

Subject Pronoun Expression in an L2-only Environment:

The Case of Equatorial Guinea

by

Lillie Vivian K. Padilla

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2020 by the
Graduate Supervisory Committee:

Álvaro Cerrón-Palomino, Chair
Barbara Lafford
Omar Beas
Joseph-Désiré Otabela

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

Subject Pronoun Expression (SPE) has been extensively studied in monolingual and bilingual varieties of Spanish using the variationist framework. The goal of these studies has been to examine the linguistic and extra-linguistic factors that condition the expression and the omission of personal subject pronouns. Nonetheless, to date, there is no study of SPE in the Spanish of Equatorial Guinea, the only African country where it is an official language, and the single country where Spanish is exclusively a second language (L2). This dissertation fills this gap in the literature by accounting for SPE in Equatoguinean Spanish.

The research questions guiding this study concern the rates of Subject Pronoun Expression, its conditioning factors, and universal accounts of L2 acquisition, in particular, the Interface Hypothesis (IH). The study had 30 participants from Malabo, who took part in sociolinguistic interviews. These interviews were transcribed and analyzed using the mixed effects software Rbrul. Along the lines of the literature reviewed, the linguistic factor groups studied were grammatical person and number, reference, reflexivity, verb type, and ambiguity. By the same token, the extra linguistic factors analyzed were age, sex, education, native language (L1), and speaker as a random factor.

The results indicate that the Equatoguinean variety of Spanish has one of the lowest pronoun rates (19.1%), a finding that goes against the predictions of the IH. With regard to the linguistic factor groups that condition Subject Pronoun Expression, Equatoguinean Spanish shows an unorthodox ranking: grammatical person and number, ambiguity, verb class, and reference. Interestingly, the low ranking of reference gives

support to the IH, which argues that L2 speakers have problems with constraints like the switch of the reference in subjects because it integrates discourse and pragmatic interfaces. The only significant extra-linguistic factor was education, whereas speakers' L1 exerted no effect on SPE. Individual speaker was a significant random factor group, indicating that variation is great even in speakers with comparable education.

In sum, this study of a unique speech community provides new information on SPE of L2 Spanish. It also contributes to the fields of language contact, language variation, and second language acquisition.

ACKNOWLEDGMENTS

I would like to express my deep gratitude to my dissertation committee members Dr. Alvaro Cerrón-Palomino, Dr. Barbara Lafford, Dr. Omar Beas and Dr. Joseph-Désiré Otabela for their patience, professional guidance, encouragement and critique of this project. I would also like to thank Dr. Sara Beaudrie and Dr. Michael Gradoville for their feedback during every presentation of this project-To each of you, my deepest thanks. I could not have done it without you!

I am also grateful to my friends as SILC, who have become my family during these five years-Fred, Rosti, Angelica, Nayibe, Nancy, Silvana, Obenewaa, Melisa, Edurne and many others. I am grateful for the support, encouragement and all the good times we shared as colleagues. Also, I would like to thank my uncle, Dr. Victor Lamptey for encouraging me to pursue a graduate degree in Spanish. I would also like to thank my parents, Mr. Edmund Essah and Mrs. Sophia Essah, who have been with me from the very beginning of this journey. They are my role models. Most importantly, I wish to thank my kind, loving and supportive husband, Eddie, and my two lovely children, Jayden and Justin, who inspire me to be my very best.

Finally, this dissertation would not have been possible without the financial support of the School of International Letters and Cultures at Arizona State University. Apart from making the collection of data possible, the financial support allowed me to participate in professional conferences and receive feedback on segments of the dissertation.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
1 INTRODUCTION	1
Sociolinguistic Variation in the Speech of Native Bilingual and L2 Speakers	4
Universal Tendencies in L2 Speech.....	5
Organization of the Dissertation.....	8
2 A BRIEF OVERVIEW OF EQUATORIAL GUINEA	10
Colonization and European Settlement.....	10
Independence and Post Colonization	13
Linguistic Overview of Native Languages in Equatorial Guinea	15
Spanish in Equatorial Guinea.....	18
Summary.....	20
3 THEORETICAL ASPECTS	21
Variationist Sociolinguistics.....	21
Language Contact.....	24
Interface Hypothesis.....	30
Summary	33
4 LITERATURE REVIEW	34
SPE in Monolingual Varieties of Spanish	35

CHAPTER	Page
SPE in Bilingual Varieties of Spanish	45
Previous Studies on Equatoguinean Spanish.....	54
Justification of the Present Study.....	57
Summary	58
5 METHODOLOGY	60
Research Questions	61
Participants.....	64
Data Collection.....	67
Data Analysis.....	71
Factor Groups Analyzed and their Predictions.....	73
Linguistic Factors.....	73
Extra-linguistic factors.....	77
Summary.....	80
6 RESULTS AND DISCUSSION.....	82
Analyzing with Rbrul.....	82
Research Question 1.....	85
Research Question 2.....	87
Research Question 3.....	95
Research Question 4.....	97
A Comparison of Pronoun Rates.....	99
Significant Factors.....	101

