanish of some sort. This could be caused by insufficient input in the context of a primarily English speaking environment or due to an overall shift toward the majority language, in this case English. Sessarego's argument for Afro-Bolivian Spanish, while different than that for transitional bilinguals, would also fall into this category, as he credits the use of one invariant verb form to imperfect processes of second language acquisition.

The current dissertation will add to the small body of work on variability in number marking and non-agreement in verb morphology in Spanish. Historical contact with African languages has been used to explain why this phenomenon appears in the varieties it does (Lipski, 2008a). My analysis of MCS will offer information on how more recent contact has played a role in the development of this phenomenon in Nicaragua and how this relates to historical contact in places such as the Caribbean and Bolivia. Due to the fact that the language contact that occurs on the Caribbean coast of Nicaragua is currently taking place, the analysis I present will inform how contact affects number marking and agreement in the present day. This will allow for a glimpse into a change in progress and an examination of what constraints are at work in the variation of this phenomenon in MCS. This dissertation will also seek to explain how the process of non-agreement is different in MCS when compared to other varieties

that show variability in number marking on verbs, primarily that this process cannot be categorized as coda consonant deletion.

A Note on Final /s/ Deletion in Nicaragua

While limited in number, previous studies on the Spanish spoken in Nicaragua (Lipski, 1984, 1985, 1996; Chappell, 2013, 2014, 2015a, 2015b, 2016) provide a foundation for the description of a few notable features, one of which is /s/ reduction (one of the only topics that has been presented from a variationist perspective). This research is relevant to the current topic of variable number marking because it reflects what has been described as non-agreement between subjects and verbs in Caribbean Spanish.

/s/ reduction is a topic discussed extensively in the literature on Spanish linguistics and represents a case of weakening most commonly observed in the coda position of syllables, whether word-internal or word-final. According to Lipski (2011, p. 72), the syllable is the central axis of variation and represents the best place to isolate cases of /s/ reduction. The coda is considered to be the weakest of all syllable positions in regard to phonological processes, prone to reduction and replacement with weaker versions of the consonant in question (Lipski, 2011, p. 73). The realization of /s/ in this position represents a case of phonological variation and involves the analysis of a variety of allophones realized in different phonological contexts (Silva-Corvalán, 2001, p. 87). It has been observed that in Spanish, the aspiration of /s/ occurs more frequently in word-final position, where it is morphemic and contextually redundant (Silva-Corvalán, 2001, p. 241). The most common variants in word-final position are sibilance [s], aspiration [h], and deletion [Ø]. According to Silva-Corvalán (2001, p. 87), word-final /s/ reduction is the most interesting because the phoneme /s/ acts as either a plural morpheme (*la casa* vs. *las casas*) or as an inflectional morpheme (*canta* vs. *cantas*) in this position, both of which represent cases of high functional output. Schwenter (2011, p. 123) asserts that /s/ weakening is most frequently found with plural nouns, where the realization of /s/ is somewhat irrelevant due to the availability of plural information from other sources.

Penny (2000) discusses the distribution of /s/ weakening in Latin America and states that lowland, island, and coastal varieties tend to demonstrate this type of reduction the most. According to Penny (2000, p. 229), these geographic zones were settled primarily by individuals from Andalusia. The same phenomenon is found in southern Spain today; however, one main difference between /s/ weakening in Peninsular Spanish and Latin American Spanish is the number of variants that exist (Penny, 2000, p. 230). While there are a variety of realizations of weakened /s/ in Peninsular Spanish, aspiration [h] and

deletion [Ø] are the variants most commonly found in Latin America (Penny, 2000, p. 230).

The literature on /s/ reduction has identified four main factors that predict weakening: phonological context, following stress, word length, and frequency. Across dialects, reduction occurs before a consonant, however before a vowel or a pause, it is less common. In varieties that do show weakening before vowels, reduced variants are more likely to be produced when preceding an unstressed vowel. In addition to these tendencies, /s/ reduction is more frequent in longer, low-frequency words (Chappell, 2015a).

Nicaragua is identified as a country where frequent /s/ weakening can be found (Penny, 2000, p. 230). In fact, reduction of syllable-final and word-final /s/ occurs at a rate higher than in any of the other Central American countries, with coda /s/ being aspirated or elided completely almost categorically, particularly at the end of a constituent (Lipski, 1996, p. 312). One of the only contexts where /s/ reduction is blocked in Nicaraguan Spanish is when a plural determiner precedes a word beginning in a tonic vowel (3.20 a), as opposed to a non-tonic vowel (3.20 b).

(3.20) (a) [lasonse]
las once
'eleven'

(b) [lohanimale(h)]
los animales
'the animals'
(Lipski, 1996, pp. 312-313)

While few quantitative linguistic studies exist about Nicaraguan Spanish, Chappell (2013, 2014, 2015a, 2015b, 2016) has published several articles that look at features of this variety from a variationist perspective, particularly focusing on coda /s/ reduction. She writes that Nicaraguan Spanish is considered one of the most "radical aspirating dialects" in regards to this phenomenon, which makes her work an important contribution to the literature on /s/ weakening (Chappell, 2014, p. 222).

Nicaraguan Spanish is characterized by the reduction of coda /s/ to glottal friction, elision, and glottal constriction. In Nicaraguan Spanish, glottal friction, or aspiration, is found before both consonants (3.21 a) and vowels (3.21 b), while full elision of /s/ occurs before pauses (3.21 c) (Chappell, 2014).

(3.21) (a) [mohka]
mosca
'fly'

(b) [mah alto]
más alto
'higher'

(c) [pae]
pares
'pairs'
(adapted from Chappell, 2014)

Compared to other non-contact varieties of Spanish in the region, Nicaraguan Spanish is described as the most extreme /s/ weakening variety in Central America (patterning more closely with Caribbean dialects) (Lipski, 1996, p. 312).

Most of the literature on /s/ reduction has focused on non-contact varieties of Spanish. Silva-Corvalán (2001, p. 256) states that /s/ weakening, specifically aspiration, has been dated back to the sixteenth century, an indication that this phenomenon is not due to contact with other languages. Lipski (2011, p. 73) echoes this sentiment, stating that coda consonant reduction is a phenomenon that occurs frequently in Romance as a natural language process. Chappell (2016) confirms this notion in her study of bilingual speakers of Miskitu and Spanish, who like speakers of NS, display /s/ reduction, particularly those that learn Spanish at an early age. Miskitu speakers who acquire Spanish later in life have lower rates of coda /s/ reduction before consonants; but weaken /s/ in prevocalic environments at the same rate as NS speakers. These results are interesting because /s/ reduction before consonants is more common across dialects, while reduction before vowels is less common and is only found in varieties with extreme /s/ weakening. Her findings support Silva-Corvalán and Lipski's claims by showing that contact and non-contact varieties within the same country both demonstrate /s/ weakening, with higher rates of reduction being observed in NS (non-contact).

Based on the findings for both NS and MCS in previous studies, final /s/ elision in 2sg. verb forms would be expected. However, this has not been directly observed in my MCS data, nor has it been discussed in previous literature; therefore, possible variable number marking on 2sg. verbs in MCS is outside the scope of the current dissertation. In addition, the phenomenon discussed in this study does not reflect the loss of a coda consonant, as I have demonstrated and can therefore not be attributed to a phonological process. If in future studies final /s/ elision on 2sg. verbs is found, this would reflect a different process and would constitute a separate explanation and analysis.

3.3 Functional Compensation

The expression of pronominal subjects in Spanish is generally considered to be pragmatic or stylistic in nature (Silva-Corvalán, 2001). However, this notion was challenged when Hochberg (1986b) proposed the functional use of subject pronouns in Puerto Rican Spanish, a variety in which final /s/ deletion is prevalent on 2sg. verb forms. This case of coda weakening results in homophony between 2sg. and 3sg. forms (3.3), as well as 1sg, 2sg, and 3sg. in some cases (3.4). As a result of this ambiguity and as a means of compensation, Hochberg (1986b, pp. 612-613) suggests that speakers produce high rates of explicit subject pronouns, reporting a rate of 84% /s/ deletion in all 2sg. contexts. In her data the highest number of explicit pronouns is produced with the pronoun *tú*, compared with other persons/numbers. This extensive use of pronouns where ambiguity in the verb form occurs helps to identify who the subject is.

Table 3.3: Homophogy 2sg. and 3sg.

Subject Pronoun	Conjugation
yo	*hablo*
	'I speak'
tú/vos	*habla*
	'you speaks'
él/ella	*habla*
	'he/she speaks'
nosotros/nosotras	*hablamos*
	'we speak'
ellos/ellas	*hablan*
	'they speak'

Table 3.4: Homophony 1sg., 2sg. and 3sg.

Subject Pronoun	Conjugation
yo	*hablaba*
	'I spoke'
tú/vos	*hablaba*
	'you spoke'
él/ella	*hablaba*
	'he/she spoke'
nosotros/nosotras	*hablábamos*
	'we spoke'
ellos/ellas	*hablaba*
	'they spoke'

The idea that higher rates of overt pronouns compensate for ambiguity in verb morphology that results via the loss of number marking was coined the Functional Hypothesis (FH) by Hochberg (1986b). In this approach, the word 'functional' refers to the connection between "language structure and behavior to communication" (Hochberg, 1986b, p. 609).

Variable number marking on verbs and subject pronoun expression have been linked in several other studies in Spanish (Hochberg, 1986a, Hochberg, 1986b; Alba, 2004; Otheguy and Zentella 2012). Initially the FH received support in Caribbean Spanish varieties. In addition to Puerto Rican Spanish (Hochberg, 1986a,b), Dominican Spanish also shows high rates of overt 2sg. subject expression. Alba (2004, p. 124) reports numbers as high as 96% and only 31% for other persons/numbers. These findings reveal that functional constraints play a role in the production of overt and null subjects in Caribbean Spanish. In these varieties, speakers must compensate for the aspiration or deletion of final /s/ in a non-phonetic way.

Functional constraints on /s/ deletion have been studied in a variety of /s/-deleting dialects in addition to Caribbean Spanish as well. While earlier studies provided support for the FH, more recent research has not: Southern Peninsular Spanish (Hernández Campoy and Trudgill, 2002), Miami Cuban Spanish (Lynch, 2009), Mexican Spanish (Lastra and Butragueño, 2015), Argentinian Spanish (Duarte and Soares da Silva, 2016), and later studies of Puerto Rican Spanish (Cameron 1993, 1996; Duarte and Soares da Silva, 2016). These studies find that final /s/ deletion in 2sg. forms does not result in overt pronoun rates that are significantly higher when compared to other persons/numbers. Research (Bentivoglio, 1987; Ranson, 1991; Silva-Corvalán, 1994; Torres Cacoullos and Travis, 2011;) has also shown that the use of explicit pronouns prevents

contextual ambiguity and not ambiguity of the verb form (Quesada, 2015, p. 43).

While the findings for the FH have been mixed, the purpose of including the hypothesis in this dissertation is not to provide evidence for or against its validity but to justify the joint analysis of variation in number marking in verb morphology and subject pronoun expression in MCS.

3.4 Theoretical Explanations

3.4.1 Incomplete Language Acquisition

This dissertation looks for evidence that provides support for or against the notion of incomplete acquisition of Spanish by MCS speakers. Silvina Montrul, who has published extensively on the idea of incomplete acquisition in heritage language speakers of Spanish in the United States, uses the term "to refer to the grammatical development and outcome of heritage language acquisition" (Montrul and Silva-Corvalán, 2019, p. 269). According to Montrul (2006), incomplete acquisition occurs when speakers do not acquire linguistic features in a target-like manner, often due to either transfer or fossilization.

The concept of incomplete acquisition originates from notions discussed by Bley-Vroman (1989) and formalized in Schachter's (1990) *Incompleteness Hypothesis*, in which he asserts that late bilingualism is characterized by non- target-like use of linguistic features due to being acquired after the first language is fully formed. Sorace (1993) expands on this notion by distinguishing between a feature that is incomplete (non-existent) and one that is divergent (non-target-like version) in the second language grammar. Montrul (2006, 2008, p. 20) contests the claim that incompleteness is only a result of age of acquisition and instead suggests that the incompleteness present in learners' grammars is due to insufficient target language input. She bases this idea on the fact that early bilinguals (< aged three or four) also show evidence of incomplete acquisition.

The term 'incomplete acquisition' has become a debated topic in the recent years, particularly in the heritage language literature. While there has been a tendency to label heritage language competence as 'incomplete', this term has been challenged by several people (Pascual Y Cabo and Rothman, 2012) who suggest that heritage language speaker competence is not incomplete, but divergent from monolingual speakers due to different access to the target language. In addition, the comparison of heritage language speakers with monolingual speakers has been criticized for not being a viable comparison (Pascual Y Cabo and Rothman, 2012). Since heritage learners acquire their family language as

young children, as do monolinguals, their competence has traditionally been compared to monolingual speakers of the same language; however, Pascual Y Cabo and Rothman (2012) argue that the input for these groups is very different. Heritage language speakers are exposed to input that has undergone attrition across generations and additional contact-related changes. Otheguy (2016, p. 301) asserts that differences in second-generation bilinguals are not errors but a result of "normal intergenerational language change accelerated by conditions of language contact". What is clear is that heritage speakers should be compared with the speakers who provided the input and not necessarily monolingual speakers.

Another issue that has been brought up regarding the term 'incomplete acquisition' is that it subsumes both arrested development of features and attrition as it is almost impossible to know which process took place in childhood when analyzing the speech of adult heritage language speakers (Pascual Y Cabo and Rothman, 2012). In addition, Kupisch and Rothman (2018) assert that the term is pejorative and misleading to non-linguistics, and should instead be replaced with the term 'differential acquisition'. The use of 'incomplete acquisition' has continued to be defended as it does not make an evaluation of the speaker and because the word 'incomplete' is used in all acquisition contexts, not only those that describe heritage language speakers (Montrul, 2018; Silva-Corvalán, 2018; Domınguez et al., 2019; Montrul and Silva-Corvalán, 2019). Based on these points, I will use the term 'incomplete acquisition' to describe the the acquisition process undergone by the first generation of Miskitu speakers who were forced to shift to Spanish in the mid-twentieth century along the Caribbean coast of Nicaragua.

The notion of incomplete acquisition (also termed acquisition failure or imperfect acquisition) has also been discussed in the general contact linguistics literature to describe the state of learners' interlanguage during second language acquisition (Winford, 2003). Winford (2003, p. 225) states that access to the target language and interaction with target language speakers are important social constraints involved in language acquisition. As I mention previously, MCS is not considered a heritage language; however, literature about heritage speakers of Spanish is utilized in this dissertation due to the robust amount of research that has been published on this context of bilingual Spanish speakers.

3.4.2 Contact Phenomena

Part of the analysis of this dissertation will include a discussion of possible explanations for the results in MCS and how they differ from NS. Three general categories include: ancestry, innovation, and contact (Guy, 2014). This discus-

sion will focus on contact phenomena, touching on topics related to acquisition and language shift, such as L1 (substrate) influence, simplification, and internal developments.

Interlanguage is defined as the learner variety of a target language where L1 features are incorporated into the L2 version of the target language (Winford, 2003, p. 210). Learners use a variety of strategies to communicate in the L2, particularly when there is a structural difference between the two languages in question. According to Winford, 2003 (p. 209), the changes learners make to their interlanguage fall into three categories: 1) L1 or substrate influence, 2) simplification of target language structures, and 3) internal changes to the interlanguage itself. All of these strategies are considered part of partial or incomplete acquisition, generally present in situations of contact and natural second language acquisition (SLA) (Winford, 2003, p. 217).

L1, or substrate influence refers to the transfer of language structures or grammatical properties from the first language to the second language, specifically when the L1 and the L2 differ in some aspect. While morphemes do not tend to be transferred directly, one strategy that does occur is the reduction or elimination of target language inflections, such as agreement, gender, and number morphemes (Winford, 2003, p. 213). This explanation aligns with what has been observed in the variation in number marking on verb morphology in Spanish.

Substrate effects have been discussed for both Puerto Rican Spanish and Dominican Spanish in regards to variability in number marking (Guy, 2014). Guy (2014, p. 457) claims that the preference in West African languages for open syllables represents a case of substrate influence in Caribbean Spanish, where slaves from West Africa acquired Spanish as a second language. Lipski (2008a) has also used this argument in the case of Afro-Bolivian Spanish. He suggests that this variety displays semi-Creole tendencies in the use of one verb form for all subjects. In this hypothesis, Lipski attributes this phenomenon to a previous creole stage that the language passed through due to contact with African substrate languages.

Simplification, another strategy used by learners, is a broad term but in the context of language acquisition is defined as the reduction of target language structures (Winford, 2003, p. 217). Bound morphology is one of the first elements learners tend to reduce. It could be argued that an example of simplification could be the use of invariant verb forms in Afro-Bolivian Spanish, a verb system in which all person and number morphology have been reduced to one form.

Internal developments constitute the third category described by Winford (2003, p. 220), which are strategies found not only in learners of second languages but in L1 acquisition, as they are universal language tendencies. These developments involve overgeneralizations in the L2 grammar and isomorphism, the use of one invariant form with one invariant meaning. An example of isomorphism is the use of an invariant verb form which only takes on the lexical meaning of the verb and does not communicate any information about the subject, such as person or number. Since learners tend to avoid morphological distinctions, such as verb inflections, other strategies may be used to compensate for the invariant form. This idea relates to the FH, which hypothesizes that learners will use explicit subject pronouns to express who the referent is due to the lack of number marking in the verb morphology.

The development of a second language via natural SLA and group language shift, as in the case of MCS, can result in the incomplete acquisition of some features of the target language, based on the processes just described (Winford, 2003, p. 222). As a result, certain elements of the interlanguage may fossilize, or cease to be acquired in a target-like manner. Sessarego (2011, 2012) uses the term crystallization (similar to fossilization) in regards to the use of invariant verb forms in Afro-Bolivian Spanish. He suggests that the lack of subject-verb agreement found in Afro-Bolivian Spanish is due to imperfect SLA (Sessarego, 2011, p. 126). Sessarego's theory explains the use of one invariant verb form in Afro-Bolivian Spanish as the result of processing difficulties that L2 learners experience at the syntax/semantics interface.

In this dissertation, I will argue that the variation in number marking on verbs in MCS is due to incomplete language acquisition. I will use the interlanguage strategies described by Winford to discuss how this variation came to develop, as they tend to overlap. I will also discuss how the results of partial language acquisition in MCS has led to a use of subject pronouns in this variety that diverges from NS.

The unique contact situation where MCS is spoken, both in the past and present, provides important information on the effects of contact on morphosyntactic features in Spanish. MCS represents a situation, in which adult speakers acquired a second language due to social or political pressure (García, 1996;Escobar, 2013;Wilson Withe et al., 2012), a common social factor that encourages language shift (Winford, 2003, pp. 257-258). This nativization process, via group language shift, has played an important role in the development of MCS, due to the fact that speakers were isolated to some degree from the target language in question. This sociohistorical difference from the speakers of NS, is

important to keep in mind when analyzing the two linguistic features discussed in this dissertation.

Chapter 4

Methodology

In this chapter I will give an overview of the methodology used in my study. This will include an explanation of my research questions and objectives, as well as a description of my participants, instruments, and analysis.

4.1 Research Questions and Hypotheses

The current dissertation examines two morphosyntactic features of MCS in comparison with NS, analyzing the correlation between contact, number marking, and the expression of pronominal subjects. The narrow objectives include (a) determining which linguistic and social factors motivate the distribution of these two features, and, most central to the goal of the dissertation, and (b) determining what processes explain the data. Different explanations related to contact, as well as second language acquisition and language shift, were explored with the goal of proposing one that explains the results of both features in MCS. The broad objective of the dissertation is to provide more information about Central American varieties of Spanish, whose representation is greatly lacking in the literature, while also exploring the effects of bilingualism and language contact. The research questions guiding this dissertation as well as the hypotheses are described below.

1. How do the two participant groups compare in regard to the linguistic factors that predict explicit subject pronoun expression and singular number marking on 3pl. verbs?

Based on studies related to these two morphosyntactic features under investigation in other varieties of Spanish, I predict that my two participant groups will have different results for both subject pronoun expression and variable number marking. Subject pronoun expression is well-documented in the variationist literature. Overall rates for the production of explicit pronouns for non-contact

varieties of Spanish vary and tend to follow a lowland/highland split, with lowland speakers producing more explicit pronouns than highland speakers. Based on this tendency, I predict that speakers of NS will fall in the middle of the continuum but will pattern more closely with lowland varieties, based on their geographical location and the characteristics they share with Caribbean Spanish, such as /s/ reduction.

Contact varieties of Spanish are generally characterized by high rates of subject pronoun expression, especially when the contact language is a non-Romance language (Lapidus and Otheguy, 2005; Otheguy et al., 2007; Michnowicz, 2015). Based on this general trend, I would expect speakers of MCS to produce even higher rates of explicit subject pronouns, compared to speakers of NS. In terms of the factors that will favor the use of explicit pronouns, I believe grammatical person, TMA, and coreference will have an effect. This prediction is based on previous studies (Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008; Michnowicz, 2015) that have shown that overt pronouns are used more often with singular verbs (particularly 1sg.), less-distinctive tense/moods (imperfect and subjunctive), and in contexts of switch reference.

My predictions in regard to variable number marking are based on findings from Critchfield and Lívio (*forthcoming*), as this is the only study that discusses this phenomenon for MCS. They find that lack of number agreement between 3pl. subjects and their corresponding verb occurs in 21% of their data. I expect to encounter a similar rate in the data set used for this dissertation. I also predict that I will only find evidence for this phenomenon in MCS as it has never been attested in NS.

I rely on literature on variable number marking in Brazilian Portuguese to help guide my analysis since this topic has been studied in depth from a variationist perspective in this language. Based on previous findings by Critchfield and Lívio (*forthcoming*), I expect syntactic order of the subject and the verb, intervening material, and phonic salience to have an effect on agreement (Naro and Scherre, 2000). VS word order, presence of intervening material, and low phonic salience should favor lack of agreement (see 3.2 for discussion on these variables in BP and MCS).

2. What social factors predict patterns in subject pronoun expression and variable number marking in MCS?

The lack of research addressing social aspects of language in Nicaragua and in Central America as a whole, reflects the necessity for more studies done from a sociolinguistic perspective in this region. While the existing literature helps to provide a foundation for the study of social evaluation in the Spanish of Central

America, more research is needed in order to draw clear conclusions about the effect of various social factors on linguistic choices.

Of the few studies that include social factors in Spanish varieties of Nicaragua, the most important predictor has been formality of the setting or speech act (see Rey, 1995, Rey, 1997 for the study of the *voseo* in NS; see Chappell, 2013, 2014, 2015a, 2016 for the study of /s/ deletion in the noun phrase in NS and MCS). Results have shown that many features are not socially stratified in NS; however, according to Lafford (1986), social stratification is more common for features that are widespread, such as /s/ reduction.

Previous literature in Spanish has found that social factors do not tend to motivate the use of explicit subjects in monolingual varieties (Silva-Corvalán, 2001, p. 163). Based on this finding, I do not expect to find social factors as significant predictors of subject pronoun expression in NS; however, when comparing both participant groups, I believe it could be possible to see an effect of bilingualism in the analysis, with MCS speakers producing higher rates of subject pronouns. Previous research shows gender as a significant factor, with women producing higher rates of explicit subject pronouns, specifically in contact with English (Shin, 2013). Since, the contact language is different in this study, this research is not necessarily a good predictor for MCS.

In contrast, I believe that social factors will act as significant predictors in variable number agreement. Based on findings for Brazilian Portuguese, I predict that level of education and gender will play a role in the variation observed in MCS. Male speakers and participants with lower levels of education have been more likely to produce non-agreement in BP. I expect level of education to be particularly important for MCS speakers because it is connected to exposure to Spanish, the language used in educational settings in Nicaragua (Wilson Withe et al., 2012). Winford (2003, p. 225) states that access to the target language and interaction with target language speakers are important social constraints involved in language acquisition. In an area of Nicaragua where there is little face-to-face contact with speakers of NS, formal education provides one area where speakers are exposed to prescriptive language. These results would reflect what has been found in other studies in Central American, but also more broadly in the sociolinguistics literature: non-standard forms are often stigmatized and occur within lower-educated groups, while standard variants tend to be used more with higher-educated speakers. I also predict that my findings for gender in MCS will correlate with what has been attested in BP and in sociolinguistics as a whole: women tend to produce higher rates of standard variants while men produce higher rates of non-standard variants (Labov, 2001; Silva-Corvalán, 2001).

3. Do the results for SPE and number marking in MCS show evidence for incomplete acquisition due to language contact? If so, what learner strategies best account for the results for both features?

Due to the sudden shift to Spanish that took place in the Miskitu community starting in the second half of the twentieth century, I predict the data will show evidence of incomplete language acquisition, which occurs when learners do not acquire linguistic features in their full, target-like manner because of a lack of input (Montrul, 2006). The geographic boundary separating the Pacific NS-speaking population and the Miskitu community located on the Caribbean coast, prevented any sort of robust or constant contact between the two groups (Floyd, 1967, pp. 1-16). In addition to the physical barriers that existed, communication with and travel to the Pacific was historically, very limited. A reliable highway connecting the two sides of the country was not completed until the early 1980s, and directly following the Sandinista revolution, the only line of communication between Puerto Cabezas, where the MCS data was collected, and the capital of Managua, was via telegraph (*The Miskitos in Nicaragua*, p. 3). While Spanish was initially brought to the Miskitu community via literacy campaigns, outside of an educational setting, learners did not receive direct input from NS speakers. Based on the history of contact between Miskitu and Spanish, as well as the linguistic differences that exist in these two languages, I expect distinct findings for MCS and NS in regard to the distribution of explicit subject pronouns and number marking on verbs.

Since I predict evidence of incomplete acquisition, I also expect to identify the strategies learners employ in the face of limited target language input. According to Montrul ((Montrul, 2006), learners show signs of incomplete acquisition of specific features of the target language via transfer from the L1 or via fossilization. Winford (2003) discusses additional processes evident in learner interlanguage grammars: L1 influence, simplification, and internal developments. In my discussion of incomplete language acquisition of MCS, I will address resulting interlanguage strategies used by learners, and whether or not there are signs of processes such as L1 transfer, fossilization, and simplification.

4.2 Methods

4.2.1 Participants: SPE

This portion of the study contained two participant groups: 1) 10 monolingual speakers of NS and 2) 10 bilingual (Miskitu-Spanish) speakers of MCS. The ten

monolingual speakers acted as a control group. All participants lived in cities surrounding the capital of Managua. Since these speakers' data was not analyzed for social factors, the distribution of gender, age, and educational background was not taken into consideration. The NS speakers were all monolingual and had no contact with other languages, including Miskitu or MCS.

The MCS speakers lived in the city of Puerto Cabezas, on the Caribbean coast. All participants self-identified as Miskitu-dominant or as a balanced Miskitu-Spanish bilingual. No speakers who learned Miskitu after Spanish, or as a second language, were included. The group of MCS speakers was distributed in terms of social categories, including age, gender, level of education, and age of acquisition of Spanish(4.1).

Table 4.1: Demographic Distribution of MCS Participants for SPE Analysis

	Number of Participants
Age	
18-29	5
30 +	5
Gender	
Female	5
Male	5
Level of Education	
Secondary	5
University	5
Age of Acquisition	
Child	7
Adolescent	3

4.2.2 Participants: Variable Number Marking

The MCS participant group used for the analysis of variable number marking consisted of 20 individuals from the Caribbean coast, who either self-identified as Miskitu-dominant or as a balanced Miskitu-Spanish bilingual. No speakers who learned Miskitu after Spanish, or as a second language, were included. The group of MCS speakers was distributed in terms of social categories, including age, gender, level of education, and age of acquisition of Spanish (4.2).

Table 4.2: Demographic Distribution of MCS Participants for Variable Number Marking Analysis

	Number of Participants
Age	
18-29	11
30 +	9
Gender	
Female	10
Male	10
Level of Education	
Secondary	9
University	11
Age of Acquisition	
Child	15
Adolescent	5

Bilingual Information of MCS Participants

An official bilingual profile was not used as part of the methodology for this study; however, participants were asked a variety of questions about their linguistic background. Their answers to these questions can be used to provide a better understanding of the bilingualism of the MCS speakers and their language use.

During the sociolinguistic interview, participants were asked what languages they spoke. All 20 individuals stated that they spoke Miskitu and Spanish, but no other languages. They were also asked when and where they learned both languages. All participants learned Miskitu at home with their families as babies but the details surrounding their acquisition of Spanish varied. 15 of the 20 speakers learned Spanish either at home from a family member or upon entering primary school. This group reported learning Spanish before the age of 6. The other 5 speakers learned Spanish in school upon entering secondary school but had little contact with the language at home. Based on these answers, Miskitu would be considered the first language for the MCS participants and Spanish would be either an additional first language for those with a more balanced level of bilingualism, or a second language.

The final question that was asked of the speakers during the interview was related to the situations in which they tend to use Miskitu and Spanish. All participants stated that they use Miskitu at home with friends and family, while Spanish is used depending on who they are with. While some seem to use Span-

ish as a way to accommodate non-Miskitu speaking people: *"el español también, bueno, lo usamos siempre porque siempre tenemos amigos españoles"* 'Spanish also, well, we always use it because we always have Spanish friends', others seem to use Spanish only when it is truly necessary for communication: *"yo casi hablo más el miskito y una persona me habla en español, hablo en español"* 'I speak mostly Miskitu and if a person speaks to me in Spanish, I speak in Spanish'. Many said they base their language on who they are talking with. If the person does not speak Miskitu, they will use Spanish but if the person does speak Miskitu, that is the preference. Spanish was discussed as being a language used at school and in other public spheres, while Miskitu could be used both at home or in public if the interlocutor also spoke Miskitu.

4.2.3 Instruments and Data Collection

There were two different instruments used. The data for the analysis of SPE came from two narration tasks while the data for the analysis of variable number marking came from sociolinguistic interviews. Due to travel safety restrictions, the data was collected at different times.

The narration task contained two parts. First, participants were asked to narrate a children's picture book which had the words removed. The book, *Here Comes the Strikeout* (Kessler, 1992), tells the story of a boy who wants to learn to play baseball. The portion used for the narration task was 28 pages in length and only contained illustrations. For the second part, participants were asked to tell a cultural legend or story from memory. The narration task was designed to elicit third person forms, as this is what was analyzed for SPE. I collected the narrations during the summer of 2016 in the cities of Puerto Cabezas, Diriamba, and Masaya. The NS speakers were recorded in their respective homes. All participants were people I knew well while living in Nicaragua; therefore the formality of the setting was very familiar. The MCS speakers were recorded in a variety of locations including the homes of my participants and a local hotel. Each individual set of narrations ranged from two to five minutes in length. I recorded both parts of the narration task and later transcribed each orthographically. These transcriptions were later reviewed for accuracy by a speaker of NS.

The sociolinguistic interviews conducted with the MCS speakers included questions specific to their cultural background speakers were collected during the summer of 2017 in Puerto Cabezas, Nicaragua. In addition to the oral interview, participants were asked to fill out a short demographic questionnaire. The location of the interview varied depending on what was available, including the homes of my participants, a local hotel, and the city market. For all of the

interviews, my research assistant, a local MCS speaker, was present and knew the interviewee personally, resulting a familiar setting in terms of formality. The individual interviews ranged from 15-45 minutes each and totaled approximately 10 hours all together. Half of the recordings were transcribed by me and the other half were transcribed by a speaker of NS. All were reviewed for accuracy by both parties.

4.2.4 Coding and Analysis of Data

In order to address my three research questions, I used a variationist approach to analyze the oral data collected from the two participant groups. I performed a logistic regression on my data to determine which predictors motivated the use of explicit subject pronouns and singular number marking by NS speakers compared to MCS speakers. The data was subjected to both a binomial logistic regression and mixed-effects logistic regression to model the relationship between the categorical response variable (explicit vs. implicit subject pronoun; singular number marking vs. plural number marking) and the predictive social and linguistic factors.

For the analysis of subject pronoun expression, all cases of finite verbs that permitted both an explicit or an implicit subject pronoun were included. No cases of other subject forms, such as noun phrases or proper names, were included. Only third person forms were examined. Cases that were excluded from the analysis were the following: verbs like *gustar*, *dicen que* and other generalized third person plural constructions, and quoted speech. Previous studies suggest that Spanish speakers use first and second person pronouns at a much higher rate than third person pronouns, providing evidence for a functional distinction between person and number (Ranson, 1991; Davidson, 1996). It has therefore been suggested that pronouns should be analyzed separately (Lozano, 2009; Quesada and Blackwell, 2009). I coded for a variety of linguistic factors, many of which have been found to be significant predictors in previous studies on Spanish, including number of the referent (3sg. or 3pl.)[3] , TMA of the verb (more or less distinctive), coreference (same or switch), and reflexivity (reflexive or non-reflexive). Since social factors have not proven to motivate the use of subject pronouns in monolingual varieties, I only included these variables in the MCS analysis: age (18-29 or 30+), gender (male or female), level of education (secondary or university), and age of acquisition (child or adolescent). In total, 347 tokens were analyzed for NS and 436 tokens were analyzed for MCS.

[3] The coding of number was based on the pronoun unless it was implicit in which case it was based on the verb form.

For the analysis of variable number marking, all the explicit 3pl. subjects and their corresponding verbs were extracted from the transcribed data. All tokens that were followed by a nasal or involved verbs like *gustar*, were excluded from

the analysis. The coding was based on previous predictors found to be significant in a similar phenomenon in Brazilian Portuguese (Mendes and Oushiro, 2015; Oushiro, 2015): phonic salience of the verb (high or low salience), subject position relative to the verb (immediately preceding, distantly preceding, or postponed), subject type (pronoun or noun phrase), and animacy of the subject ([+human] or [-human]). The social factors that were included in the analysis were age (18-29 or 30+), gender (female or male), level of education (secondary or university), and age of acquisition (child or adolescent). 595 tokens were analyzed.

4.2.5 Description of Linguistic Factors for Subject Pronoun Expression

In this section, I will provide a short description of each factor used in the analysis of subject pronoun expression.

Number

Number was coded based on the person and number of the referent. Since only third person forms were analyzed the two levels were 3sg. and 3pl. Due to variable number marking that exists in MCS, the pronoun was used to determine number unless it was implicit, in which case the number of the verb was used. There were some cases where the number of the pronoun and verb did not match (4.1). I expect 3sg. verb forms to favor the use of explicit pronouns for both groups of speakers, based on findings for other varieties of Spanish (Otheguy, Zentella, and Livert, 2007; Michnowicz, 2015).

(4.1) *Ella$_i$ est-aba dormida en la cama de los*
she be-PST.IPFV.3SG. asleep in the bed of the
duendecitos. Y se hiz-o$_i$ amiga con ellos$_j$.
elves and REFL become-PST.3SG friend with them
Ello − s$_j$ nunca se bañ-aba.
they-PL never REFL bathe-PST.IPFV.3SG
'She was asleep in the elves' bed. And she became friends with them. They never bathed.'

TMA of Verb

The tense, mood, and aspect of the verb was divided into two groups: less distinctive and more distinctive, based on the coding from previous studies (Orozco and Guy, 2008; Michnowicz, 2015). Less distinctive referred to cases

where 1sg. and 3sg. forms are the same including present indicative, present subjunctive, imperfect indicative, and imperfect subjunctive. All other TMAs were considered to be more distinctive. I expect less distinctive forms to result in higher rates of explicit subject pronoun production by both participant groups (Orozco and Guy, 2008; Michnowicz, 2015).

Coreference

Data was coded for coreference as being either same or switch reference. In (4.2), *el amigo* 'the friend' is an example of a switch from the previous referent, as is *él* he, referring to the main character of the story. This character continues to be the subject of the following verb *podía*, an example of same reference. Following the methodology used by Geeslin and Gudmestad, coreference was based on the subject of the closest preceding tensed verb that occurred in a clause (Geeslin and Gudmestad, 2011). The coding of the data from the narration of the book excluded any narration by or reference to the speaker such as *parece que* or *veo*. I expect MCS speakers to produce more explicit subject pronouns in cases of switch reference; however at lower rates compared to NS speakers (Cameron, 1992; Silva-Corvalán, 1994; Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008; Michnowicz, 2015).

(4.2) *Y el amigo$_i$ de nuevo le estaba*
and the friend of new him be-PST.IPFV.3SG
enseñ-ando como bate-ar y él$_j$ se puso
teaching-GER how bat-INF and he REFL become-PST.3SG
muy triste porque Ø no podía$_j$ bate-ar.
very sad because (he) no can-PST.IPFV.3SG bat-INF
'And the friend was teaching him how to bat again and he became very sad because he couldn't bat.'

Reflexivity

All tokens were coded as containing a reflexive verb or not. Previous studies that examine SPE across grammatical person and number have found non-reflexive verbs to result in higher rates of explicit subject pronoun production due to the additional disambiguation reflexive pronouns provide (Abreu, 2012; Otheguy and Zentella, 2012; Michnowicz, 2015). The current dissertation only includes the analysis of third person forms, which both utilize *se* as the reflexive pronoun. While this form is the same for both 3sg. and 3pl., I still expect non-reflexive verbs to significantly predict explicit subject pronoun use based on the findings of the previously mentioned studies.

4.2.6 Description of Linguistic Factors for Variable Number Marking

In this section, I will provide a short description of each factor used in the analysis of variable number marking, including examples from the data set.

Phonic Salience

In this dissertation, phonic salience is related to the morpheme on the verb that marks number, either singular or plural. The designation of low phonic salience was used in cases where the plural form differed from the singular form by only a final nasal consonant (*trabajan* vs. *trabaja* in 4.3, 4.4). The designation of high phonic salience was used in cases where there was a greater perceptible distinction between singular and plural forms (*dijeron* vs. *dijo* in 4.5, 4.6). I expect verb forms with low phonic salience to result in more cases of non-agreement (Naro and Scherre, 2000; Mendes and Oushiro, 2015; Critchfield and Lívio, *forthcoming*) (for a more thorough explanation of phonic salience, see section 3.2.1).

(4.3) *Los* *hombre-s* *trabaj-an*
the.PL man-PL work-PRS.3PL
'The men work'

(4.4) *Ellos* *trabaj-a* *de* *la* *agricultura* *ganadería*
they work-PRS.3SG of the agriculture ranching
'They work in agriculture, ranching'

(4.5) *Ellos* *le* *dij-eron* *que* *no* *les* *llev-ara*
they him tell-PST.3PL that no them take-PST.SBJV.3SG
'They told him not to take them'

(4.6) *Entonces* *los* *pajarito-s* *dij-o* *que*
then the.PL bird-PL say-PST.3SG that
hab-ia *un* *techo*
there-was-PST.3SG a roof
'Then the birds said that there was a roof'

Subject Position

Subject position was coded based on the syntactic order of the subject in relation to the verb, as well as the presence of linguistic material between the two: immediately preceding (4.7), distantly preceding (4.8), and post posed (4.9)[4].