CHAPTER	Page
L2 Speakers and the Discourse/Pragmatic Interface.....	106
Summary.....	108
7 CONCLUSION	109
Contribution to the Field of Language Variation Research.....	109
Contribution to the Field of Language Contact Research	111
Contribution to Second Language Acquisition Studies.....	111
Limitations of the Study.....	111
Future Research.....	113
Summary.....	114
REFERENCES	117
APPENDIX	
A LANGUAGE BACKGROUND QUESTIONNAIRE	134
B INTERVIEW QUESTIONS	136
C IRB APPROVAL	139

LIST OF TABLES

Table	Page
1. Participant Information	66
2. Overall Distribution of Null and Overt Pronouns in Equatoguinean Spanish.....	85
3. Rbrul Analysis of Significant Linguistic Factors	88
4. Rbrul Analysis of Significant Extra-Linguistic Factors.....	91
5. Rbrul Analysis of Overt SPE by Speakers.....	92
6. Non-Significant Linguistic Predictors of Overt SPE in Equatoguinean Spanish.....	93
7. Non-Significant Extra-Linguistic Predictor of Overt SPE in Equatoguinean Spanish.....	94
8. Overt SPE's by Rate.....	96
9. Cross Tabulation of Education and Person/Number.....	104

LIST OF FIGURES

Figure	Page
1. Map of Local Languages Spoken in Equatorial Guinea	17

CHAPTER 1

INTRODUCTION

Subject pronoun expression (SPE) is one of the most studied topics in sociolinguistics (Bayley, Cárdenas, Schouten & Salas, 2012; Carvalho, Orozco, & Shin, 2015; Otheguy & Zentella, 2012; Silva-Corvalán, 1982). The reason for this interest is that the features +/- SPE groups languages into those that allow for the expression of a null subject pronoun (e.g., Spanish, Italian, Greek) and languages that do not (e.g., English, French). The following examples illustrate this tendency in Spanish.

(1) *Porque la música española **Ø** escucho bien lo que dice, en inglés también una parte **Ø** escucho.* (Speaker 6)

Because Spanish music, I hear very well what it says, in English too, I understand a portion.

(2) *Todo lo que **yo** veo en mi país, me gusta, está bien* (Speaker 10)

Everything that I see in my country, I like it, it's good.

In these examples and throughout the rest of this dissertation, overt SPPs will be represented in bold, and null pronouns will be represented by **Ø**. In sentence (1) there is a null subject (**Ø** *escucho*) whereas in (2) there is an overt subject (**yo** *escucho*).

Nonetheless, because Spanish is a language with a rich verbal suffix morphology, the referent of (1) can still be inferred as the first-person singular form of the verb due to the fact that it ends in “o”.

Spanish thus allows for the expression of null subjects in all tenses and in all grammatical persons, and the presence or the absence of the subject pronoun does not

change the basic meaning of any of the utterances. The choice, however, to use a null or overt subject pronoun is not random. It is governed by semantic, discourse and pragmatic constraints. For this reason, there is immense research in generative studies, second language studies and sociolinguistic studies regarding the expression of the subject pronoun. For instance in studies in the field of syntax, the focus has been on the characteristics of the Pro-drop parameter (e.g. Toribio, 1994) and the Null Subject Parameter (e.g. Lozano, 2002). In the field of Second Language Acquisition, the focus has been on resetting parameters in the second language (e.g. Lafond, Hayes, & Bhatt, 2001), as well as the pragmatic rules regarding Subject Pronoun Expression (e.g. Lozano, 2009). In the field of sociolinguistics, the focus has been on the linguistic and extra-linguistic factors that condition the expression and the omission of the subject pronoun. An exploration of these factors will be the focus of the present study.

Several studies have employed the *variationist sociolinguistics approach*, an approach that views language variation as orderly and structured, to examine the linguistic and extra-linguistic factors that affect the alternation of the null and expressed pronoun in monolingual and bilingual varieties of Spanish spoken in the Caribbean, Americas and Europe (e.g., Barrenechea & Alonso, 1977; Bayley & Pease-Alvarez, 1997; Bentivoglio, 1987; Cameron, 1993; Erker & Guy, 2012; Flores-Ferrán, 2004; Morales, 1980; Orozco & Guy, 2008; Otheguy & Zentella, 2007, 2012; Otheguy, Zentella & Livert, 2007; Posio, 2011; Prada Pérez, 2009; Shin & Otheguy, 2013; Silva-Corvalán, 1982, 1994). However, despite this extensive research, there is very little information on factors that condition SPE in the variety of Spanish spoken in Africa,

specifically, in Equatorial Guinea, the only African country with Spanish as an official language.

This question has never been previously addressed because there is very little literature on SPE in the Equatoguinean variety of Spanish. Previous studies conducted on this variety have almost exclusively focused on dialectology, for instance, the works of Castillo Barril (1964), Echegaray (1951), De Granda (1985, 1990), Lipski (1984, 2002, 2004, 2008), and Quilis and Fresnillo (1995). Others have focused on phonology (e.g., Granda, 1985; Quilis, 1992), semantics (Nistal, 2009; Tiogang, 2007), syntax (Vuskovich, 2013) and language attitudes (Gomashie, 2019). To date, Essah (forthcoming) is so far the only study that has examined Spanish in Equatorial Guinea using the language variation approach in sociolinguistics.