[4] The term 'post posed' was used to refer to post verbal position of the subject.

Initially there was a preceding category (tokens contained one to four syllables between the subject and the verb) as well as a distantly preceding category (tokens contained five or more syllables between the two 4.10); however these were combined due to the small number of distantly preceding tokens in the data set. The coding of levels for this variable was adopted from Oushiro (2015). Based on her analysis and other research on BP and MCS (Critchfield and Lívio, *forthcoming*; Naro and Scherre, 2000), I expect distantly preceding and post posed subjects to favor singular number marking.

(4.7) *ellos viv-en allí*
they live-PRS.3PL there
'They live there'

(4.8) *ellos nunca aguant-aba de nada*
they never tolerate-PST.IPFV.3SG of nothing
'They never tolerated anything'

(4.9) *vien-en mi-s prima-s*
come-PRS.3PL my cousin-PL
'My cousins come'

(4.10) *muchos de mi raza miskita habl-an el inglés*
many from my race Miskitu speak-PRS.3PL the English
'Many from my Miskitu race speak English'

Subject Type

Subject type was divided into the following categories: pronoun (personal, indefinite, relative, or demonstrative) (4.11) and noun phrase (both simple and compound). I expect the use of noun phrases to result in more cases of non-agreement based on Oushiro's (2015) findings for BP.

(4.11) *ellos mand-aban en el parlamento*
they send-PST.IPFV.3PL in the parliament
'They were in charge in the parliament'

(4.12) *mi-s ancestro-s and-aba en el mar*
my.pl ancestor-PL go-PST.IPFV.3SG in the sea
'My ancestors went around in the sea'

Animacy

Subjects were coded for animacy, using the following designations: [+human] and [-human]. This variable only distinguished humanness and not animacy due to the small number of [-human, +animate] tokens.[5] Subjects that were human were coded as such (4.13), while subjects that were non-human were combined regardless of being animate (4.14) or inanimate (4.15). I expect [-human] subjects to favor singular marked forms (Mendes and Oushiro, 2015).

(4.13) *la-s* *mujer-es* *miskita-s* *nac-ieron* *para*
the-PL woman-PL Miskitu-PL born-COND.3SG for
tener *hijos*
have.INF children
'The Miskitu women were born to have children'

(4.14) *liwa mairin* *er-a* *una* *parte* *mujer* *y* *un*
mermaid.SG be-PST.IPFV.3SG a part woman and a
parte *de* *pescado...* *ellos* *viven* *más* *en* *el* *profundo*
part of fish... they live-PRS.3PL more in the depths
de-l *mar*
of-the sea
'Mermaid was one part woman and one part fish...they live mostly in the depths of the sea'

(4.15) *la-s* *bebida-s* *típica-s* *ser-ía* *chicha*
the-PL drink-PL typical-PL be-COND.3SG chicha
'The typical drinks would be chicha'

4.2.7 Social Factors for MCS Analysis

Social factors will be analyzed in the MCS data for both SPE and variable number marking. Research on SPE has generally found no effect for social factors; however, I will include the following variables in order to provide a more robust analysis of the distribution of explicit subject pronouns in MCS. Regarding number marking, I expect males and speakers with lower levels of formal education to favor singular marked forms (see Mendes and Oushiro, 2015, as well as section 3.2.1 for further discussion). These predictions are based on findings for BP (age and age of acquisition have not been researched).

Age

Age was divided into two categories: 18-29 and 30+. Participants shared their age on the demographic information questionnaire and were assigned to the

appropriate group. The separation of age groups was done in order to evenly distribute speakers to fall on both ends of a spectrum. The groups were divided around the age of 30 because this corresponds with the counterrevolution that took place in Nicaragua in the late 1980s. Speakers born after the counterrevolution would have been required to speak Spanish in public school due to the Spanish literacy campaign promoted by the Sandinistas (García, 1996; Wilson Wither et al., 2012; Escobar, 2013).

Gender

Gender was divided into two categories: female and male. Participants were grouped based on how they answered a written demographic questionnaire filled out prior to being interviewed. In the form, this question was open and could be answered freely based on how each individual identified. No answers other than 'female' and 'male' were given.

Level of Education

Level of education referred only to formal education and was divided into two categories: secondary and university. Speakers were placed into the highest level they had completed or were currently at in their studies. Participants were asked to fill out written demographic information prior to their interview. They wrote what their level of education was at the time of the interview and were assigned to the corresponding category.

Age of Acquisition

Age of acquisition was divided into two categories: child and adolescent. Child was used for speakers who reported learning Spanish at or before the age of 6 (generally upon entering primary school) and adolescent was used for speakers who reported learning Spanish at or after the age of 12 (generally upon entering secondary school). This division is based on descriptions of early and late child bilinguals (Montrul, 2008, pp. 17-20). The participants were asked when they learned Spanish in the interview. Each answered orally and was placed in the corresponding category depending on the age they gave. This social category only applied to the bilingual speakers of MCS.

4.3 Summary

In this chapter, I outlined the methodology used in the dissertation, including a description of the participants, instruments, data collection, and coding. In

addition, I discussed my research questions and predictions for the data analysis results, based on findings and patterns from previous research. Both linguistic and social factors will be examined in the subsequent chapters as part of the data analysis of SPE and variable number marking.

Chapter 5

Data Analysis: Subject Pronoun Expression

In this chapter I will present the data analysis for subject pronoun expression. The discussion of the results includes an overview of the distribution of the two variants in the data set for NS vs. the data set for MCS, an explanation of the statistical models used, and a presentation of which factors predicted the production of explicit subject pronouns. In addition to four internal linguistic factors examined for NS and MCS, I analyzed an additional four external factors for the MCS data, as bilingual speakers have previously shown some sensitivity to social constraints and SPE.

5.1 General Frequency Results

The general frequency results (Figure 5.1) for NS showed that explicit subject pronouns occurred in 23.9% of the data set overall, with 83 occurrences out of 347. In comparison, speakers of MCS produced overt subject pronouns at a rate of 16.5%, with 72 occurrences out of 436.

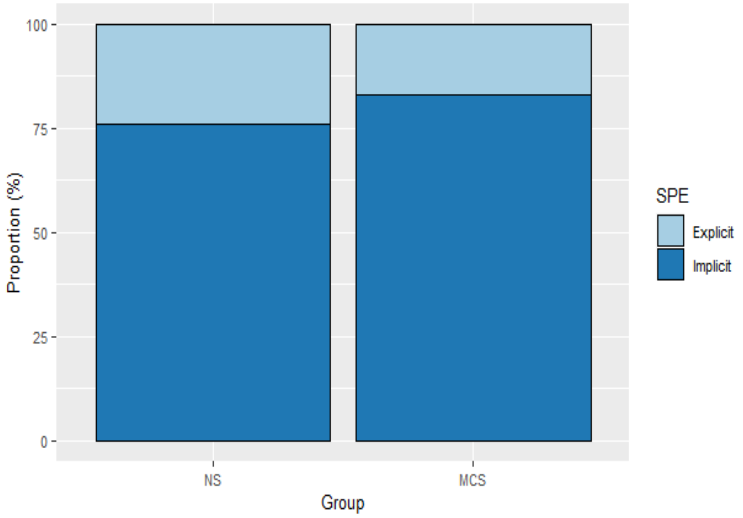

Figure 5.1: SPE by Speaker Group

Referring back to the trends previously found for monolingual varieties of Spanish (Table 3.1), NS speakers fell in the middle of the spectrum. Overall rates for the production of explicit pronouns in non-contact varieties of Spanish vary and tend to follow a lowland/highland split, with lowland speakers producing more explicit pronouns than highland speakers. The current results showed that the NS speakers patterned more closely with highland varieties despite geographical location and shared characteristics with Caribbean Spanish, such as /s/ reduction (Hochberg, 1986a, b; Cameron, 1993, 1996; Alba, 2004; Foote and Bock, 2012; Guy, 2014, p. 446; Chappell, 2014, 2016).

Previous studies show a tendency for speakers of contact varieties of Spanish to produce higher rates of explicit subject pronouns compared to speakers of non-contact varieties. MCS speakers in this data set do not coincide with this general trend, as they produce explicit pronouns at a lower rate than the NS speakers.[6]

[6] It is important to note that this study only analyzes the context of third person. It is possible that the overall pronoun rate may change if first person and second person tokens were included.

5.2 Model Selection: Part 1

The data used for this analysis was comprised of a categorical dependent variable: implicit or explicit subject pronoun expression, with 347 total tokens in the NS data set and 436 total tokens in the MCS data set. Each token was coded for four different independent linguistic variables: number, TMA, coreference, and reflexivity.

I started the model selection by running initial logistic regression models in R using lrm() for both data sets. This output was used to check the overall

significance and goodness-of-fit of the models with the four linguistic variables. The results showed that overall the models had predictive power and were significant, based on small *p*-values of the Model Likelihood Ratio Test (<0.0001 for NS and 0.0068 for MCS). In terms of goodness-of-fit, the concordance index C was .749 for the NS data set and .616 for the MCS data set, both of which signal discrimination (Hosmer and Lemeshow, 2000, p. 162).

Next, I reviewed the output of the model to see if any of the factors should be removed. I checked for any signs of data sparseness or multicollinearity by examining the standard errors for each object in the models. All of the standard errors were small, showing no evidence of these issues. I used a generalized linear regression to run three different stepwise tests: forward, backward, and bidirectional. All results showed that only number and coreference significantly improved the model for NS, while only coreference significantly improved the model for MCS. This was evaluated using the AIC, which goes down as a model becomes more significant. The NS model's AIC with TMA and reflexivity included as predictors was 344.9 and without them was 342.89. The MCS model's AIC with number, TMA, and reflexivity included as predictors was 386.57 and without them was 386.25. In order to determine whether the difference in AIC values was significant, I ran an ANOVA with the two models for each data set. The results showed that the difference in the AICs was not significant so I decided to keep all four factors in both models. Finally, I looked for possible interactions between my independent variables and did not find anything statistically significant.[7]

[7] $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

My two final models (Table 5.1 for NS and Table 5.2 for MCS) were mixed-effects logistic regressions with speaker as a random factor and number, TMA, coreference, and reflexivity as fixed-effects (Table 5.1). The final model for NS had an AIC of 338.9 and showed number and coreference being significant predictors, while the final model for MCS had an AIC of 368.1 and showed coreference as being the only significant predictor (based on low *p*-values).

Table 5.1: Mixed-Effects Logistic Regression Predicting Explicit SPE in NS: Linguistic Factors (n=347, AIC=338.9)

Variable	Estimate/Odds	SE	P r(>|z|)	n	%Explicit
Pronoun expression					
(vs. Implicit)				264	76.1%
Explicit	-3.99/.019	.74	9.66e-08 ***	83	23.9%
Number					
(vs. Plural)				51	9.8%
Singular	1.57/4.79	0.53	.00292 **	296	26.3%
TMA					
(vs. More distinctive)				196	21.9%
Less distinctive	0.09/1.09	.29	.75588	151	26.5%
Coreference					
(vs. Same)				217	13.8%
Switch	1.67/5.31	.29	1.06e-08 ***	130	40.8%
Reflexivity					
(vs. Reflexive)				53	15.1%
Non-reflexive	.59/1.81	.46	.20235	294	25.5%

Table 5.2: Mixed-Effects Logistic Regression Predicting Explicit SPE in MCS: Linguistic Factors (n=436, AIC=368.1)

Variable	Estimate/Odds	SE	P r(>|z|)	n	% Explicit
Pronoun expression					
(vs. Implicit)				364	83.5%
Explicit	-1.84/.16	.56	.00112 **	72	16.5%
Number					
(vs. Plural)				85	24.7%
Singular	-.2/.82	0.34	.54869	351	14.5%
TMA					
(vs. More distinctive)				349	15.7%
Less distinctive	.02/1.02	.34	.94518	87	19.5%
Coreference					
(vs. Same)				290	12.8%
Switch	.87/2.39	.29	.00239 **	146	24%
Reflexivity					
(vs. Reflexive)				56	23.2%
Non-reflexive	-.34/.71	.38	.37327	380	15.5%

Tables 5.1 and 5.2 give information about the number of tokens observed for each level of any given factor ('*n*'), as well as the percentage of that number that were explicit pronouns ('% Explicit'). The coefficient estimates in Tables 5.1 and 5.2 were used to determine the simple log odds of the outcome. These estimates are shown for the levels listed (in comparison to the reference level). The reference level for the dependent variable was implicit SPE, meaning the model compared the probability of the second level (explicit SPE) with the reference levels (implicit, plural, more distinctive, same, reflexive). Exponentiating the coefficient gave the simple log odds of explicit SPE vs. implicit SPE (with the reference levels of the predictors): .019/1.9% probability for NS and .16/16% probability for MCS. When the reference levels were reset to reflect the categories that favored explicit pronoun production the most (for NS: singular, less distinctive, switch reference, non-reflexive; for MCS: plural, less distinctive, switch reference, reflexive), the odds increased to .937/93.7% for NS and to .805/80.5% for MCS.

The p-values for the independent variables in the NS model show that number and coreference were significant, which means there was no significant difference between more and less distinctive TMA, or between reflexive and non-reflexive verbs. For the MCS model, only coreference was significant, which means there was no significant difference between singular and plural verbs, between more and less distinctive TMA, or between reflexive and non-reflexive verbs. Any *p*-value smaller than 0.05 was considered to be the threshold for statistical significance. The effect size for each predictor, presented in the following sections, can be calculated by exponentiating the coefficient estimate. These numbers show how much more likely the level shown was to result in explicit SPE, compared to the reference level for that specific predictor.

5.3 Linguistic Factors

In this section I will discuss the findings for each of the four linguistic variables used in the analysis: number, TMA, coreference, and reflexivity.

5.3.1 Number

Number of the verb was coded as being either 3sg. or 3pl. This factor was only found to be a significant predictor in NS (p-value = .00292 **). The results showed that when compared to 3pl., contexts with 3sg. were 4.79 times more likely to predict explicit subject pronouns. In 3pl. contexts, NS speakers only used explicit pronouns 9.8% of the time while in 3sg. contexts, they used them

26.3% of the time (Figure 5.2). In comparison, MCS speakers produced less explicit pronouns with 3sg. (14.5% vs. 24.7% with 3pl.). Previous findings for both monolingual and bilingual speakers find higher rates of explicit pronoun use with singular referents (Otheguy, Zentella, and Livert, 2007; Michnowicz, 2015), therefore the MCS data here does not follow the general trend.

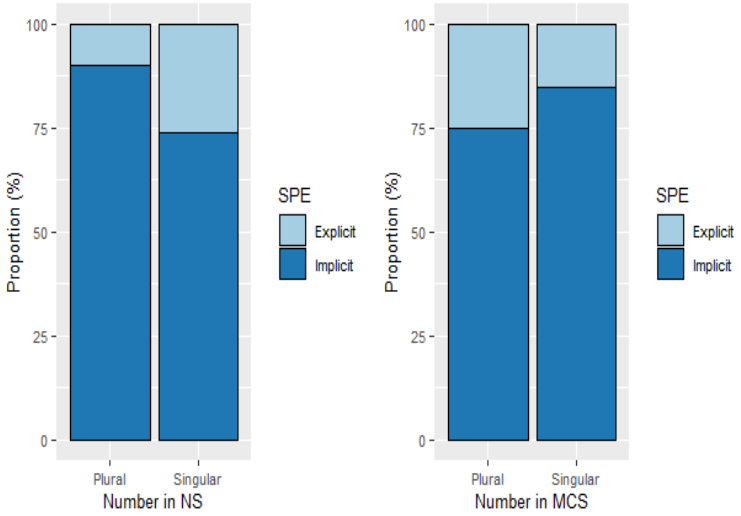

Figure 5.2: Number in NS vs. MCS

According to Miskitu grammars, singular subjects can be either explicit or implicit, while plural subjects are obligatorily explicit (CIDCA, 1985, p. 129; Salamanca, 1988, p. 252). The findings here for MCS may reflect this distribution in Miskitu. While the participants did use implicit pronouns with 3pl., it occurred at a lower rate compared to 3sg. (5.1).

(5.1) *Pero ellos$_i$ sí sab-ían que hacer. Ellos$_i$*
but they.PL yes know-PST.IPFV.3PL that do.INF they.PL
le dij-eron que no les llevar-a hacia
her tell-PST.3PL that no them take-SBJV.PST.3SG. toward
el lugar directamente. Ellos$_i$ se ib-an a
the place directly they.PL REFL go-PST.IPFV.3PL to
entrar al amanecer.
enter.INF at sunrise
'But they did know what to do. They told her not to take them to directly to the place. They were going to enter at sunrise.'

5.3.2 TMA

TMA was coded as being either less distinctive or more distinctive, based on morphological ambiguity between 1sg. and 3sg. forms. This factor was not significant in either of the data sets; however, participants produced more explicit pronouns with less distinctive TMA (for NS: 26.5% vs. 21.9% with more distinctive forms; for MCS: 19.5% vs. 15.7% with more distinctive forms) (Figure 5.3).

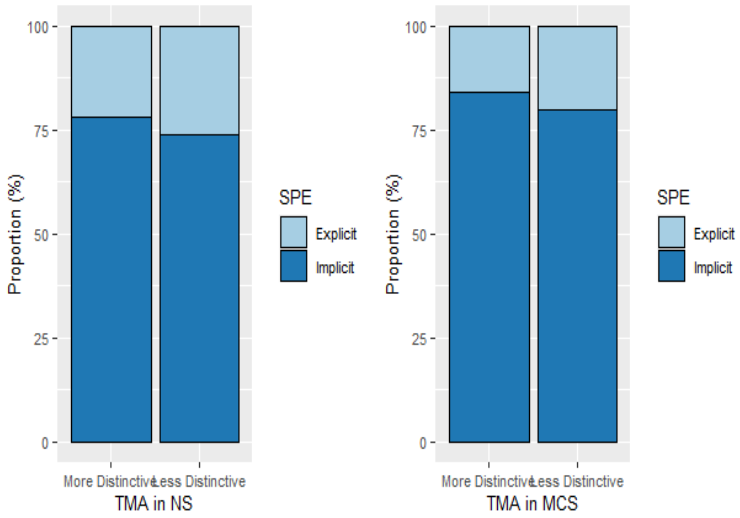

Figure 5.3: Verb Form in NS vs. MCS

This finding reflects what has been observed in both non-contact and contact varieties of Spanish (Orozco and Guy, 2008; Michnowicz, 2015), where less distinctive TMA motivates the use of explicit subjects as a way of avoiding potential ambiguity of the referent. Due to the context of the two tasks in the present study, 1sg. tokens did not occur, which may explain why less distinctive forms did not favor explicit pronoun production more strongly.

5.3.3 Coreference

Coreference was coded as being either same or switch. This was the only factor that was significant for both participant groups (p-value = 1.06e-08 *** for NS and p-value = .00239 ** for MCS). The results showed that when compared to contexts of same reference, in contexts of switch reference NS speakers were 5.31 times more likely to produce an explicit subject pronoun, while MCS speakers were 2.39 times more likely.

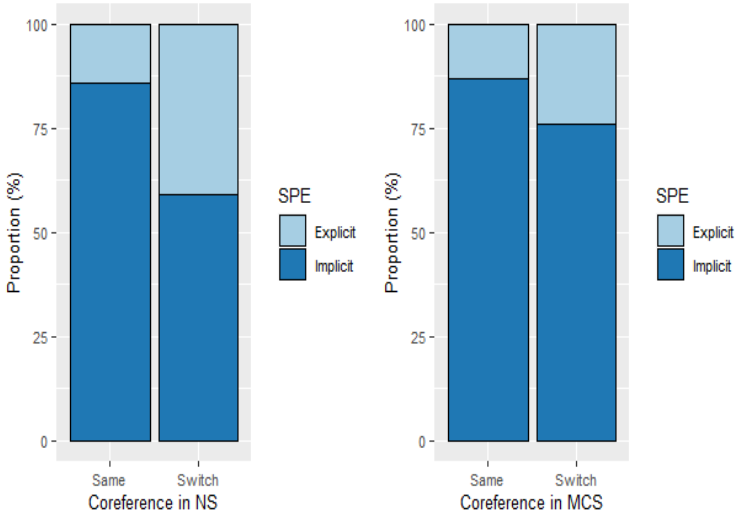

Figure 5.4: Coreference in NS vs. MCS

These results coincide with what has been found previously, with bilingual speakers of Spanish more likely to show a weakening of pragmatic constraints (Silva-Corvalán, 1994; Otheguy, Zentella, and Livert, 2007; Michnowicz, 2015). While the MCS speakers were still sensitive to the context of switch reference, they seemed to be less so compared to their NS-speaking counterparts: 24% explicit SPE with switch reference by MCS participants vs. 40.8% by NS participants (Figure 5.4). There were many examples in the data set where an explicit pronoun would be expected but was not produced, resulting in ambiguity for the listener (5.2).

(5.2) *Y el hombre$_i$ est-aba corr-iendo y la*
and the man be-PST.IPFV.3SG run-GER and the
mujer$_j$ se d-io cuenta que el hombre$_i$ no
woman REFL give-PST.3SG sum that the man no
est-aba en la casa y agarr-ó$_j$ al niño
be-PST.IPFV.3SG in the house and grab-PST.3SG the boy
y se pus − o$_j$ a llorar y a correr detrás
and REFL put-PST.3SG to cry.INF and to run.INF behind
del hombre.
the man
'And the man was running and the woman realized that the man was not in the house and grabbed the boy and started to cry and to run after the man.'

5.3.4 Reflexivity

Reflexivity was coded as reflexive or non-reflexive depending on the verb form used. This factor was not significant in either data set; however, NS data followed the general trend found in previous studies (Abreu, 2012; Otheguy and Zentella, 2012; Michnowicz, 2015): speakers produced more explicit pronouns with non-reflexive verbs (25.5% vs. 15.1% with reflexive verbs). Reflexivity tends to decrease the odds of producing an explicit pronoun due to the additional referential information encoded in the reflexive pronoun. MCS speakers diverged from this general trend, with explicit pronouns being used in 15.5% of cases with non-reflexive verbs and in 23.2% of cases with reflexive verbs (Figure 5.5).

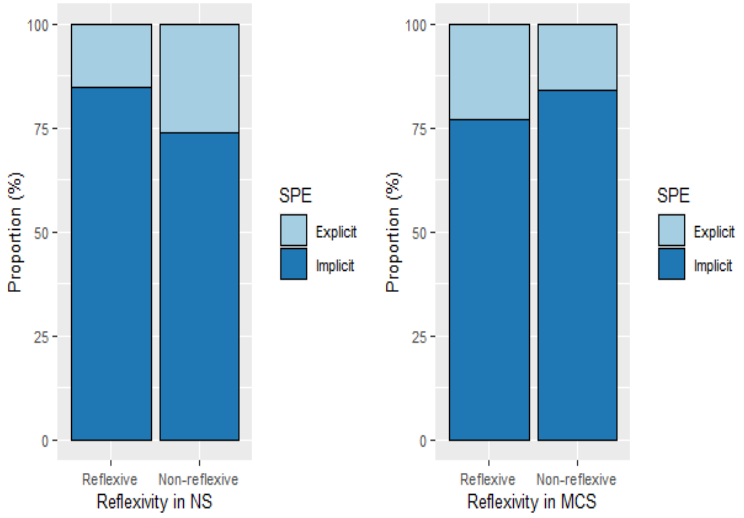

Figure 5.5: Reflexivity in NS vs. MCS

5.4 Summary of Results: Linguistic Factors

The results for the two participant groups showed very different behavior in regard to the production of subject pronouns. While not all factors were statistically significant in predicting explicit SPE, the NS speaker data aligned with what has been found in other monolingual varieties of Spanish, with explicit pronouns more likely to occur under the following conditions: singular referents (Otheguy, Zentella, and Livert, 2007, Michnowicz, 2015), less distinctive TMA (Orozco and Guy, 2008; Michnowicz, 2015), contexts of switch reference (Cameron, 1992; Silva-Corvalán, 1994; Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008, Michnowicz, 2015), and with non-reflexive verbs (Abreu, 2012; Otheguy and Zentella, 2012; Michnowicz, 2015). Unlike

NS speakers, MCS speakers did not align with all of these trends. The findings of the analysis showed that switch reference and less distinctive TMA were the only expected contexts where speakers were more likely to produce an explicit pronoun (albeit at a rate lower than NS speakers), a sign of weakened pragmatic constraints. The MCS speakers demonstrated a complete opposite effect of what was expected for the other two variables analyzed; they were more likely to produce explicit pronouns with reflexive verbs and plural referents. In addition, the MCS participants produced subject pronouns at a much lower rate overall when compared to the NS participants, a finding that diverges from patterns in other bilingual varieties of Spanish.

5.5 Model Selection: Part 2

In addition to the four linguistic factors analyzed, the MCS data was coded for four social factors: speaker age, gender, level of education, and age of acquisition. An initial logistic regression with these additional factors resulted in what appeared to be a stronger model as the concordance index C went up from .616 to .740 and the p-value of the Model Likelihood Ratio Test got smaller (<0.0001.). The AIC also went down from 368.1 to 358.7 with the social factors included in the model. However, I ended up removing age of acquisition from the final analysis because it was not contributing to the model (high *p*-values) and because an interaction between this factor and age was found. In addition, the maximum number of variables recommended for my data set was seven based on the number of tokens so I decided to only keep age, gender, and level of education in order to have a more accurate model.

The final model was a mixed-effects logistic regression with speaker as a random variable and number, TMA, coreference, reflexivity, age, gender, and level of education as fixed-effects (Table 5.3). The final model had an AIC of 361.6 and showed both age and gender as significant predictors. All linguistic factors were kept in the model in order to maintain consistency with the previous analysis. The favoring/disfavoring effect for each of these variables remained the same with the new social factors included.

Table 5.3: Mixed-Effects Logistic Regression Predicting Explicit SPE in MCS: Linguistic and Social Factors (n=436, AIC=361.6)

Variable	Estimate/Odds	SE	P r(>|z|)	n	% Explicit
Pronoun Expression					
(vs. Implicit)				364	83.5%
Explicit	-3.12/.044	.63	8.61e-07 ***	72	16.5%
Number					
(vs. Plural)				85	24.7%
Singular	-.28/.76	.33	.400390	351	14.5%
TMA					
(vs. More Distinctive)				349	15.7%
Less Distinctive	.19/1.2	.34	.579404	87	19.5%
Coreference					
(vs. Same)				290	12.8%
Switch	.83/2.29	.28	.003636 **	146	24%
Reflexivity					
(vs. Reflexive)				56	23.2%
Non-reflexive	-.2/.82	.38	.606441	380	15.5%
Age					
(vs. 30+)				230	15.6%
18-29	1.23/3.41	.33	.000221 ***	206	17.5%
Gender					
(vs. Female)				202	7.9%
Male	1.68/5.38	.41	3.60e-05 ***	234	23.9%
Level of Education					
(vs. University)				266	21%
Secondary	-.45/.63	.36	.214215	170	9.4%

With the reference levels set to reflect the categories that favored explicit pronoun production the most (plural, less distinctive, switch reference, reflexive, 18-29, male, university), the odds were 1.42/142%. Under these conditions, explicit SPE was 1.42 times more likely than implicit SPE, an increase from the original model.

5.6 Social Factors

In this section, I will discuss the findings for social factors in MCS. These variables were only analyzed with the bilingual data set, as previous research shows that monolingual varieties of Spanish are not constrained by these variables.

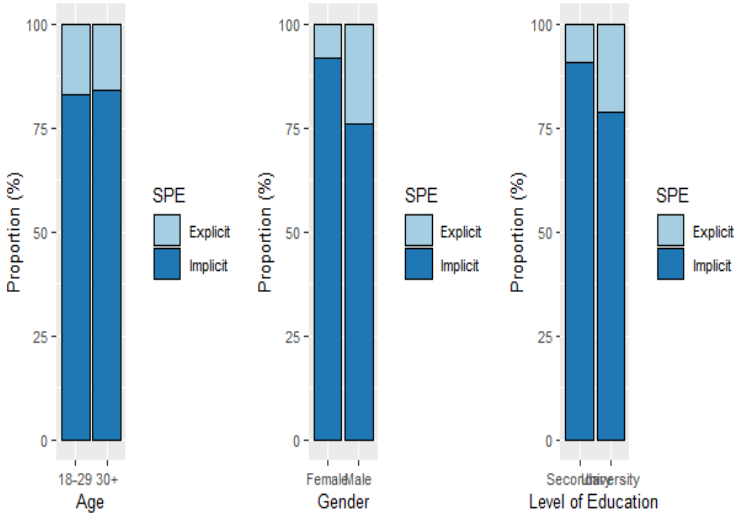

Figure 5.6: Social Factors for SPE in MCS

5.6.1 Age

Age was coded into two groups: 18-29 and 30+. This was the second most significant predictor in the model (p-value = .000221 ***), with speakers 18-29 years of age being 3.41 times more likely to produce explicit pronouns, compared to speakers 30 and over. While both age groups had similar rates of production, the younger participants in the data set were more consistent in their use of explicit subject pronouns across speakers. As the model accounted for speaker as a random variable, this consistency contributed to the results of statistical significance, despite the rate of explicit SPE. The finding for age indicates that younger speakers are approximating more to rates found in NS. The younger speakers in the data set acquired Spanish at an earlier age than older speakers. While participants 30 and over learned Spanish at ages 6, 8, 10, 12, and 13, all participants 18-29 reported learning Spanish as infants, apart from one speaker who reported learning Spanish at age 16. This early exposure to Spanish, particularly prescriptive language taught in the education system, may explain the results for this social factor.

5.6.2 Gender

Gender was divided into female and male. This was the most significant predictor of explicit pronoun production in the model (p-value = 3.60e-05 ***), with male speakers being 5.38 times more likely to use explicit pronouns compared to female speakers. This finding does not coincide with previous studies on bilingual SPE or with general sociolinguistic trends (Silva-Corvalán, 2001; Shin, 2013). Women tend to be either leaders of linguistic change or use more standard forms. In the context of this study, increased explicit pronoun production, closer to NS rates would reflect these tendencies; however, the females speakers in the data set showed the opposite effect. Another possible explanation could be that men are more mobile and therefore more exposed to NS due to the necessity to look for work away from the Caribbean Coast. This explanation has also been used to explain standard language use by men in BP (Rodrigues, 1987; Mendes and Oushiro, 2015), in addition to be used more broadly in sociolinguistic theory (Labov, 2001).

5.6.3 Level of Education

Level of education was coded as secondary or university. The logistic regression did not show this variable as being a significant predictor of SPE; however, the negative estimated log odds (-.45) for secondary, showed speakers with a lower level of education were less likely to produce explicit subject pronouns. The less formal education a speaker had, the more they diverged from the SPE rate of NS speakers. Three of the five speakers with a secondary level of education produced explicit pronouns at extremely low rates (8.5%, 6.25%, 3.23%). The other two were closer to the average for NS speakers but still fell below 23.9% (20.83% and 18.75%). While level of education is not necessarily reflective of level of bilingualism, speakers with higher levels of formal education would have more exposure to prescriptive language norms.

5.7 Summary of Results

The results from the statistical analysis of social factors in MCS showed age and gender being significant factors that predicted SPE, with younger speakers and male speakers producing higher rates of explicit pronouns. While not a statistically significant variable, speakers with higher levels of education also produced more explicit pronouns. The common thread among these three factors is the exposure to a non-contact variety of Spanish (NS), in which explicit pronouns were produced around 24% of the time. Participants with higher

levels of education and belonging to the younger age group were closer to this number, possibly due to their exposure to more prescriptive language in the education system. In addition, male speakers in the community have a tendency to be more mobile and therefore come into contact with speakers outside the Caribbean coast more frequently than women. The trends regarding social factors showed that the more a group of speakers was exposed to either prescriptive language or NS norms, the more likely they were to produce higher rates of explicit subject pronouns, coming closer to SPE patterns in NS.

Chapter 6

Data Analysis: Variable Number Marking

In this chapter I will present the data analysis for variable number marking. The discussion of the results includes an overview of the overall distribution of the two variants in the data set, an explanation of the statistical model used, and a presentation of which factors predicted singular number marking on 3pl. verb forms in MCS. This analysis does not include data from NS as this group of speakers did not display any variation in regards to the dependent variable (6.1 and 6.2).

(6.1) *ellos no pod-ían enamorarse porque*
they no can-PST.IPFV.3PL fall.in.love because
er-an de diferentes razas
be-PST.IPFV.3PL of different races
'They could not fall in love because they were from different races'

(6.2) *ellos ib-an a tener mucha comida en*
they go-PST.IPFV.3PL to have.INF a.lot food in
abundancia
abundance
'They were going to have a lot of food in abundance'

6.1 General Frequency Results

The general frequency results (Figure 6.1) for MCS show that lack of agreement occurred in 23.87% of the data set overall, with 142 occurrences out of 595. All 20 speakers displayed variability in number marking.

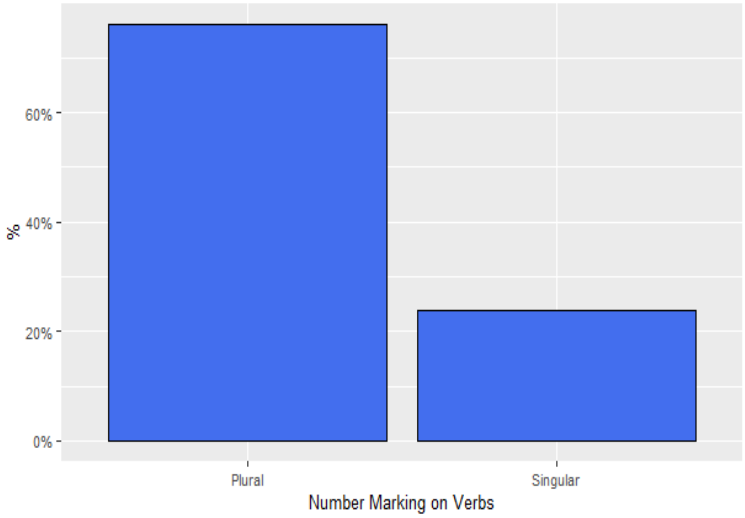

Figure 6.1: Distribution of Number Marking in MCS

The overall distribution of the variable was very close to what was attested by Critchfield and Lívio (*forthcoming*), who found lack of agreement at a rate of 20.3% in their MCS data. The variation found in the current data from MCS speakers was only observed within the context of 3pl. While there were cases of plural marking on 3sg. verb forms, this was limited to the subject *gente* 'people', a non-count noun in Spanish.

6.2 Model Selection

The data set used for this analysis was comprised of a categorical dependent variable: plural number marking or singular number marking, with 595 total tokens. Each token was coded for eight different independent variables: phonic salience, subject position, subject type, animacy of the subject, speaker gender, age, level of education, and age of acquisition.

I started the model selection by running an initial logistic regression model in R using lrm(). This output was used to check the overall significance and goodness-of-fit of the model with all eight independent variables. The results showed that overall the model had predictive power and was significant, based on a small *p*-value of the Model Likelihood Ratio Test (<0.0001). In terms of goodness-of-fit, the concordance index C was 0.757, which signals acceptable discrimination (Hosmer and Lemeshow, 2000, p. 162).

Next, I reviewed the output of the model to see if any of the factors should be removed. I checked for any signs of data sparseness or multicollinearity by examining the standard errors for each object in the model. All of the standard

errors were small, showing no evidence of multicollinearity. I used a generalized linear regression to run three different stepwise tests: forward, backward, and bidirectional. All results showed that all predictors, other than subject type and age of acquisition, significantly improved the model. This was evaluated using the AIC, which goes down as a model becomes more significant. The model's AIC with subject type and age of acquisition included as predictors was 581.3 and without them was 577.91. In order to determine whether the difference in AIC values was significant, I ran an ANOVA with the two models. The results showed that the difference in the AIC was not significant so I decided to keep both subject type and age of acquisition in the model. Finally, I looked for possible interactions between my independent variables and did not find anything statistically significant.[8] My final model was a mixed-effects logistic regression with speaker as a random factor and phonic salience, subject position, subject type, animacy of the subject, speaker gender, age, level of education, and age of acquisition as fixed-effects (Table 6.1). This final model had an AIC of 577.5 and showed phonic salience, subject position, animacy, age and gender being significant predictors (based on low p-values).

[8] $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 6.1: Mixed-Effects Logistic Regression Predicting 3sg. Number Marking with 3pl. Subjects in MCS: Linguistic and Social Factors (n=595, AIC=577.5)

Variable	Estimate/Odds	SE	P r(>|z|)	n	%3sg.
Number marking					
(vs. Plural)				453	76.13%
Singular	-4.5/.011	.71	3.00e-10 ***	142	23.87%
Phonic salience					
(vs. High)				103	6.8%
Low	1.67/5.29	0.43	9.31e-05 ***	492	27.4%
Subject position					
(vs. Immediately preceding)				317	16.1%
Distantly preceding	.75/2.13	.25	.00240 **	189	30.7%
Post posed	.68/1.97	.32	.03266 *	89	37.1%
Subject type					
(vs. Pronoun)				277	19.1%
Noun phrase	.18/1.2	.24	.43998	318	28%
Animacy					
(vs. [+human])				518	22.4%
-human	.77/2.16	.32	.01654 *	77	33.8%
Age					
(vs. 18-29)				332	16.9%
30+	.96/2.6	.39	.01395 *	263	32.7%
Gender					
(vs. Female)				369	17.1%
Male	.98/2.67	.36	.00608 **	226	35%
Level of education					
(vs. University)				343	22.2%
Secondary	.67/1.95	.35	.05653	252	26.2%
Age of acquisition					
(vs. Child)				435	20.5%
Adolescent	.04/1.04	.44	.93001	160	33.1%

Table 6.1 gives information about the number of tokens observed for each level of any given factor ('n'), as well as the percentage of that number that were 3sg. ('% Explicit'). The coefficient estimates in Table 6.1 were used to determine the simple log odds of the outcome. These estimates are shown for the levels listed (in comparison to the reference level). The reference level for the dependent variable was plural number marking, meaning the model compared the probability of the second level (singular) with the reference levels (plural,

high, immediately preceding, pronoun, [+human], 18-29, female, university, child). Exponentiating the coefficient gave the simple log odds of singular verb morphology vs. plural verb morphology (with the reference levels of the predictors): .011, or 1.1% probability. When the reference levels were reset to reflect the categories that favored singular tokens the most (low, preceding, noun phrase, [-human], 30+, male, secondary, adolescent), the odds increased to 4.38, which means that under these conditions, singular number marking was a little over four times more likely than plural number marking.