Thus, the present study fills this gap in the literature by examining SPE in Equatoguinean Spanish. Specifically, the study asks the following research questions related to the frequency, the factors, and the universal trends regarding Subject Pronoun Expression in Equatorial Guinea:

- (a) What is the null and overt pronoun rate in Equatoguinean Spanish? How does this rate compare with pronoun rates in other varieties of Spanish?
- (b) What linguistic and social constraints favor the use of SPE in Equatoguinean Spanish?
- (c) Does SPE in Equatoguinean Spanish support or contradict universal trends of SPE in bilinguals, as proposed by the Interface Hypothesis?

(d) What does the use of SPE reveal about language contact between Spanish and the local languages that are spoken in Equatorial Guinea?

In the section that follows, a discussion of sociolinguistic variation in the speech of native, bilingual and second language speakers is provided to orient the reader.

Sociolinguistic Variation in the Speech of Native, Bilingual and Second language Speakers

Sociolinguistic variation is the study of how language varies and changes among a specific population of speakers. Extensive research conducted on first and second languages shows that even though linguistic structures provide more than one way of saying the same thing, language variation is indeed systematic.

Studies on sociolinguistic variation began with native speakers of English, with Labov (1963, 1972) groundbreaking research on Martha's Vineyard. Since then, several studies have been conducted to understand sociolinguistic variation in many other languages and these studies have been conducted in both monolingual and bilingual varieties. To analyze these studies, statistical programs that handle various variables, specifically multivariate regression analysis models such as Goldvarb, Rbrul, Varbrul among others are used. These programs have been fruitful in predicting the linguistic and extra-linguistic factors that influence language variation. In Spanish, Barrenechea and Alonso's (1977) study in Buenos Aires pioneered this trend. Subsequent studies were later conducted on the Spanish varieties in Spain, Latin America and the United States (e.g., Bentivoglio, 1987; Cameron, 1992, 1993; 1995; Orozco, 2015; Orozco & Guy, 2008).

Interestingly, although variationist studies originated with investigations of native speakers' speech, the past decade has seen studies being conducted on second language speech, also known as *Interlanguage* (Selinker, 1972), which will be described in the next section. An analysis of this kind not only sheds light on the underlying grammars of second language speakers but also on the extra-linguistic factors that come into play. According to Bayley and Tarone (2012, p.49), "the variationist approach to SLA is unique in its ability to offer powerful tools for the quantitative analysis of learner language in either cross-sectional or longitudinal studies..." Like native speakers, second language speaker's speech has also been found to vary systematically (Bayley, 1994; Geeslin, Linford & Fafulas, 2015; Geeslin & Gudmestad, 2008). It is for this reason that the present study investigates SPE in an L2-Only environment.

The next section discusses some universal tendencies in second language speech.

Universal Tendencies in Second Language (L2) Speech

Second language speech is often referred to as *Interlanguage*. This term was proposed by Selinker (1972) to describe the speech of second language speakers who possess features belonging to their native language, their second language and features not belonging to either the second or first language. The interlanguage can be seen as a set of evolving linguistics systems between the first and the second language. According to Selinker, interlanguage is systematic (it is governed by rules), it is dynamic, variable and it is a reduced system (it avoids complex systems).

Selinker (1972) categorized second language speech into five central processes. The first is *language transfer*. Language transfer occurs when learners transfer the rules

in their L1 to their L2. This transfer could be a *positive transfer* or a *negative transfer*. A positive transfer is when the L1 structure facilitates the acquisition of the L2 structure. For instance, a native speaker of Spanish being able to transfer the Spanish possession structure (e.g. *el amigo de mi hermana* “the friend of my sister”) to form the equivalent English possession structure. A negative transfer on the other hand is when the L1 structure impedes the acquisition of the L2 structure. For instance, the use of false cognates in Spanish such as *embarazada*, which means pregnant, to incorrectly mean one is embarrassed. The second process he mentions is the *transfer of training*. Selinker (1972, p.37) states that “if these fossilizable items, rules, and subsystems [which occur in IL performance] are a result of identifiable items in training procedures, then we are dealing with the process known as a transfer of training.” An example is a second language learner using the formal form *usted* (meaning *you*) when speaking to friends because the learner has been trained to use this form when speaking with the teacher in the formal classroom setting.

The third process is the strategies for second language learning. Strategies such as *repetition, deduction, inferencing, and simplification* are often employed by second-language speakers to reduce the cognitive load of having to remember two language systems. *Repetition* refers to the reproduction of a word after hearing it while with *deduction*, the second language speaker consciously applies certain rules to the second language. For instance, a Spanish learner deducing the conjugation of an unfamiliar verb ending in *ar* based on his/her knowledge of the rules regarding the conjugation of such verbs. With respect to *inferencing*, the second language learner makes guesses based on what the learner already knows, for instance, the meaning of new words. Regarding

simplification, the second language learner produces a simpler structure in the interlanguage than the structure found in the target language. For instance, a learner using the simple present (e.g. I am play) instead of a required present continuous tense (e.g. I am playing).

The fourth process mentioned by Selinker (1972) is the strategies of second language communication. For instance, in attempting to communicate with native speakers, second language speakers may employ certain strategies such as *language switching*, *circumlocution*, and *literal translations*. *Language switching* occurs when one switches between languages (e.g. “*gracias* for coming”) and *circumlocution* happens when one uses multiple words as a way of speaking indirectly (e.g. “he passed away from this earth”, instead of “he died”). Concerning *literal translations*, the learner translates word for word into the target language, irrespective of its meaning. For instance, translating the Spanish idiom “*te estoy tomando el pelo*” as “I am taking your hair” instead of the appropriate English equivalent “I am pulling your legs.”

The final strategy Selinker mentions is *an overgeneralization*. When using *overgeneralization*, learners apply certain rules related to the target language to contexts in which they do not apply. For instance, a Spanish learner wrongly conjugating an irregular verb such as *traer* ‘to bring’ as ‘*traí*’ instead of ‘*traje*.’

A controversial, yet often discussed, aspect relating to interlanguage is *fossilization*. According to Selinker (1972), *fossilization* occurs when advanced second language learners fail to progress to a natively like proficiency. This is due to the retainment of rules and sub-systems in their interlanguage. Long (2003) criticized the term

fossilization arguing that the claims were based on assumptions instead of evidence, inappropriate participants and insufficient data. He instead proposed that the use of the term *stabilization* as a more appropriate term to explain the errors of advanced second language learners. According to Long (2003), fossilization is a product whereas *stabilization* is a process of language learning that could be overcome as the learner progresses with the target language.

However, according to Long (2003), if the learner resorts to an avoidance of the target form, stabilization could be worsened and long-term stabilization may occur. For instance, adult second language learners may avoid certain target forms because of problems in mapping out certain target forms and their meanings, especially when there is more than one form that exists. The avoidance of the form could lead to long-term stabilization (Long, 2003).

The organization of the dissertation is described just below.

Organization of the Dissertation

The present study has seven chapters. The first chapter is a general introduction to the study. In this chapter, I have presented the phenomena being studied, important concepts related to language variation, as well as universal trends in second language speech.

In the second chapter, I offer a brief history of colonization and the European settlement in Equatorial Guinea, independence and the post-colonial era. In the same chapter, I also provide a linguistic overview of Equatorial Guinea by describing the

linguistic situation in Equatorial Guinea as well as the various native languages that are spoken.

In chapter three, the theoretical concepts relating to the study are explained in order to orient the reader. Chapter four presents a review of the literature on both monolingual and bilingual studies conducted on SPE. In chapter five, the methodology used in this study, as well as the linguistic and extra-linguistic factors to be examined, are discussed. Also, this chapter describes participants, the data collection method, the initial hypothesis and the method of analysis.

Chapter six provides the results of the study. After which the results are discussed in light of the research questions posed. Finally, in chapter seven, the contributions, limitations of this study, and suggested areas for future research are provided.

CHAPTER 2

A BRIEF OVERVIEW OF EQUATORIAL GUINEA

The goal of this chapter is to present a brief overview of Equatorial Guinea. Thus, the chapter is divided into four main sections. The first two sections focus on the history of Equatorial Guinea. In these sections, details regarding colonization and European settlement from 1471 to independence in 1968 are provided. There is also a description of the era of independence and post-colonization under the rule of Francisco Macías Nguema. The next two sections focus on the languages spoken in Equatorial Guinea. There is a focus on the local languages, the corresponding ethnic groups, and finally, a description of Spanish spoken in Equatorial Guinea.

Colonization and European Settlement

Geographically, Equatorial Guinea (E.G) is divided into two parts: the mainland region, which is bordered by Cameroon and Gabon, and the insular region, which contains Corisco, *Elobey Grande* and *Elobey Chico* islets, as well as the islands of Annobon and Bioko (formerly known as Fernando Poo). Bioko contains Malabo, which is the capital of Equatorial Guinea. Juan de Santarem and Pedro Escobar discovered the island of Annobon on January 1 in 1471. They named the island *do anno bon*, which means on the day of the New Year (Quilis & Casado-Fresnillo, 1995). Fernando Poo discovered the island of Bioko in 1474. He named the island *Formosa* which means beautiful. With time, this island took the name of its discoverer, Fernando Poo, before being changed to its present name, Bioko.

In 1778, Spain gained control over Fernando Poo, Annobon, and Corisco through the Treaty of Pardo and San Ildefonso they signed with Portugal. Spain's primary interest was in the slave trade and the establishment of ports between the Metropoli and Filipinas. Thus in 1798, the Spaniards conducted an expedition to Equatorial Guinea to manage the island. However, upon arrival in Malabo, they began to suffer from Malaria and other life-threatening diseases consequently leading them to abandon the island for many years (Quilis & Casado-Fresnillo, 1995).