The p-values for the independent variables show that all predictors were significant other than subject type, level of education, and age of acquisition, which means there was no significant difference between pronouns and noun phrases, secondary and university level of education, and speakers who acquired Spanish as children versus as adolescents with regard to the odds of singular vs. plural number marking. Any *p*-value smaller than 0.05 was considered to be the threshold for statistical significance. The effect size for each predictor, presented in the following sections, can be calculated as well by exponentiating the coefficient estimate. These numbers show how much more likely the level shown was to result in singular number marking (non-agreement), compared to the reference level for that specific predictor.

6.3 Linguistic Factors

The linguistic factors were analyzed using a mixed-effects logistic regression in order to determine which predictors motivated the production of singular verb forms with 3pl. subjects.

6.3.1 Phonic Salience

Phonic salience was divided into two categories: high and low. This was one of the factors with the strongest predictive power, with a *p*-value of 9.31e-05 ***. The model showed that verbs with low phonic salience were 5.29 times more likely to result in non-agreement. In this context, singular number marking occurred 27.4% of the time versus 6.8% of the time with high phonic salience verbs (Figure 6.2).

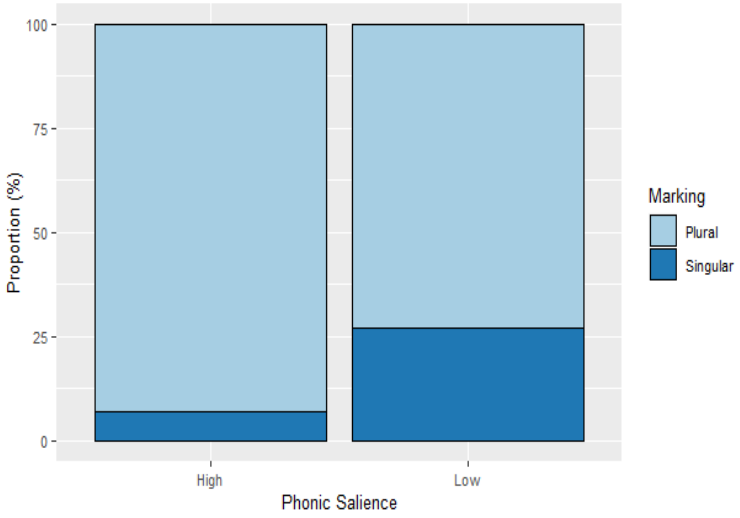

Figure 6.2: Phonic Salience and Number Marking

Low phonic salience tense/moods such as present (*tiene* in 6.3) and imperfect (*permitía* in 6.4) were the strongest predictors of singular number marking in the data set.

(6.3) *los* *miskito-s* *también* *tien-e* *derecho* *a* *su*
the.PL Miskitu-PL also have-PRS.3SG right to their
tierra *natal*
land native
'The Miskitus also have a right to their home land'

(6.4) *antes* *los* *líder-es* *no* *permit-ía* *eso*
before the.PL leader-PL no permit-PST.IPFV.3SG that
'Before the leaders didn't permit that'

High phonic salience verb forms did not frequently motivate the production of non agreement, with only seven cases being observed in the data (*vino* in 6.5 and *fue* in 6.5).

(6.5) *eso-s* *buzo-s* *como* *vin-o* *aquí,* *hac-en*
that-PL diver-PL like come-PST.3SG here make-PRS.3PL
huelgas
strikes
'Those divers, like they came here, they make strikes'

(6.6) *ellos* *fue* *guerrilleros*
they be-PST.3SG warriors
'They were warriors'

These results align with what was previously found in MCS (Critchfield and Lívio, *forthcoming*), as well as with what has been attested for Brazilian Portuguese (Naro and Scherre, 2000; Mendes and Oushiro, 2015). Non-agreement is more likely to occur when the difference between singular and plural verb morphology is less perceptible, which is directly related to tense/mood (Naro and Scherre, 2000, p. 243). Present and imperfect forms are therefore less salient with the difference only being a final nasal (see sections 3.2.1 and 4.2.6 for a review of phonic salience in BP and MCS, respectively).

6.3.2 Subject Position

Subject position combined both syntactic order of the subject and the verb, as well as the presence of any linguistic material between the two. The three levels included in the analysis were immediately preceding, distantly preceding (one or more syllables between the subject and verb), and post posed. The results from the logistic regression showed that the probability of getting singular number marking was 2.13 times more likely when the subject was distantly preceding and 1.97 times more likely when it was post posed, compared to when it was immediately preceding the verb. When the subject was immediately preceding the verb, 3sg. tokens were produced 16.1% of the time, versus 30.7% when distantly preceding and 37.1% when post posed. Both distantly preceding and post posed were statistically significant in predicting singular number marking, with *p*-values of .00240** and .03266* respectively.

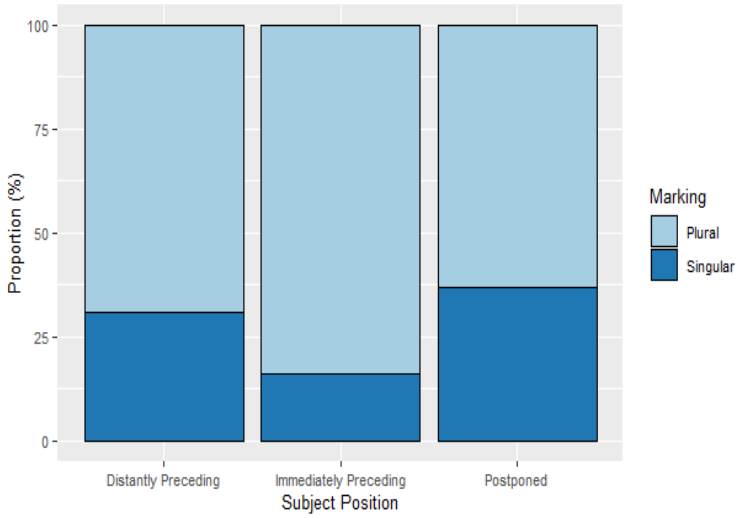

Figure 6.3: Subject Position and Number Marking

These results also align with previous findings for MCS and Brazilian Portuguese, where higher rates of non-agreement occur when the subject is post posed and when there is linguistic material between the subject and the verb (Critchfield and Lívio, *forthcoming*; Mendes and Oushiro, 2015, p. 363).

6.3.3 Subject Type

Subject type was divided into two categories: pronoun and noun phrase. This was the only linguistic factor that did not significantly favor or disfavor singular number marking; however, with noun phrases, 3sg. tokens occurred 28% of the time (versus 19.1% with pronouns) and were 1.2 times more likely to result in non-agreement (Figure 6.3.

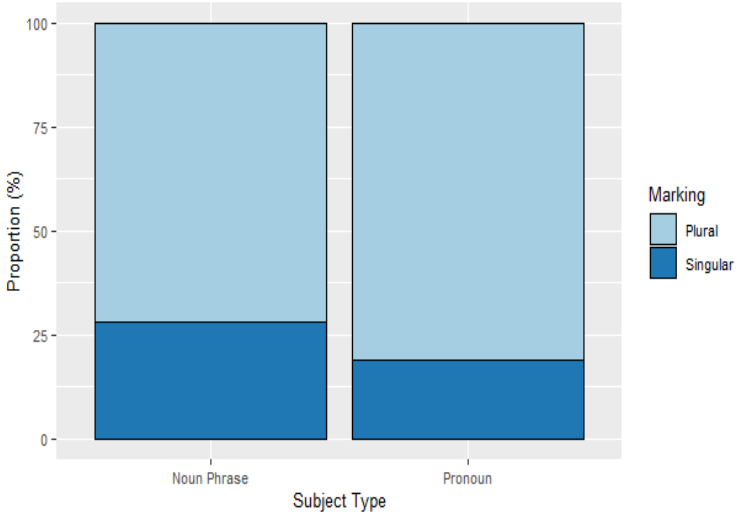

Figure 6.4: Subject Type and Number Marking

We can draw on Gundel et al.'s (1993) Givenness Hierarchy to explain the results for subject type. They propose the notion that the cognitive status of the referent conditions the subject form used by the speaker, which is based on how prominent the referent is in the discourse. According to the Givenness Hierarchy, the more activated the referent is in the mind of the speaker, the more minimal the subject form used (Gundel et al., 1993). According to Zanini, Chiara, et al., 2020 and their work on the processing of morphological number, the more salient the referent, the easier it is to assign number; therefore, the fact that more cases of non-agreement occurred with pronouns fits with this reasoning, as they are more minimal in form compared to noun phrases.

6.3.4 Animacy

Animacy of the subject had two levels: [+human] and [-human]. The results showed that when compared to [+human] tokens, [-human] subjects (6.7 and 6.8) were 2.16 times more likely to predict singular number marking, therefore being a statistically significant predictor with a *p*-value of .0165*. When a subject was [-human], non-agreement occurred 33.8% of the time versus 22.4% of the time when the subject was [+human] (Figure 6.5).

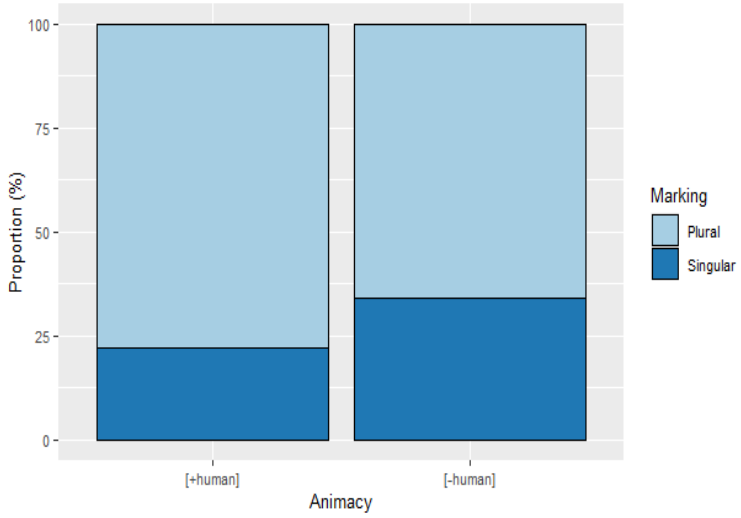

Figure 6.5: Animacy and Number Marking

(6.7) *antes* *ven-ía* *uno-s* *proyecto-s*
before come-PST.IPFV.3SG some-PL project-PL
'Some projects used to come here'

(6.8) *eso-s* *canal-es* *no* *es* *bueno* *para* *nosotros*
that-PL canal-PL no be-PST.3SG good for us
'Those canals are not good for us'

This factor was not found to be significant in MCS by Critchfield and Lívio (*forthcoming*); however, the results do reflect what has been found for Brazilian Portuguese, where [-human], [-animate] subjects favor singular number marking (Mendes and Oushiro, 2015, p. 363). This tendency can be explained by the connection between saliency and animacy. The more animate a referent is, the more salient it is in the mind of the speaker. This saliency results in a greater ease of assigning number to any given referent (Zanini, Chiara, et al., 2020).

6.4 Social Factors

In this section, I will discuss the results for the four social factors included in the analysis: age, gender, level of education, and age of acquisition. The results for this section are of particular importance due to the lack of sociolinguistic studies on Nicaraguan varieties of Spanish.

6.4.1 Age

The age of participants was broken down into two groups: 18-29 and 30+. The logistic regression showed that speakers over the age of 30 were 2.6 times more likely to produce non-agreement, a statistically significant predictor with a *p*-value of .01395*. The younger group of participants produced singular number marking 16.9% of the time, while the older group of participants produced it 32.7% of the time (Figure 6.6).

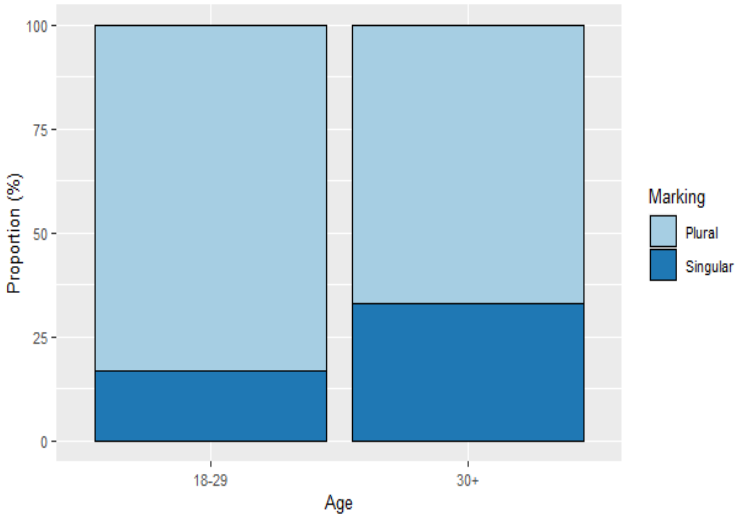

Figure 6.6: Age and Number Marking

General sociolinguistic findings for age show that younger speakers tend to use more non-standard variants, unless highly stigmatized (Silva-Corvalán, 2001, p. 102; Chappell, 2015a, p. 225). The results of the current analysis therefore suggest that this feature may be stigmatized in the community or within the country as a whole. They also may suggest the direction of change for variable number marking in MCS. Since younger speakers are producing less non-agreement, this feature may decrease in the speech community over time.

6.4.2 Gender

Gender was one of the factors with the greatest statistical significance, displaying a *p*-value of .00608**. The model showed that male speakers were 2.67 times more likely to produce singular verb marking with 3pl. subjects when compared to female speakers. Males produced non-agreement 35% of the time while females only produced non-agreement 17.1% of the time (Figure 6.7).

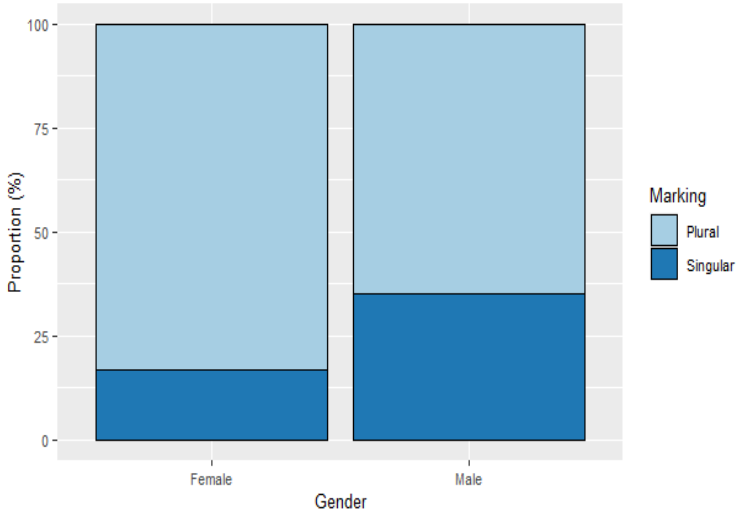

Figure 6.7: Gender and Number Marking

The results for this data set mirror what has been attested in Brazilian Portuguese. In addition, findings across languages show that female speakers use standard, more prestigious variants when compared to male speakers, often being evaluated as speaking in a more "correct" manner (Labov, 2001; Silva-Corvalán, 2001, p. 98). This is particularly true in formal speech, where females make more self-corrections than males, a sign that they are more aware of speech choices and their social consequences. Silva-Corvalán (2001, p. 98) explains that a double standard exists in regards to gender, with women being expected to act in a polite and courteous manner, while it is acceptable for men to break rules and behave in a rude or vulgar manner. In addition, traditionally women have had fewer opportunities for advancement, therefore socioeconomic status was communicated through their physical appearance and speech.

6.4.3 Level of Education

Level of education was broken down into two categories: secondary and university. The results of the analysis showed that this factor did not quite meet the <

.05 threshold for statistical significance in predicting singular number marking (p-value = .05653); however, speakers with a secondary level of education were still almost 2 times more likely to produce lack of agreement when compared to speakers with a university level of education. The participant groups produced singular number marking at a similar rate: 22.2% for those with a university-level education and 26.2% for those with a secondary-level education (Figure 6.8).

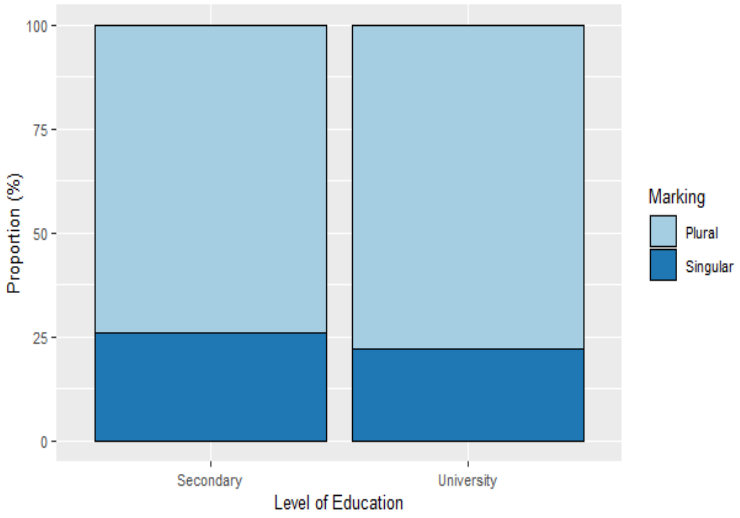

Figure 6.8: Level of Education and Number Marking

Level of education is one of the strongest predictors of this variation in Brazilian Portuguese, as speakers with less formal education produce higher rates of singular marking with 3pl. subjects (Brandão and Vieira, 2012; Naro and Schere, 2013; Oushiro, 2015). These findings for both MCS and Brazilian Portuguese show that there is a correlation between level of education and the use of standard language forms.

6.4.4 Age of Acquisition

Age of acquisition was divided into two groups: child and adolescent, in order to account for those speakers who acquired Spanish at or before the age of 6 (generally upon entering primary school) and those that acquired Spanish at or after the age of 12 (generally upon entering secondary school). The model did not show this factor to be significant in predicting singular number marking; however, speakers who acquired Spanish at an older age produced non-agreement 33.1% of the time, while speakers who acquired Spanish at a younger age produced non-agreement 20.5% of the time (Figure 6.9).

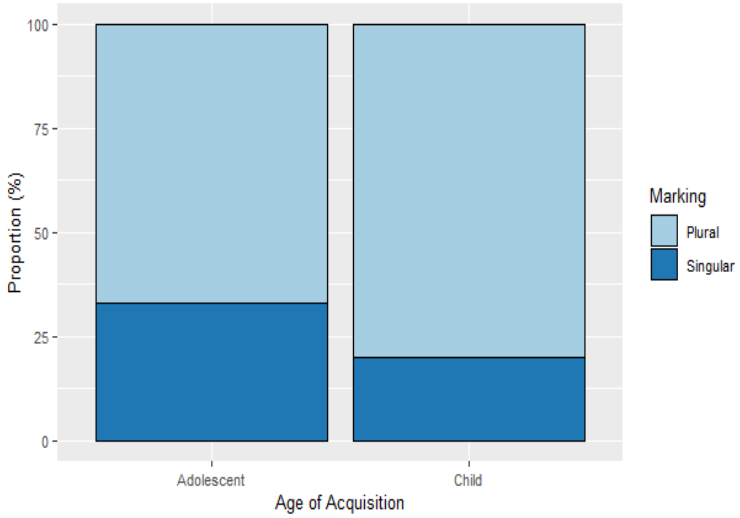

Figure 6.9: Age of Acquisition and Number Marking

These results may indicate that this is a fixed feature in the MCS dialect established when Spanish was first implemented in the education system on the Caribbean coast, and therefore not based on current speakers' acquisition process.