Upon the abolition of slavery at the Vienna Congress in 1815, Spain allowed the British government to set up an anti-slavery base in Equatorial Guinea (Quilis & Casado-Fresnillo, 1995). Emancipated creole English-speaking slaves from Sierra Leone, Nigeria and Liberia arrived on the island of Fernando Poo. Their presence brought about the diffusion of Pidgin English, a creole of English based on the Krio language of Sierra Leone. The slaves who were referred to as *Fernandinos* became the bourgeoisie class in Equatorial Guinea. They not only engaged in commercial activities with the residents but also became intermediaries between the Europeans and the natives (Granda, 1990; Lipski, 1985; Quilis & Casado-Fresnillo, 1995). Moreover, it was a fernandino who introduced cacao, which later became a lucrative source of income for the country (Sundiata, 1990).

Interestingly, in spite of the early arrival of the Spaniards on the island, Spanish presence was not felt in Fernando Poo until the British were expelled in 1858. In the mainland, Spanish presence was only noticeable after the Spanish Civil War (1936-1939). Spain began economic development in Fernando Poo because of their interest in

the exportation of cocoa (Castillo Barril, 1964; Lipski, 1985; Quilis & Casado-Fresnillo, 1995) and they relied on the local population for cheap labor to work on the cacao plantations. A divide-and-rule technique was also used by Spain as a form of governance. The population were divided into two groups, the *emancipados*, who were citizens the Spaniards felt had assimilated to Spanish culture, and the *menores*, who were described as primitives or unassimilated citizens (Sundiata, 1990). While the latter were labor and plantation workers, the former were given administrative positions.

Moreover, the presence of the Catholic Church was highly felt during this period. Missionaries learned local languages to evangelize the people while the locals learned Spanish. This was because the ability to speak Spanish meant one could acquire property, share public spaces with the colonizers or occupy prestigious positions (Manso & Bibang Oyee, 2014).

The Spaniards also renamed many cities and establishments in Equatorial Guinea with Spanish names. For instance, Clarence City and West bay were renamed Santa Isabel and San Carlos respectively. Towns such as Niefang, Añisok, and Sampaka had their names modified; Sevilla de Niefang, Valladolid de Los Bimbiles, and Zaragoza respectively (Besari, 2005). In addition, Equatorial Guineans were forced to have Spanish Christian names placed before their traditional Bantu names. In the island of Annobon, not only were first names replaced but also native surnames were replaced with Spanish towns and cities (Manso & Bibang Oyee, 2014).

In 1968, amid pressures from grassroots movements as well as many international bodies, Spain gave in to nationalism. On October 12, 1968, Equatorial Guinea achieved its independence and elected its first president, Francisco Macías Nguema.

Independence and Post-Colonization.

Francisco Macías Nguema was president from 1968-1979 (Campos, 2003). His 11-year rule was marked by dictatorship and various human rights abuses. In July 1970, Nguema made Equatorial Guinea a single-party state with himself as a president for life. To get rid of all colonial influence, Nguema replaced Spanish as an official language with Fang, his maternal language. This resulted in a generation of Guineans with low proficiency levels and literacy skills in Spanish (Besari, 2005; Gomashie, 2019; Lipski, 2004; Nistal Rosique, 2007; Quilis & Casado-Fresnillo, 1995). With Fang as the official language, all other languages were relegated to rural areas or within homogeneous or interethnic communications (Granda, 1985). In effect, Fang became the language of administration and the only language used in public places.

Moreover, a significant portion of Guineans were exiled, all foreigners were expelled and pre-colonization infrastructure was destroyed (Lipski, 1985). Fernando Poo was renamed Macías Nguema. Macias also showed a complete disregard for academia and this was evident in his hunting down of all intellectuals, closing down of all libraries, censoring the press and journalists thus resulting in not a single university graduate in Equatorial Guinea after his rule in 1979 (Klinteberg, 1979). Furthermore, he committed several human rights violations such as forced labor on all citizens over the age of 15, as well as the murder and torture of all political opponents. In 1979, his nephew Teodoro

Obiang Nguema who was then the lieutenant colonel and minister of defense overthrew Macías. Macías was executed on September 29, 1979, by a firing squad. By the end of his rule, an estimated 50,000 to 100,000 of the country's population of 300,000 people were murdered (Sundiata, 1990). Besides, the Guinean economy had collapsed and one-third of all educated elite either had been exiled or murdered (Klinteberg, 1979).

Upon taking office, Teodoro Obiang Nguema restored Spanish once again as the official language of Equatorial Guinea. Equatorial Guinea also began bilateral relations with France in 1985 and became a member of Francophonie in 1989. French became elevated to the second official language of Equatorial Guinea and in 1997, it became a compulsory subject in schools. Portuguese was added as the third official language in 2010. Despite the country having three official languages (Spanish, French and Portuguese), most Guineans only speak their mother tongue and Spanish, which is the only language used in administration, education, religion and in the media.

Presently, the estimated population of E.G is 1,222,442 people: an average of 339,395 people in the insular region (made up of Little Elobey, Corisco, Bioko, Great Elobey, Annobón) and 882,747 the mainland (which is Rio Muni) ("Datos del Censo de Población de Guinea Ecuatorial" 2015). It is one of the least populated countries in Africa and concerning size, it is one of the smallest as well at 28, 051 sq. km. (the size of Maryland or Vermont in the United States). According to the United Nations Educational, Scientific and Cultural Organization (UNESCO), E.G had the highest literacy rate in sub-Saharan Africa (a literacy rate of 94%) as of 2014. Like most countries, four levels can be found in the educational sector of Equatorial Guinea-

preschool, elementary, secondary and higher education. Spanish is the only language used from elementary school through to secondary (Lipski, 1985).

Administratively, Equatorial Guinea is divided into the regions Bioko Norte, Bioko Sur, Litoral, Centro Sur, Kie-Ntem, Wele-Nzas and Annobon. Each region is subdivided into districts. Historically, Equatorial Guinea has depended on agriculture products for its survival-cocoa, coffee, and timber, which were exported to Spain, Germany, and the United Kingdom. Obiang Nguema's greatest achievement has been the discovery of oil and gas in 1991 (Kraus, 2010). With this discovery, Equatorial Guinea went from being an unknown country to one of the leading oil exporters in the world today (Frynas, 2004).

Linguistic Overview of Local Languages in Equatorial Guinea.

Despite Spanish being the official language in Equatorial Guinea, there is a wide variety of local languages spoken by each ethnic group as native languages (Lipski, 1985, 2004). These languages are Fang, Bube, the Playero languages, Fá d'Ambu (also known as Annobonese). These indigenous languages are used mostly in interethnic communications, as well as in family settings, radio, television programs, and political rallies (Lipski, 1985, Gomashie, 2019). Among these languages, Fang is the most widely used (Lipski, 1985, 2002, 2004; Simmons & Fennig, 2018). Fang speakers originally occupied the mainland region, Rio Muni. However, with time, they expanded their territory to the island of Annobon and Bioko as well (Lipski, 1984, 2004, 2008). Moreover, due to the implementation of Fang as a national language during the Macías

government, Fang is not just spoken by native speakers but by non-natives as well (Lipski, 2002).

Bube is spoken in the island of Bioko. According to Quilis and Casado Fresnillo (1995), Bube has about six unintelligible dialects. It was introduced into Equatorial Guinea before the 16th century after a series of immigration. The Playero speakers are found on the coast of Rio Muni and their languages include Kombe, Bujeba, Balengue, Batanga, and Benga (Lipski, 2004). According to Lipski (1985), although Fang, Bube and the Playero languages are Bantu languages they are mutually unintelligible (except for some Playero dialects).

Fá d'Ambu is a creole of Portuguese spoken in the island of Annobon. Its speakers were of Bantu origin and came from Angola, São Tomé and Príncipe (Quilis & Casado-Fresnillo, 1995). Pichinglis, also known as Pichi, is a creole of English originally brought to Equatorial Guinea by Nigerian contract workers. It is widely used on the island of Bioko (Quilis & Casado-Fresnillo, 1995). According to Lipski (2004, p. 117), Pichinglis is the lingua franca of Fernando Poo, even in spite of its stigmatized status and various campaigns to abolish its use (Lipski, 2000). It is used in informal settings, familiar and in-group conversations (Yakpo, 2009). After Fang, it is the second most widely spoken local language in Equatorial Guinea (Ethnologue, 2015).

A visual representation of the linguistic situation in Equatorial Guinea is provided in the map below.

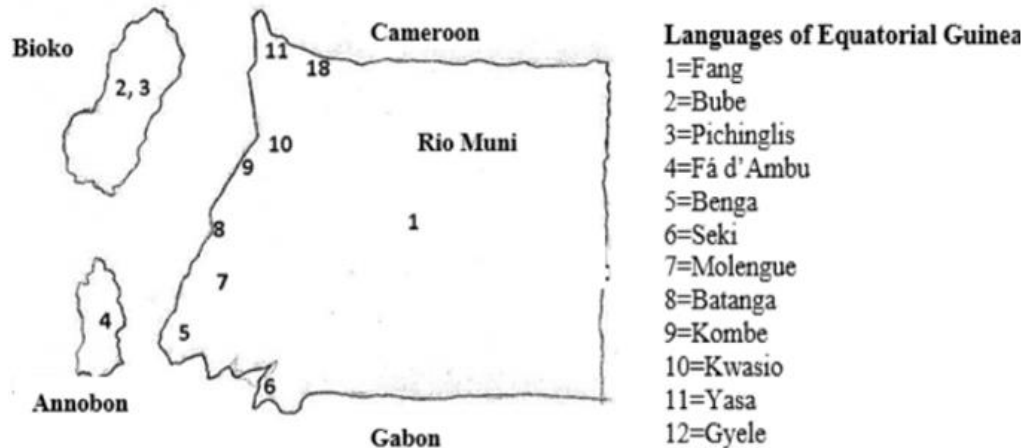


Fig 1. Map of local languages spoken in Equatorial Guinea.