6.5 Summary of Results

The data analysis of variable number marking in MCS showed that a mix of linguistic and social factors predicted the production of singular verb morphology with 3pl. subjects (non-agreement): phonic salience, subject position, animacy, age, and gender. The mixed-effects regression model confirmed that the chances of getting singular number marking with 3pl. subjects is 4.38 times higher when the phonic salience of the verb is low, the subject distantly precedes the verb, the subject is a noun phrase, the subject is [-human], the speaker is over the age of 30, male, with a secondary education, and who acquired Spanish as an adolescent. Initial plotting of the dependent variable therefore correlates with the trends discussed for MCS previously and with Brazilian Portuguese.

Chapter 7

Discussion

The final piece of the analysis will include a discussion of how language contact, acquisition, and language shift explain the results from the statistical analysis. I will explain how speakers of MCS compensated for differences in the L1 and L2 regarding SPE and number marking by utilizing both language universal cognitive strategies and L1 transfer. The research questions will be used to guide the discussion.

7.1 Linguistic Factors

1. How do the two participant groups compare in regard to the linguistic factors that predict explicit subject pronoun expression and singular number marking on 3pl. verbs?

A variety of linguistic factors were included in the two analyses. The NS speaker data was only examined for SPE as this group did not demonstrate any variation in number marking. Regarding the production of implicit versus explicit pronouns, NS participants aligned with what has been found in previous research for Spanish (Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008; Abreu, 2012; Otheguy and Zentella, 2012; Michnowicz, 2015): singular referents, less distinctive TMA, contexts of switch reference, and non-reflexive verbs increased the production of explicit subject pronouns. The NS group, therefore, acted as a point of comparison for the MCS participants.

MCS data was analyzed for both SPE and variable number marking. Regarding SPE, previous research predicts that contact varieties of Spanish will produce higher rates of explicit pronouns as a result of weakened pragmatic constraints (Silva-Corvalán, 1994; Otheguy, Zentella, and Livert, 2007; Michnowicz, 2015). The MCS data showed the opposite, with overall lower rates of explicit pronoun production; however, the participants did show some evidence for pragmatic

weakening as they appeared to be sensitive to discourse-pragmatic constraints but not to the same degree as the monolingual group. While less distinctive TMA and switch reference contexts resulted in higher rates of explicit pronoun production, these rates were much smaller than what was observed in the NS data. According to Silva-Corvalán (2001, p. 154), on average speakers produce an explicit subject 50% of the time when there is a change in referent. The NS speakers came close to this number, using overt pronouns in 40.8% of contexts with switch reference, while the MCS speakers only produced overt pronouns in 24% of these cases. The MCS results for number diverged completely from what was found for NS and other varieties of Spanish (Otheguy et al., 2007; Michnowicz, 2015): plural referents motivated higher rates of explicit subject pronouns instead of the opposite effect. The last variable, reflexivity, was not a significant motivator.

Since Brazilian Portuguese is the only related language or variety in which variable number marking has been analyzed from a variationist perspective, it was used as a point of comparison in the MCS analysis. Research shows that several linguistic factors motivate the production of singular verb forms with 3pl. subjects in Brazilian Portuguese, including phonic salience of the verb, subject position, subject type, and animacy of the referent. The MCS speakers in the present data set aligned with these trends for all four factors. Low phonic salience, distanced or inverted subject and verb, subjects that are noun phrases, and non-human referents all favored non-agreement.

7.2 Social Factors

2. What social factors predict patterns in subject pronoun expression and variable number marking in MCS?

The analysis of SPE and variable number marking included social predictors for MCS. Clear patterns emerged in both analyses. The findings for social factors showed that the less a speaker was exposed to prescriptive norms, the more likely they were to diverge from patterns found in NS, the majority dialect in Nicaragua.

In the analysis of SPE, older speakers, females, and speakers with lower levels of formal education differed the most from NS, producing much lower rates of explicit subject pronouns. In contrast, younger speakers, males, and speakers with higher levels of formal education behaved more similarly to the control group of NS participants.

In the analysis of variable number marking, the non-standard variant (singular number marking) was produced at higher rates by older speakers, males,

speakers with lower levels of education, and those who acquired Spanish at a later age.

Age was a statistically significant predictor for both SPE and variable number marking, with younger speakers more closely resembling NS. This finding suggests that both features analyzed have some level of stigmatization attached to them as younger speakers tend to use more non-standard forms unless a feature is stigmatized (Silva-Corvalán, 2001, p. 102). Another explanation is that due to increased exposure to standard language norms via education and media, younger speakers approximate more to NS. Regardless of the reasoning, this finding could suggest the direction of change for both SPE and variable number marking in MCS. These features may change over time to reflect tendencies in NS as contact between the two coasts increases.

Level of education was not a statistically significant variable in either of the analyses but it did show a favoring effect. Speakers with a secondary level of education produced more non-standard language, contributing to more cases of low explicit SPE and more cases of singular number marking with 3pl. verbs. In contrast, speakers with a university level of education approximated more to language norms in NS, producing explicit subject pronouns at a more similar rate to NS participants and producing fewer examples of subject-verb non-agreement. General findings in sociolinguistics reflect this tendency: speakers with lower levels of formal education deviate more from standard language norms (Brandão and Vieira, 2012; Naro and Schere, 2013; Chappell, 2015a; Oushiro, 2015). Level of bilingualism has therefore been connected to use of non-standard language. While level of education is different than level of bilingualism, speakers with higher levels of formal education have more exposure to prescriptive language norms in class. This trend is apparent in the findings for age of acquisition in the analysis of variable number marking where speakers who acquired Spanish at a later age produced more non-agreement.

The only social factor that had different patterns for the two features was gender, with females deviating more from standard norms in regard to SPE and males deviating more from standard norms in regard to variable number marking. The relationship between gender and use of non-standard language has several explanations in the linguistics literature. Women for the most part have the tendency to use more standard forms. This has been attributed to their awareness of the connection between speech choices and social consequences, which is directly related to the fact that historically women have had fewer opportunities for advancement. They therefore have relied more heavily on the use of standard language as a way to communicate their socioeconomic status (Silva-Corvalán, 2001, p. 98). Other cases where women use more standard

forms have been tied to their lack of mobility and not to socioeconomic status. Rodrigues (1987) found women in rural areas were more likely to produce standard agreement between subjects and verbs than men were. This was attributed to men leaving the community to work and subsequently being exposed to working-class populations in which variable number marking is more frequent.

Other findings have shown women to use more non-standard language and therefore be at the forefront of linguistic change, particularly in contexts of language contact. In her research Shin (2013) found female Spanish speakers in the United States used higher rates of explicit subject pronouns due to contact with English (a non-pro drop language). In this case, the tendency to adhere to language norms in the L2 resulted in the use of more non-standard language patterns in the L1. These examples show the tendency for females speakers to adhere to language norms, whether they be in the L1 or L2, or whether those be standard or non-standard forms depending on the speech community.

Based on the findings for the two features analyzed here, it seems that variable number marking might be more highly stigmatized, which explains why the female speakers produced non-agreement at a lower rate. However, in the case of SPE, women differed more than men when compared to NS speakers, using fewer explicit subjects pronouns. Since SPE is already variable and depends on several discourse-pragmatic factors, the underproduction or overproduction of explicit pronouns would be less stigmatized. Men, who as a whole have more contact with NS speaker norms, via mobility and employment, approximate more to the standard rate of explicit subject pronouns.

7.3 Language Contact

3. Do the results for SPE and number marking in MCS show evidence for incomplete acquisition due to language contact? If so, what learner strategies best account for the results for both features?

The Caribbean coast of Nicaragua, where MCS is spoken, provides an interesting case study for how language contact can affect the acquisition of morphosyntactic features in Spanish. MCS developed in the mid twentieth century when the Miskitu community, both adults and children, acquired Spanish as a second language due to social and political pressure (García, 1996; Escobar, 2013; Wilson Withe et al., 2012). This acquisition process has played an important role in the development of MCS, as the speakers were geographically isolated from the target language, NS. This isolation has led to what I believe to be the incomplete or divergent representation of certain features in MCS. According to Montrul (2006, 2008), incompleteness is a result of insufficient target lan-

guage input; therefore it is important to identify what the target variety was for the speakers in this study in order to determine whether or not grammatical properties are in fact incomplete. When the shift to Spanish began, Miskitu people were being taught by NS speakers via the Río Coco Pilot Project for Basic Education and the Sandinista literacy campaign. At the time, there was no direct ground access to the Pacific coast capital, and telegraph was the only means of communication (*The Miskitos in Nicaragua, 1981-1984*. 1984). This isolation would have meant limited contact with NS speaking populations, outside of the educational sphere. The lack of input these first speakers of MCS experienced would have greatly impacted the acquisition of Spanish in a target-like manner. In addition, the majority of these learners would have already fully formed their first language, therefore being more susceptible to incomplete acquisition.

The current target language for learners is not necessarily NS. Most speakers now acquire Spanish from other speakers of MCS: family, friends, or teachers. Based on this assumption, it would be expected that learners would acquire features of MCS. Assuming the infrequent use of explicit subject pronouns and the variable number marking on 3pl. verbs are fixed features in MCS, it is logical that these would be aspects of the grammar that would continue to be passed down; therefore, their use would not represent incompleteness, when compared to the target variety.

We can conclude that incomplete acquisition of SPE and number marking occurred in the speech of MCS speakers during initial language shift in the past but that the distribution of these features in current MCS speakers is not due to incomplete acquisition, but is simply characteristic of the variety. The relationship between the acquisition of grammatical properties, such as SPE and number marking, and input in this contact situation provides insight into the acquisition of MCS. Montrul (2006, 2008) asserts incomplete acquisition is more related to input and not an effect of the age the second language was acquired. This stance goes against what was proposed by Schachter's (1990) *Incompleteness Hypothesis*, as well as Sorace's (2013) claim, that incomplete acquisition occurs when a language is acquired outside the critical period. The results of this dissertation support Montrul's claims, as age of acquisition was not a significant predictor of either feature under investigation. I assume, because of the contact situation surrounding MCS when it first developed, that the features I observe in MCS related to SPE and number marking are remnants of the incomplete acquisition of certain grammatical properties that took place during initial language shift. These were then passed to new generations via input and have become fixed features in the language and not a result of current incomplete acquisition.

While low explicit subject pronoun use and variable number marking on 3pl. verbs are current features of MCS, their distribution is changing. Younger speakers appear to be approximating closer to distributions found in NS and producing higher rates of explicit pronouns and more subject-verb agreement. Why is this the case? Based on the results for age, the features analyzed in this dissertation are stigmatized, aligning with findings by Silva-Corvalán (2001, p. 102) and Chappell (2015a, p. 225). This, coupled with more exposure to standard language norms via education, media, and travel, the younger MCS speakers in the study behaved more similarly to NS speakers.

The findings for subject pronoun expression and number marking also appear to be strongly connected to bilingualism and language contact. For the analysis of SPE, the NS speakers produced rates of overt pronouns that fell in the middle of the spectrum of other monolingual varieties (23.9%). MCS speakers, however, diverged from what was found for NS and for other contact varieties of Spanish. Instead of using more explicit pronouns, MCS participants produced fewer (16.5%). The analysis of number marking also resulted in differences between the two participant groups as this phenomenon was only observed in MCS and not NS.

Studies of SPE in contact varieties show that speakers are likely to produce different rates of explicit pronouns compared to non-contact varieties, which is supported by the current findings for MCS. Previous research on variable number marking also bolsters the claim for contact-induced effects as there are no attested varieties of Spanish or Portuguese that display this type of non-agreement that do not have significant contact with other languages as part of their past or present. That being said, similar phenomena have been found in contexts with no recent contact with other languages (such as Caribbean Spanish) but final /s/ and /n/ deletion are not the same as the variable number marking presented here. The only other varieties or languages with parallel features are Afro-Bolivian Spanish and Brazilian Portuguese, both of which have a history of language contact and acquisition.

MCS displays evidence of past incomplete acquisition of SPE and number marking as speakers behave differently than NS speakers in regard to explicit subject pronoun distribution and number marking on 3pl. verb forms. Based on the patterning of these features in younger speakers, it appears that MCS is moving toward a more standard distribution, similar to that found in NS, which is likely due to increased exposure to NS in the present day. Despite this change, we can still expect to find evidence in the data analysis of learner strategies that were used in the face of insufficient target language input during initial language shift. Is there a clear learner strategy that explains the results for

both features? The results of the two statistical analyses provide evidence for both universal acquisition processes related to cognition and L1 influence.

7.3.1 Salience

The term salience describes the way in which a stimulus stands out in our minds: "Salient items or features are attended, are more likely to be perceived, and are more likely to enter into subsequent cognitive processing and learning." (Ellis, 2017, p. 21). Learners draw on salience as a cognitive strategy when given a set of linguistic choices, such as implicit versus explicit subject pronoun. One example of a tendency found in second language acquisition, regardless of L1, involves subject expression and processing difficulties at the syntax-discourse interface. Previous studies on contact varieties of Spanish find that while still sensitive to referent status; speakers of Spanish as a second language tend to overproduce explicit pronouns in contexts of same reference and under produce implicit pronouns in contexts of switch reference (Cameron, 1992; Silva-Corvalán, 1994; Otheguy, Zentella, and Livert, 2007; Orozco and Guy, 2008, Michnowicz, 2015). These difficulties have not been attributed to L1 transfer but instead to the complex discourse-pragmatic constraints that exist in Spanish with regard to SPE. The ability of learners to operate within these constraints, albeit with a less standard distribution, demonstrates reliance on cognitive processes, namely salience. This notion has been used in the acquisition literature to describe the activation of the referent in the mind of the speaker (Gundel et al. 1993; Lubbers Quesada, 2015). More minimal forms, which in the context of this analysis would be implicit pronouns, indicate greater referent salience; therefore explicit pronouns would indicate lesser salience. A switch moves the focus from the most recent, salient referent, therefore requiring an explicit form. MCS displayed a weakening of pragmatic constraints in regard to coreference. While speakers produced explicit pronouns at a much lower rate than NS speakers (24% vs. 40.8%), they favored overt forms in contexts of switch reference (p=.003636 **). These results show that despite complex constraints on SPE in Spanish, MCS speakers still display sensitivity to a change in referent, drawing on salience of the referent to determine subject pronoun form.

Another example of salience as a strategy used in language choice can be found in the distribution of the linguistic factors in variable number marking. In this context, the term 'salience' is used to describe how the verb form stands out from others and also how activated the referent is in the mind of the speaker. The more salient the referent, the more likely the speaker was to produce agreement. In terms of form, more minimal forms, such as pronouns, reflect greater salience of the referent in the mind of the speaker (Levinson, 2000; Lubbers

Quesada, 2015, p. 265). Human referents are also more salient for speakers, compared to nonhuman referents (Zanini, Chiara, et al., 2020). Finally, high phonic salience verbs should favor agreement more due to the greater perceptible difference between 3sg. and 3pl. forms. These favoring effects were found in the MCS data, with pronouns, human referents, and high phonic salience verbs more likely to result in agreement. This reflects the way in which speakers utilize salience as a cognitive strategy when producing either singular or plural verb morphology in the context of 3pl.

Simplification

Simplification, another strategy used by learners, involves the reduction of target language structures (Winford, 2003, p. 217). It has also been used to explain the tendency to overproduce explicit subject pronouns in contact varieties of Spanish, specifically with non-Romance languages. These findings are due to typological differences that result in explicit pronouns being used as the default, as both overt and null forms are permissible in Spanish. This is a strategy used by speakers in the face of the complex discourse-pragmatic constraints that determine pronoun expression in Spanish. This pattern, however, does not align with the results for the current study as MCS speakers produced lower rates of explicit subject pronouns compared to NS speakers, not higher rates.

It has been suggested that simplification can result in an increase in SPE or weakening of pragmatic constraints (Sorace, 2004, 2005; Prada Pérez, 2009; Michnowicz, 2015). We see the second of these predictions in the decreased sensitivity to contexts of switch reference on the part of MCS speakers; however, overall explicit subject pronoun rates in MCS do not align with this claim. I propose that the simplification process used by learners of Spanish can not only result in an overproduction of subject pronouns but also an underproduction. MCS participants speak Miskitu as a first language and Spanish as a second language. As previously outlined, Miskitu only allows null subjects with singular referents while Spanish allows them for all referents. It is possible that MCS speakers use fewer subject pronouns because null subjects are possible for all grammatical persons/number. In the case of MCS, speakers seem to avoid the complex discourse-pragmatic constraints involved in Spanish SPE by using fewer explicit forms all together.

While MCS speakers draw upon universal cognitive strategies to compensate for incomplete language acquisition, influence from Miskitu is also clear in both SPE and variable number marking.

L1 Transfer

L1 transfer is one of the strategies utilized by learners who have not acquired a second language in its entirety and involves any type of cross-linguistic influence (Winford, 2003, 2005). Speakers tend to employ this strategy when language structures or grammatical properties are transferred from the first language to the second language and is more likely when the L1 and the L2 differ, as is the case with Miskitu and Spanish (in terms of their use of subject pronouns and in their verb morphology).

The MCS data provides evidence of L1 transfer in regard to how number impacts SPE and verb morphology. Spanish is a null-subject language that permits both explicit and implicit subjects, primarily constrained by discourse-pragmatic factors. Miskitu is a partial null-subject language that relies on a distinct distribution of null and overt subjects, with singular referents permitting both explicit and implicit subjects, and plural referents requiring the use of explicit subjects exclusively. Evidence of L1 transfer is apparent in the results for SPE in MCS, specifically in the distribution as it relates to number. The logistic regression shows plural verbs were stronger predictors of explicit SPE. Speakers produced an overt pronoun 24.7% of the time when the referent was plural (vs. 14.5% when the referent was singular). These findings do not correlate with other research on contact varieties of Spanish and SPE (Otheguy, Zentella, and Livert, 2007; Michnowicz, 2015), where singular referents tend to favor more explicit subject pronoun use. The pattern found in the current MCS data, therefore provides evidence of transfer from Miskitu.

The linguistic factor of number in the analysis of variable number marking also provides evidence of L1 transfer. Number marking in Miskitu is different from Spanish as number is not marked in the morphology of the verb. The same form is used for singular and plural within each grammatical person, resulting in three total verb forms for any given TMA of a verb: one for first person, one for second person, and one for third person. This typological difference has resulted in the presence of variable number marking in MCS due to partial transfer from Miskitu. While speakers appear to have acquired the singular/plural distinction for first and second person, the context of third person is more susceptible to L1 transfer because the perceptibility in forms is minimal due to low phonic salience for many TMAs. It is common for learners to employ the strategy of reduction or elimination of target language inflections such as morphemes used to mark gender, number, or agreement (Winford, 2003, p. 213). The use of 3sg. verb forms with 3pl. referents can therefore be attributed to the differences in number marking that exist in Miskitu and Spanish.