In spite of this apparent multilingualism in Equatorial Guinea, many of the above listed local languages are at risk of endangerment (Gomashie, 2019; Manso & Bibang-Oyee, 2014). This is because they are neither taught in schools nor are they formally recognized as *lingua francas*. According to the Expanded Graded Intergenerational Disruption Scale (EGIDS) which measures language vitality, Fang and Pichininglis are described as widely used. Kwasio, Seki, Kombe, and Fà d' Ambu are described as active languages. However, Bube, Yasa, and Gyele are described as languages at risk of extinction while Molengue is described as nearly extinct. For these reasons, institutions such as *Instituto de Lingüística del Consejo de Investigaciones Científicas y Tecnológicas* (CICTE) (the Linguistic Institute of the Council of Scientific and Technological Research), the *Promotora Española de Lingüística* (PROEL) and the United Nations

International Children's Emergency Fund (UNICEF) are establishing language agencies in order to keep these local languages alive (Gomashie, 2019).

Spanish in Equatorial Guinea

Spanish has been the official language in Equatorial Guinea since 1844 (Besari, 2005). However, unlike many Spanish-speaking countries, it is not a native language in Equatorial Guinea. According to researchers, it can best be described as a second language (Nistal Rosique, 2007; Vuskovic, 2013), or a bilingual dialect (Bibang Oyee, 2002; Lipski, 1999, 2008; Quilis & Casado Fresnillo, 1995). It is both the medium of instruction and a compulsory subject in schools.

The earliest studies conducted on Equatoguinean Spanish focused on documenting the peculiar characteristics of Equatoguinean Spanish and how it differed from that of the Hispanic world (Bibang Oyee, 2002; Castillo Barril, 1969; Granda Gutiérrez, 1984, 1988; González Echegaray, 1951, 1959; Lipski, 1985; Quilis & Casado-Fresnillo, 1992). Granda (1984) described the frequent neutralization between /l/ and /r/ in Equatoguinean Spanish and attributed it to the interference from Fang. Lipski (1984) also described the main features of Equatoguinean Spanish which he listed as the variability of *tú* and *usted*, the combination of *usted* with the second person conjugations in Equatoguinean Spanish, the elision of /s/, the neutralization of rhotic sound (the tap /r/ and the trill /rr/), the use of the preposition *en* with motion verbs, inconsistent use of seseo and ceceo sounds, and the occlusive articulation of /b/, /d/ and /g/. These distinct characteristics have led some researchers (e.g. Lipski, 2008, 2014) to argue that in spite of the earlier characterizations of Equatoguinean Spanish as a deficient variety

characterized by various errors, Equatoguinean Spanish is best described as a dialect on its own. Other researchers disagree that there is an Equatoguinean dialect of Spanish (e.g., Bibang Oyee, 2002, Nsue Otong, 1986) and find Equatoguinean Spanish to be similar to that of other Spanish speaking countries.

Nonetheless, according to Lipski (2002), Equatorial Guineans are highly proficient Spanish speakers and Spanish is used in all parts of the country- on the island of Annobon (in spite of its isolation from the mainland), on the island of Bioko and in Rio Muni. However, Spanish language proficiency is not uniform among the Equatoguinean population in spite of the high percentage of Guineans that communicate in Spanish (Lipski, 2002; Besari, 2005; Nistal Rosique, 2007). Mohamadou (2008) describes three varieties of speech that can be found in Equatorial Guinea-the basilectal variety or *español popular*, mesolect variety or the *español común* and the acrolect or official Spanish. The uneducated, agricultural workers, laborers or anyone who needs the language to transmit a message speaks the basilectal variety of Spanish. This variety is learned outside the school and as such, speakers deviate from the rules of standard Spanish. The second variety is mesolectal variety. This type of Spanish adheres to the rules of standard Spanish much more than the aforementioned and it is spoken by administration workers. The third level is the acrolect or the official Spanish variety. The elite and intellectuals speak this type of Spanish. Based on these distinctions, Mohamadou (2008) states that in Equatorial Guinea, Spanish proficiency not only varies greatly but it is also in a continuum. It is affected by social variables such as one's level of education, profession, the relationship between speakers, and the context of the conversation.

Gomashie (2019) studied the vitality of Spanish in Equatorial Guinea and affirmed that Spanish is under no threat of extinction in Equatorial Guinea, as opposed to the other indigenous languages of the island. It is the most spoken language in Equatorial Guinea. It is used in administration, religion, education and mass media, and it enjoys prestige in Equatorial Guinea. It is a functional language in all sociocultural contexts. Moreover, various researchers have found that Equatorial Guineans have positive attitudes towards Spanish (Chirilă, 2015; Gomashie, 2019; Quilis, 1983, 1988). Chirilă (2015) for instance found that Equatoguinean students in Malabo had positive attitudes towards Spanish, especially among the younger generation. These Equatoguinean students were aware of the worldwide status of the language as well as the numerous benefits it provides them.