7.4 Summary

Guy (2014) discusses three general categories of explanations that can be used to determine why certain features exist in some varieties of the same language but not in others: ancestry, innovation, and contact. Here I focused on the contribution of contact on resulting SPE and number marking patterns in MCS. This is particularly relevant in the context of Nicaragua's Caribbean coast as the Miskitu community has undergone acquisition of a second language and group language shift. When this initially occurred beginning in the 1950s with the Río Coco Pilot Project for Basic Education and later into the late 1980s with the expansion of the Sandinista literacy campaign, speakers would have developed an interlanguage, or learner variety of Spanish. In order to compensate for the many structural differences between Miskitu and Spanish, it is likely that the speakers in this community would have used a variety of strategies to communicate in their L2, including L1 transfer and simplification (Winford, 2003, pp. 209-210). These strategies are considered part of incomplete language acquisition, which according to Winford (2003, p. 217), is present in situations of contact and natural second language acquisition, which is the case for MCS. The findings for MCS also provide evidence for cognitive strategies expected during acquisition, regardless of L1, such as reliance on salience to determine language choice.

The development of a second language via natural SLA and group language shift, as in the case of MCS, can result in the incomplete acquisition of some features of the target language (Winford, 2003, p. 222). An important aspect of the acquisition process on the Caribbean coast was the lack of input from the target language. Apart from school, learners did not have access to many speakers of NS (due to geographic boundaries). As a result, certain elements of the interlanguage fossilized, or ceased to be acquired in a target-like manner. This has been seen in other contact-varieties of Spanish, such as Afro-Bolivian Spanish. Sessarego (2011, 2012) suggests that the lack of subject-verb agreement found in Afro-Bolivian Spanish is due to incomplete second language acquisition (Sessarego, 2011, p. 126). When the population first acquired Spanish, they were geographically isolated from the general population of speakers so their acquisition process fossilized at a certain point and remained that way. The speakers of Afro-Bolivian Spanish closely resemble MCS speakers in regard to geographic isolation, providing additional support for the effect of insufficient input and incomplete language acquisition during initial language shift.

In regard to MCS, the second language acquisition process seems to have stopped due to a lack of target language input when the community first began

to shift to Spanish. This explains why the features studied here are still present in MCS. However, more recent contact with NS speakers due to education, media, and migration has resulted in the reduced effect of number from Miskitu, particularly in younger generations.

The divergent results for MCS (compared to NS) regarding SPE and variable number marking provide evidence for incomplete acquisition of thees features. While the MCS speakers display advanced second language acquisition, certain learner strategies related to interlanguage development are present, particularly in regard to number, as this is the common component between SPE and morphological system in verbs that differs between Spanish and Miskitu. By examining the patterns found for the individual linguistic and social factors, the L2 strategies associated with this process: L1 transfer, simplification, as well as salience, explain the present findings for the linguistic features in MCS analyzed in this dissertation.

Chapter 8

Conclusion

This dissertation analyzed two linguistic features found in MCS (NS was used as a comparison group): subject pronoun expression and variable number marking, both of which can be attributed to incomplete language acquisition stemming from the initial language shift to Spanish that occurred in the Miskitu community during the second half of the twentieth century. The linguistic factors analyzed for each of the features in MCS provide evidence of universal cognitive strategies, as well as L1 transfer from Miskitu, utilized by speakers in the past who acquired Spanish with insufficient target language input. The social factors analyzed support the claim for the incomplete acquisition of SPE and number marking due to contact, both of which today appear to have a stigmatized distribution in MCS; however, the results suggest that speakers are moving toward a more standard distribution of these features, as younger speakers with increased exposure to prescriptive language norms approximate closer to patterns found in NS.

NS and MCS behaved differently in their production of explicit and implicit subject pronouns but in both data sets, similar linguistic factors motivated the use of overt forms. For NS, singular referents and contexts of switch reference were the two statistically significant linguistic factors that predicted speakers would produce an explicit subject pronoun. While a change in referent was also a significant predictor in MCS, plural referents had a higher probability of resulting in explicit subject pronouns when compared to singular referents. This, along with the low rate of explicit pronoun production (16.5%) deviate from what has been found for other contact varieties of Spanish. MCS was the only participant group analyzed for social factors, with age and gender emerging as statistically significant factors.

Linguistic and social factors both motivated the use of 3sg. verb morphology with 3pl. subjects in the MCS data set, including phonic salience, subject

position, animacy, age, and gender, These variables were all significant predictors in the statistical model and when the phonic salience of the verb was low, the subject distantly preceded the verb, was a noun phrase, was [-human], the speaker was over the age of 30, a male, had a secondary level of education, and acquired Spanish as an adolescent, the probability of getting non-agreement in the data set was 4.38 times higher. These results coincided with findings discussed for MCS and BP.

8.1 Linguistic Factors

The results for the linguistic factors in both parts of the data analysis provide evidence for the compensation of incomplete acquisition of Spanish via universal cognitive strategies and via L1 transfer.

For subject pronoun expression, learner strategies that compensate for processing difficulties at the syntax-discourse interface can be attributed to cognition and not the L1. Past research confirms that learners of Spanish show sensitivity to changes in referents; however, when in contact with a non-Romance language, learners overproduce explicit forms in contexts of same reference and null forms in contexts of switch reference. This tendency is due to the complex discourse-pragmatic constraints that exist in Spanish. While learners display non-standard distribution of explicit and implicit pronouns, the fact that they continue to be sensitive to some degree to changes in reference is evidence that the cognitive process of salience is being utilized. Findings for coreference in MCS reflect this reliance on salience, as participants significantly favored explicit pronouns in contexts of switch reference, though at a much lower rate compared to NS speakers.

For variable number marking, universal cognitive processes are evident in the effect salience has on the distribution of the linguistic factors. The more semantically or structurally salient the referent, the more likely the speaker was to produce standard 3pl. number marking. Structurally minimal forms, like pronouns, tend to be used by speakers when the referent is more salient in their mind (Levinson, 2000; Lubbers Quesada, 2015, p. 265). In addition, human referents are also more salient (Zanini et al., 2020). Along with high phonic salience verbs, pronouns and [+human] subjects favored agreement in the MCS data. Higher production of standard agreement in more salient contexts is evidence of cognitive patterns not related to the L1.

While some of the findings for SPE and variable number marking in MCS can be attributed to learner strategies related to cognition, transfer from Miskitu is also evident, specifically in how number affected the distribution of linguistic

factors. Speakers have a tendency to rely on their L1 as a learning strategy, specifically when acquiring aspects of language that differ in the L1 and L2 (Winford, 2003). Number as it relates to SPE and number marking on verb morphology have different distributions in Miskitu and Spanish and are therefore more susceptible to L1 influence.

Explicit pronouns are obligatory in Miskitu for plural referents and optional for singular referents, while both overt and null pronouns are possible for all referents in Spanish, with SPE being based on discourse-pragmatic constraints. In the data set, MCS speakers produced more explicit pronouns with plural referents, while NS speakers produced more with singular referents. MCS therefore, displayed evidence of L1 influence and deviated from what has been found previously for both contact and non-contact varieties in regard to this distribution.

Unlike in Spanish, in Miskitu number is not morphologically marked in the verb. The use of 3sg. verb forms with both 3sg. and 3pl. referents can therefore be attributed to the lack of number distinction in Miskitu verb morphology. This reduction of target language inflectional morphemes is common when typological difference exists in the L1 and L2 (Winford, 2003, p. 213). The fact that this variation only occurs in the context of third person, however, is related to the lack of salience between singular and plural forms.

Finally, the overall distribution of subject pronouns in MCS shows evidence of simplification, a frequent finding in contact varieties. While this simplification of discourse-pragmatic constraints usually results in an overproduction of explicit pronouns, the MCS data shows an opposite effect in which speakers demonstrate an underproduction of overt forms.

8.2 Social Factors

The results of the statistical analysis of social factors show the less a speaker was exposed to standard language norms, the more likely they were to diverge from the linguistic patterns in NS.

Age was a statistically significant motivator in both analyses. Speakers aged 18-29 produced more explicit pronouns and higher rates of subject-verb agreement. These results show younger speakers approximate closer to NS norms by using more standard language, a sign that the non-standard distribution for each feature is stigmatized (Silva-Corvalán, 2001, p. 102). Younger speakers tend to be more exposed to standard language norms through formal education and the media. These findings suggest that over time, MCS will become more similar to NS in regard to SPE and number marking.

Findings for gender were mixed as women produced fewer explicit subject pronouns and men produced more subject-verb non-agreement. While results for age suggest both features are stigmatized, the results for gender provide evidence for variable number marking being the more stigmatized of the two, as women are more likely to use standard language norms as a way to achieve upward mobility (Silva-Corvalán, 2001, p. 98). In contrast, the distribution of gender for SPE can be attributed to the increased mobility and exposure to NS men have due to employment outside the community.

Level of education was not a statistically significant factor in either analysis; however, speakers with a secondary level of education produced fewer explicit pronouns and more non-agreement. These findings show speakers with lower levels of formal education deviate more from language norms in NS, a general trend found in the sociolinguistics literature (Brandão and Vieira, 2012; Naro and Scherre, 2013; Chappell, 2015; Oushiro, 2015). The use of non-standard language has been attributed to level of bilingualism. While level of education is different than level of bilingualism, speakers with higher levels of education also have more exposure to prescriptive language via the classroom.

8.3 Limitations

This dissertation contains some notable limitations regarding sample size and the number of participants. The analysis of SPE included 10 NS participants and 10 MCS participants, while the analysis of variable number marking included 20 MCS speakers. This limited number of participants was due primarily to difficulties in data collection, related to political unrest in Nicaragua from 2018-2019 and the COVID-19 pandemic. It is important to note that an increased participant group may better reflect the conditioning factors involved in SPE and number marking in Nicaraguan varieties of Spanish. Due to the small number of participants, the sample sizes were also fairly limited: 783 tokens for SPE and 595 tokens for variable number marking. As this sample size is not necessarily representative of the entire MCS population, any claims I make are based solely on the data specific to this dissertation. While I do suggest that the data presented here offers evidence for the inter-generational transfer of features, the sample size must be considered as a limitation. The statistical analysis did however follow guidelines proposed by Hosmer and Lemeshow (2000, pp. 346-347), and cited in Levshina (2015, p. 257), regarding the number of factors that should be included in any given model, according to data set size. With that in mind, I believe that the research presented in this dissertation is an important contribution to the field of Hispanic linguistics, particularly the study

of indigenous languages in contact with Spanish. As both MCS, and to a lesser extent, NS are both understudied varieties of Spanish, this work offers insight into their distribution of morphosyntactic features. The findings in this dissertation, while based on a small sample of speakers, will allow for more informed predictions in future research projects about NS and MCS.

An additional limitation in this research involves the amount of information published about the Miskitu community and language. While several books and articles have been published about both the history of contact with outside groups and the Miskitu grammar, many gaps still exist in the literature, including, but not limited to, discourse-pragmatic constraints in Miskitu, bilingualism in the RAAN, and current contact with monolingual Spanish speakers in Miskitu communities. In this dissertation, I attempt to use available resources to the best of my ability in order to provide relevant background about MCS and its speakers. While the study did not include a bilingual profile of the participants, each person was asked a variety of questions related to language use during the sociolinguistic interview. A qualitative analysis of these findings was included in the methodology section in order to provide a better understanding of the linguistic demographics of the participants.

8.4 Contributions and Future Directions

This dissertation examined two linguistic features in MCS: subject pronoun expression and variable number marking. MCS is a contact variety that emerged from a case of rapid language shift along the Caribbean coast of Nicaragua, beginning in the middle of the twentieth century. As a result, the incomplete acquisition of SPE and number marking occurred, both of which differ in their distribution in Spanish and Miskitu. During initial language shift, MCS speakers relied on both cognitive universals as well as their L1 to fill gaps due to these differences and to compensate for incomplete acquisition. These strategies are becoming less relevant in younger generations as speakers are now more exposed to standard language norms and as a result are approximating closer to NS trends.

Future studies of these features would benefit from a better understanding of Miskitu, particularly the distribution of explicit versus implicit pronouns for singular referents. The available grammars do not discuss pragmatic constraints for SPE in Miskitu, or more generally across the language. In addition, it would be beneficial to analyze first and second person forms for both SPE and variable number marking, and include a more precise way to measure Spanish competency, as this may reveal an effect for level of bilingualism that would

contribute to the understanding of these features. Finally, additional analyses of morphosyntactic features in would further help to clarify language internal and language external factors that motivate variation in MCS. My previous research and observations of this contact variety include object pronoun use, /s/ deletion on 2sg. verb forms, and lack of gender agreement between nouns and adjectives. More information about these features in MCS would provide more support for the learner strategies discussed in this dissertation.

Examining the distribution of two distinct features in MCS allows for a more consistent and robust conclusion in regard to what processes are at play in this contact situation. In addition to expanding the literature on SPE in contact varieties of Spanish, this dissertation shed light on variable number marking, a feature that has only been attested in a few varieties of Spanish. The findings for the statistical analyses confirm that there is indeed evidence of that incomplete acquisition of the features under investigation occurred in the past, with diverging distributions in MCS, compared to NS. In addition, the linguistic factors analyzed revealed strategies used by speakers to compensate for lack of input during language shift. They drew on not only their L1 but universal cognitive patterns, such as salience, to navigate the acquisition of a typologically different language. Finally, the social factors analyzed show some level of stigmatization attached to the two features and suggest that a shift toward a more standard distribution of explicit subject pronouns and number marking on verb morphology is taking place in the community on the part of younger speakers. These findings indicate that a possible change is in progress in regard to SPE and variable number marking in MCS. The findings in this dissertation provide an example of the strategies speakers use when acquiring a second language in a natural contact situation. They also illustrate how these strategies can impact morphosyntactic features in Spanish and how this might change when current speakers have continued exposure to standard language norms.

Appendix A

Interview Questions with MCS Speakers

1. ¿De dónde es Usted?
2. ¿De dónde es su familia?
3. Cuénteme de su familia.
4. ¿En qué trabaja Usted?
5. ¿Qué tipos de trabajos existen aquí?
6. ¿Ha viajado Usted fuera de Puerto Cabezas?
7. ¿Qué hace Usted en su tiempo libre?
8. ¿Cuáles idiomas habla Usted?
9. ¿Cuándo empezó a aprender cada lengua?
10. ¿Dónde aprendió cada lengua?
11. ¿En qué situaciones utiliza cada lengua?
12. ¿Cuénteme de los diferentes grupos de personas que viven aquí?
13. ¿Cuénteme de la comunidad miskita aquí?
14. ¿Cómo es la cultura miskita?
15. ¿Cómo se relaciona la gente miskita aquí?
16. ¿Qué va a pasar con la lengua miskita en esta comunidad?
17. ¿Cuáles son los pensamientos sobre la gente del Pacífico?
18. Cuénteme una leyenda cultural.
19. ¿Me puede contar una historia o un recuerdo de su niñez?
20. ¿Cuáles son sus metas para el futuro?
21. ¿Cómo es la política aquí?
22. ¿Qué piensan del medio ambiente aquí?
23. ¿Se escucha mucho del canal interoceánico?
24. ¿Qué está pasando con los pescadores y los buzos?
25. ¿Cuál es el problema más grande en su comunidad?

26. ¿En su cultura cuáles son los papeles de las mujeres y de los hombres?
27. ¿Se nota alguna diferencia entre las generaciones en cuanto a sus pensamientos?
28. ¿Hay muchas personas que hablan el inglés aquí?
29. ¿Cómo son las escuelas aquí?
30. ¿En qué idioma reciben clase los estudiantes?
31. ¿A qué edad los niños empiezan a aprender el español aquí?
32. ¿Me puede contar un poco de Sikru Tara?
33. ¿Cuáles son las comidas típicas aquí?
34. ¿Cuál es la relación que tienen aquí con los miskitos de Honduras?
35. ¿Cuál es la diferencia entre la lengua miskita de Nicaragua y de Honduras?
36. ¿Cómo son las comunidades afuera de la ciudad?
37. ¿Hay mucho turismo aquí?
38. ¿Dónde hay un lugar bonito para visitar aquí?
39. ¿Muchas personas de aquí viajan o viven en otros lugares?
40. ¿Qué es lo que le gusta más de su cultura?
41. ¿Qué no le gusta de su cultura?
42. ¿Conoce a unos miskitos que viven en el sur? ¿Piensa que su nivel de orgullo sobre la cultura miskita es igual a la de aquí?

Bibliography

Abreu, L. (2012). Subject pronoun expression and priming effects among bilingual speakers of puerto rican spanish (K. Geeslin & M. Díaz-Campos, Eds.). In K. Geeslin & M. Díaz-Campos (Eds.), *Selected proceedings of the 14th Hispanic Linguistics Symposium*. Citeseer.

Alba, O. (2004). *Cómo hablamos los dominicanos: Un enfoque sociolingüístico*. Grupo León Jimenes.

Badlato, M. P. (2000). *Miskitu discourse*. The University of Texas at Austin.

Barnes, H. (2010). Subject pronoun expression in bilinguals of two null subject languages (K. Arregi, Z. Fagyal, S. Montrul, & A. Tremblay, Eds.). In K. Arregi, Z. Fagyal, S. Montrul, & A. Tremblay (Eds.), *Romance linguistics 2008: Interactions in Romance. Selected papers from the 38th Linguistic Symposium on Romance Languages, Urbana-Champaign, April 2008*. John Benjamins Publishing.

Barrenechea, A. M., & Alonso, A. (1973). Los pronombres personales sujetos en el español hablado en Buenos Aires, In *Studia iberica: Festschrift für hans flasche*. Francke Verlag.

Bentivoglio, P. (1987). *Los sujetos pronominales de primera persona en el habla de caracas* (Vol. 13). Universidad Central de Venezuela, Consejo de Desarrollo Científico y Humanístico.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning?. In *Linguistic perspectives on second language acquisition* (pp. 41–68). Cambridge University Press.

Brandão, S. F., & Vieira, S. R. (2012). Concordáncia nominal e verbal: Contribuições para o debate sobre o estatuto da variação em três variedades urbanas do portugués. *Alfa: Revista de Lingística (São José do Rio Preto)*, *56*(3), 1035–1064.

Bruyne, J. D., & Pountain, C. J. (1995). *A comprehensive spanish grammar*. Blackwell Publishing.

Butler, J. (2000). Yatama: Rebellion with a cause? *Revista Envío*, (232).

Camacho, J. (2013). *Null subjects. [electronic resource]*. Cambridge University Press.

Cameron, R. (1992). *Pronominal and null subject variation in spanish: Constraints, dialects, and functional compensation* (Doctoral dissertation). University of Pennsylvania.

Cameron, R. (1993). Ambiguous agreement, functional compensation, and nonspecific tú in the spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change*, *5*(3), 305–334.

Cameron, R. (1996). A community-based test of a linguistic hypothesis. *Language in Society*, 61–111.

Carvalho, A. M., & Bessett, R. M. (2015). Subject pronoun expression in spanish in contact with portuguese. In A. Carvalho, R. Orozco, & N. L. Shin (Eds.), *Subject pronoun expression in spanish: A cross-dialectal perspective* (pp. 143–168). Georgetown University Press Washington, DC.

Catter, T. V. (2011). Intercultural bilingual education in nicaragua: Contextualisation for improving the quality of education. *International Review of Education / Internationale Zeitschrift für Erziehungswissenschaft / Revue Internationale de l'Education*, *57*(5/6), 721–735.

Chappell, W. (2013). *Social and linguistic factors conditioning the glottal stop in nicaraguan spanish* (Doctoral dissertation). The Ohio State University.

Chappell, W. (2014). Reanalysis and hypercorrection among extreme/s/reducers. *University of Pennsylvania Working Papers in Linguistics*, *20*(2), 5.

Chappell, W. (2015a). Formality strategies in managua, nicaragua: A local vs. global approach. *Spanish in Context*, *12*(2), 221–254.

Chappell, W. (2015b). Linguistic factors conditioning glottal constriction in nicaraguan spanish. *Italian Journal of Linguistics*, *27*(2), 1–42.

Chappell, W. (2016). Bilingualism and aspiration. *Spanish Language and Sociolinguistic Analysis*, *8*, 261–284.

CIA. (2019). Nicaragua. https://www.cia.gov/library/publications//the-world-factbook/geos/nu.html

CIDCA. (1985). *Miskitu bila aisanka. gramática misquita*. Managua.

Craig, C. G. (1985). Current knowledge of amerindian languages of nicaragua. *International Journal of American Linguistics*, *51*(4), 381–384.

Critchfield, M., & Lívio, C. (forthcoming). Variable concord in two american romance varieties: Using brazilian portuguese to inform mosquito coast spanish. In C. Howe, P. Chamorro, T. Gupton, & P. Renwick (Eds.), *Proceedings from the 49th linguistics symposium on romance languages*. University of Chicago Press.

Davidson, B. (1996). 'Pragmatic weight' and spanish subject pronouns: The pragmatic and discourse uses of 'tú' and 'yo' in spoken madrid spanish. *Journal of Pragmatics*, *26*(4), 543–565.

Dennis, P. A. (2004). *The miskitu people of awastara*. University of Texas Press.

Domınguez, L., Hicks, G., & Slabakova, R. (2019). Terminology choice in generative acquisition research: The case of "incomplete acquisition" in heritage language grammars. *Studies in Second Language Acquisition*, *41*(2), 241–255.

Dozier, C. L. (1985). *Nicaragua's mosquito shore: The years of british and american presence*. University Alabama Press.

Duarte, M. E., & Soares da Silva, H. (2016). Microparametric variation in spanish and portuguese. In M. A. Kato & F. Ordóñez (Eds.), *The morphosyntax of portuguese and spanish in latin america*. Oxford University Press.

Ellis, N. (2017). Salience in usage-based SLA. In S. M. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 21–40). Routledge.

Escobar, A. M. (2013). Bilingualism in latin america. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (p. 725). Wiley Online Library.

Ethnologue. (2020). Nicaragua. https://www.ethnologue.com/country/ni/languages

Floyd, T. S. (1967). *The anglo-spanish struggle for mosquitia*. University of New Mexico Press.

Foote, R., & Bock, K. (2012). The role of morphology in subject–verb number agreement: A comparison of mexican and dominican spanish. *Language and Cognitive Processes*, *27*(3), 429–461.

García, C. (1996). *The making of the miskitu people of nicaragua : The social construction of ethnic identity*. Universitetet.

Geeslin, K., & Gudmestad, A. (2011). Using sociolinguistic analyses of discourse-level features to expand research on l2 variation in forms of spanish subject expression (L. Plonsky & M. Schierloh, Eds.). In L. Plonsky & M. Schierloh (Eds.), *Selected proceedings of the 2009 Second Language Research Forum*. Cascadilla Proceedings Project Somerville, MA.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274–307.

Guy, G. (1981). *Linguistic variation in brazilian portuguese: Aspects of the phonology, syntax, and language history* (Doctoral dissertation). University of Pennsylvania.

Guy, G. (2005). A questão da crioulização no português do brasil (A. M. S. Zilles, Ed.). In A. M. S. Zilles (Ed.), *Estudos de variação lingüística no*

brasil e no cone sul. Editora da Universidade Federal do Rio Grande do Sul.

Guy, G. (2014). Variation and change in latin american spanish and portuguese. In P. Amaral & A. M. Carvalho (Eds.), *Portuguese-spanish interfaces: Diachrony, synchrony, and contact* (pp. 443–464). John Benjamins.

Guy, G. (2017). The african diaspora in latin america. In C. Cutler, Z. Vrzić, & P. Angermeyer (Eds.), *Language contact in africa and the african diaspora in the americas* (pp. 49–78).

Heath, G. R. (1927). *Grammar of the Miskito language*. F. Lindenbein.

Helms, M. W. (1971). Asang: Adaptations to culture contact in a misquito community. University of Florida Press. Gainesville.

Hernández Campoy, J. M., & Trudgill, P. (2002). Functional compensation and southern peninsular spanish/s/loss. *Folia Linguistica Historica*, *23*, 141–167.

Hochberg, J. (1986a). /s/ deletion and pronoun usage in Puerto Rican Spanish. In D. Sankoff (Ed.), *Diversity and diachrony*. Philadelphia: John Benjamins.

Hochberg, J. (1986b). Functional compensation for/s/ deletion in Puerto Rican Spanish. *Language*, *62*(3), 609–621.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. Wiley.

INETER. (2020). Mapa de la división política departamental. http://www.ineter.gob.ni/Geodecia/files/nicaragua.jpg.

Jiménez Sabater, M. A. (1977). Estructuras morfosintácticas en el español dominicano: Algunas implicaciones sociolingüisticas. *Ciencia y Sociedad*.

Kessler, L. (1992). *Here comes the strikeout*. Harper Collins.

Kupisch, T., & Rothman, J. (2018). Terminology matters! why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, *22*(5), 564–582.

Labov, W. (1969). Contraction, deletion, and inherent variability of the english copula. *Language*, 715–762.

Labov, W. (2001). *Principles of linguistic change volume 2: Social factors*. Blackwell.

Lafford, B. (1986). Valor diagnóstico-social del uso de ciertas variantes de/s/en el español de cartagena, colombia. *Estudios sobre la fonologıa del español del Caribe*, *53*, 74.

Lapidus, N., & Otheguy, R. (2005). Contact induced change? overt nonspecific ellos in spanish in new york (L. Sayahi & M. Westmoreland, Eds.). In L. Sayahi & M. Westmoreland (Eds.), *Selected proceedings of the 2nd workshop on spanish sociolinguistics*. Cascadilla Somerville, MA.

Lastra, Y., & Butragueño, P. M. (2015). Subject pronoun expression in oral mexican spanish. In A. Carvalho, R. Orozco, & N. L. Shin (Eds.), *Subject pronoun expression in spanish: A cross-dialectal perspective* (p. 57). Georgetown University Press Washington, DC.

Levinson, S. C. (2000). *Presumptive meanings : The theory of generalized conversational implicature*. MIT Press.

Levshina, N. (2015). *How to do linguistics with r: Data exploration and statistical analysis*. John Benjamins Publishing Company.

Limerick, P. P. (2017). Language contact in the us southeast: The case of spanish subject expression in an emerging bilingual community in georgia. *Spanish in Context*, *14*(1), 53–78.

Lipski, J. (1984). /s/ in the spanish of nicaragua. *Orbis*, *33*(1-2), 171–181.

Lipski, J. (1985). /s/ in central american spanish. *Hispania*, *68*(1), 143–149.

Lipski, J. (1993). Creoloid phenomena in the spanish of transitional bilinguals. In A. Roca & J. M. Lipski (Eds.), *Spanish in the united states: Linguistic contact and diversity* (pp. 155–182). Mouton de Gruyter New York.

Lipski, J. (1996). *El español de américa*. Cátedra Madrid.

Lipski, J. (2008a). *Afro-bolivian spanish*. Iberoamericana Vervuert.

Lipski, J. (2008b). Afro-bolivian spanish and helvécia portuguese: Semi-creole parallels. *PAPIA-Revista Brasileira de Estudos do Contato Linguístico*, *16*(1), 96–116.

Lipski, J. (2008c). *Varieties of spanish in the united states*. Georgetown University Press.

Lipski, J. (2011). Socio-phonological variation in latin american spanish. In M. Díaz-Campos (Ed.), *The handbook of hispanic sociolinguistics* (p. 146). Wiley Online Library.

Lozano, C. (2009). Pronominal deficits at the syntax-discourse interface: A developmental study of greek learners of non-native spanish, In *Workshop on interfaces in l2 acquisition, universidade nova de lisboa, june, available at: Http://www. ugr. es/cristoballozano*.

Lucchesi, D., Baxter, A. N., & Ribeiro, I. (2009). *O portugués afro-brasileiro*. Edufba.

Lynch, A. (2009). A sociolinguistic analysis of final /s/ in miami cuban spanish. *Language Sciences*, *31*(6), 766–790.

Mendes, R. B., & Oushiro, L. (2015). Variable number agreement in brazilian portuguese: An overview. *Language and Linguistics Compass*, *9*(9), 358–368.

Meringer, E. R. (2014). Accommodating mestizaje on nicaragua's rio coco: Miskitu activism before the sandinista revolution. *AlterNative: An International Journal of Indigenous Peoples*, *10*(3), 203–215.

Michaelis, S. M., Maurer, P., Haspelmath, M., & Huber, M. (2013). *The atlas of pidgin and creole language structures*. Oxford University Press.

Michnowicz, J. (2015). Subject pronoun expression in contact with maya in yucatan spanish. In A. Carvalho, R. Orozco, & N. L. Shin (Eds.), *Subject pronoun expression in spanish: A cross-dialectal perspective* (pp. 101–120). Georgetown University Press Washington, DC.

MINED. (2019). Bandera de nicaragua. https://nicaraguaeduca.mined.gob.ni/bandera-nacional/.

The miskitos in nicaragua, 1981-1984. (1984). Americas Watch Committee.

Monguilhott, I. d. O. Et al. (2009). *Estudo sincrônico e diacrônico da concordância verbal de terceira pessoa do plural no pb e no pe*. Florianópolis, SC.

Montrul, S. (2002). Incomplete acquisition and attrition of spanish tense/aspect distinctions in adult bilinguals. *Bilingualism*, *5*(1), 39.

Montrul, S. (2004). Subject and object expression in spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism*, *7*(2), 125.

Montrul, S. (2006). Incomplete acquisition as a feature of bilingual and l2 grammars. In *Inquiries in linguistic development: In honor of lydia white* (pp. 335–359). John Benjamins Amsterdam.

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. John Benjamins Publishing.

Montrul, S. (2018). Heritage language development: Connecting the dots. *International Journal of Bilingualism*, *22*(5), 530–546.

Montrul, S., & Silva-Corvalán, C. (2019). The social context contributes to the incomplete acquisition of aspects of heritage languages. *Studies in Second Language Acquisition*, *41*(2), 269–273.

Naro, A. J. (1981). The social and structural dimensions of a syntactic change. *Language*, 63–98.

Naro, A. J., & Scherre, M. M. P. (2000). Variable concord in portuguese: The situation in brazil and portugal. *Creole Language Library*, *21*, 235–256.

Naro, A. J., & Scherre, M. M. P. (2013). Remodeling the age variable: Number concord in brazilian portuguese. *Language Variation and Change*, *25*(1), 1.

Olloqui de Montenegro, L. d. (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, San Juan, Puerto Rico: Academia Puertorriqueña de la Lengua Española.

Orozco, R., & Guy, G. (2008). El uso variable de los pronombres sujetos: ?‘ Qué pasa en la costa caribe colombiana (M. Westmoreland & J. A. Thomas, Eds.). In M. Westmoreland & J. A. Thomas (Eds.), *Selected proceedings of the 4th Workshop on Spanish Sociolinguistics*. Cascadilla Proceedings Project Somerville, MA.

Otheguy, R. (2016). The linguistic competence of second-generation bilinguals (C. Tortora, M. denDikken, I. L. Montoya, & T. O'Neill, Eds.). In C. Tortora, M. denDikken, I. L. Montoya, & T. O'Neill (Eds.), *Romance linguistics 2013: Selected papers from the 43rd Linguistic Symposium on Romance Languages, New York, 17-19 April, 2013*. John Benjamins Publishing Company.

Otheguy, R., & Zentella, A. C. (2012). *Spanish in new york : Language contact, dialectal leveling, and structural continuity.* Oxford University Press.

Otheguy, R., Zentella, A. C., & Livert, D. (2007). Language and dialect contact in spanish in new york: Toward the formation of a speech community. *Language*, 770–802.

Oushiro, L. (2015). *Identidade na pluralidade: Avaliação, produção e percepção linguıstica na cidade de são paulo* (Doctoral dissertation). Universidade de São Paulo.

Pascual Y Cabo, D., & Rothman, J. (2012). The (il) logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics*, *33*(4), 450–455.

Penny, R. J. (2000). *Variation and change in spanish*. Cambridge University Press.

Pereira, D. C. (2004). *Concordância verbal na lıngua falada nas trilhas das bandeiras paulistas* (Doctoral dissertation). Universidade de São Paulo.

Prada Pérez, A. d. (2009). *Subject expression in minorcan spanish: Consequences of contact with catalan*. The Pennsylvania State University.

Prada Pérez, A. d. (2015). First person singular subject pronoun expression in spanish in contact with catalan. In A. Carvalho, R. Orozco, & N. L. Shin (Eds.), *Subject pronoun expression in spanish: A cross-dialectal perspective* (pp. 121–142). Georgetown University Press Washington, DC.

Quesada, M. L. (2015). *The l2 acquisition of spanish subjects : Multiple perspectives.* De Gruyter Mouton.

Quesada, M. L., & Blackwell, S. E. (2009). The l2 acquisition of null and overt spanish subject pronouns: A pragmatic approach (J. Collentine, M. García, B. Lafford, & F. M. Marín, Eds.). In J. Collentine, M. García, B.

Lafford, & F. M. Marín (Eds.), *Selected proceedings of the 11th Hispanic Linguistics Symposium*. Citeseer.

Ranson, D. L. (1991). Person marking in the wake of /s/ deletion in andalusian spanish. *Language Variation and Change*, *3*(2), 133–152.

Rey, A. (1995). Social correlates of the "voseo" of Managua, Nicaragua: Family and neighborhood domains. *Hispanic Journal*, 39–53.

Rey, A. (1997). Social correlates of the "voseo" of Managua, Nicaragua: Workplace, street, and party domains. *Hispanic Journal*, 109–126.

Rodrigues, A. (1987). *A concordância verbal no português popular em são paulo* (Doctoral dissertation). São Paulo: FFLCH-USP.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, *13*(2), 155–163.

Rowland, D. (1935). Spanish occupation of the island of old providence, or santa catalina, 1641-1670. *The Hispanic American Historical Review*, *15*(3), 298–312.

Rubio, C. F., & Gonçalves, S. C. L. (2012). A fala do interior paulista no cenário da sociolinguıstica brasileira: Panorama da concordância verbal e da alternância pronominal. *Alfa: Revista de Linguıstica (São José do Rio Preto)*, *56*(3), 1003–1034.

Salamanca, D. (1988). *Elementos de gramática del miskito* (Doctoral dissertation). Massachusetts Institute of Technology.

Schachter, J. (1990). On the issue of completeness in second language acquisition. *Interlanguage Studies Bulletin (Utrecht)*, *6*(2), 93–124.

Scherre, M. M. P. (1998). Variação da concordância nominal no português do brasil: Influência das variáveis posição, classe gramatical e marcas precedentes. In S. Grosse & K. Zimmermann (Eds.), *Substandard" e mudança no português do brasil* (pp. 153–188).

Scherre, M. M. P., & Naro, A. J. (1992). The serial effect on internal and external variables. *Language Variation and Change*, *4*(1), 1–13.

Scherre, M. M. P., & Naro, A. J. (2001). Sobre as origens estruturais do português brasileiro: Crioulização ou mudança natural? *PAPIA-Revista Brasileira de Estudos do Contato Linguıstico*, *11*(1), 40–50.

Scherre, M. M. P., & Naro, A. J. (2010). Perceptual vs. grammatical constraints and social factors in subject-verb agreement in brazilian portuguese. *University of Pennsylvania Working Papers in Linguistics*, *16*(2), 20.

Scherre, M. M. P., & Naro, A. J. (2014). Sociolinguistic correlates of negative evaluation: Variable concord in rio de janeiro. *Language Variation and Change*, *26*(3), 331.

Schwenter, S. (2011). Variationist approaches to spanish morphosyntax: Internal and external factors. In M. Díaz-Campos (Ed.), *The handbook of hispanic sociolinguistics* (pp. 123–147). Wiley Online Library.

Sessarego, S. (2011). On the status of afro-bolivian spanish features: Decreolization or vernacular universals (J. Michnowicz & R. Dodsworth, Eds.). In J. Michnowicz & R. Dodsworth (Eds.), *Proceedings of the Fifth Workshop on Spanish Sociolinguistics*.

Sessarego, S. (2012). Non-creole features in the verb system of afro-hispanic languages: New insights from sla studies. *International Journal of Linguistics*, *4*(1), 146.

Shin, N. L. (2013). Women as leaders of language change: A qualification from the bilingual perspective (A. M. Carvalho & S. Beaudrie, Eds.). In A. M. Carvalho & S. Beaudrie (Eds.), *Proceedings of the 6th international workshop on spanish sociolinguistics*. Citeseer.

Silva-Corvalán, C. (1994). *Language contact and change : Spanish in los angeles.* Clarendon Press.

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

Silva-Corvalán, C. (2018). Bilingual acquisition. In N. Shin & D. Erker (Eds.), *Questioning theoretical primitives in linguistic inquiry: Papers in honor of Ricardo Otheguy* (pp. 245–268). John Benjamins Publishing Company.

Solomon, J. (1998). External constraints and variable subject expression in yucatec spanish, In *Annual meeting of New Ways of Analyzing Variation 27, Athens, Georgia*.

Solomon, J. (1999). *Phonological and syntactic variation in the spanish of valladolid, yucatán* (Doctoral dissertation). Stanford University.

Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in non native grammars of italian. *Second Language Research*, *9*(1), 22–47.

Sorace, A. (2004). Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition*, *7*(2), 143–145.

Sorace, A. (2005). Selective optionality in language development. In L. Cornips & K. P. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and the social* (pp. 55–80).

Tagliamonte, S. A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. John Wiley & Sons.

Torres Cacoullos, R., & Travis, C. E. (2011). Testing convergence via code-switching: Priming and the structure of variable subject expression. *International Journal of Bilingualism*, *15*(3), 241–267.

Valdés, G. (2001). *Learning and not learning English: Latino students in American schools*. Teachers College Press.

Wilson Withe, C., Oporta, P., Muller Oporta, J., Chow Smith, W., Bency, S., Taylor, A., Alvarado, M., Ocampo, J., Morazan Hodgson, E., Marley, S., Et al. (2012). *Sari laka apu kan piuara: Antes de los dias tristes: Cuaderno cultural miskitu.* (tech. rep.). UNESCO, San Jose (Costa Rica). Programa Conjunto de Revitalizacion Cultural.

Winford, D. (2003). *An introduction to contact linguistics*. Wiley-Blackwell.

Winford, D. (2005). Contact-induced changes: Classification and processes. *Diachronica*, *22*(2), 373–427.

Zanini, C., Rugani, R., Giomo, D., Peressotti, F., & Franzon, F. (2020). Effects of animacy on the processing of morphological number: A cognitive inheritance? *Word Structure*, *13*(1), 22–44.