Summary

The goal of this chapter has been to provide a brief overview of Equatorial Guinea, from colonial rule to post-independence. Thus, the first section focused on providing a brief account of European settlement in Equatorial Guinea, beginning with the Portuguese discovery and ending with the settlement of the Spaniards. The next section focused on the post-colonial rule, specifically the reign of terror of Macías-the destruction of colonial infrastructure, numerous human rights abuses, the murder and torture of all opponents, among many others. Finally, the linguistic situation in Equatorial Guinea has been described. Specifically, a description of both the local languages spoken as well as Spanish in Equatorial Guinea. In the next chapter, the focus will be on the theoretical aspects of this study.

CHAPTER 3

THEORETICAL ASPECTS

This study combines the variationist sociolinguistics and language contact approach to examine the alternation of the null and overt subject pronoun in Equatoguinean Spanish. Thus, the chapter examines the theoretical aspects of the study. The chapter is divided into three sections. Firstly, there is an examination of the key tenets of the variationist approach in order to point out its suitability for the present study. In the second section, there is a discussion on the effects of language contact among bilinguals. Finally, because one of the goals of this study is to test the tenets of the *Interface Hypothesis* (Sorace, 2011), the third section examines this hypothesis in order to orient the reader.

The discussion of variationist sociolinguistics is as follows.

Variationist sociolinguistics

Language variation is at the core of sociolinguistics. Tagliamonte (2006, p. 4) described variationist sociolinguistics as “a discipline that integrates social and linguistic aspects of language.” Thus, it is the study of not only how language varies and changes, but also, the interaction of social factors (e.g., age, sex, education, etc.) and linguistic structures (e.g., grammatical forms, sounds, etc.). Because SPE is a variable phenomenon, the variationist approach is most suitable for its study.

A key concept in variationist sociolinguistics is the *linguistic variable*, the presumption that there are two or more ways of saying the same thing (Labov, 1972; Sankoff, 1980). To identify linguistic variables, Tagliamonte (2006) proposed a two-step process. First, one must identify two or more variant expressions of the form in question.

In the second step, one must provide an accountable method for narrowing down the variable in question and distinguishing it from other forms. Labov (1982) referred to the latter as the *principle of accountability*. Based on this principle, one must not only give a detailed account of occurrences of the variable in question but of non-occurrences as well.

Poplack (1993, p. 251) stated, “variation analysis combines techniques from linguistics, anthropology, and statistics to investigate language use and structure.” The central claim of variationist sociolinguistics proposed by Labov (1972) is that variation is not random but instead carefully structured. Labov (1982, p. 17) referred to this as “normal” heterogeneity. In other words, not only do speakers have more than one way of saying the same thing but also the heterogeneity in language follows a pattern, a pattern best observed when one critically analyzes the *vernacular* (Tagliamonte, 2006).

The *vernacular*, according to Labov (1972), is “the style in which the minimum attention is given to the monitoring of speech” (p. 208). Sociolinguistic interviews are used to elicit the vernacular in that they provide numerous benefits to the researcher. They differ from non-sociolinguistic interviews because they have the ability to elicit both the targeted forms of speech and a casual register. Moreover, sociolinguistic interviews are carefully designed to overcome the *observer's paradox*, a situation where people alter their speech when they are being observed. They also enable the researcher to collect a large quantity of data for coding and analysis. According to Tagliamonte (2006), in conducting sociolinguistic interviews it is important to begin with questions regarding demography and then gradually progress to more personal questions. In addition, it is important to take into consideration the age of the speaker and the type of

community in formulating questions for the interview. During the interview, the goal is to record one or two hours of speech and collect demographic data about each participant (Labov, 1973). The researcher thus asks participants questions that allow them to tell stories or vividly describe events while the interviews are recorded. These recordings are then transcribed later for analysis.

Although variationist sociolinguistics has its roots in dialectology, it goes beyond description and includes statistical methods in its analysis. These statistical methods allow one not only to find out the overall distribution, but also the correlation between linguistic and extra-linguistic forms. To find out the overall distribution, the occurrences of the variable are calculated and taken out of the total number of cases in which it could occur but it did not occur (Tagliamonte, 2006). Correlations, on the other hand, present the researcher with trends and patterns in the data set, which allow the researcher to make inferences regarding the variable in question. For instance, it allows the researcher to examine whether the variable in question is affected by linguistic factors such as *grammatical person and number, reflexivity, ambiguous verb forms, verb class, and reference* in the case of the present study. By way of illustration, *grammatical person and number, reflexivity, and ambiguous verb forms* are presented.

(1) *Vivo en Basupu y de Guinea Ecuatorial* (Speaker 4)

I live in Basupu and from Equatorial Guinea.

(2) *Para llegar temprano a la escuela, me despierto muy temprano en la madrugada.* (Speaker 6)

To get to school early, I wake up very early in the morning.

(3) *No volvería a llegar tarde más* (Speaker 6)

I would no longer arrive late.

Example (1) illustrates the use of the verb in the first person singular pronoun, (2) the verb in the reflexive form, and (3) the ambiguous verb form in the conditional tense.

Correlations also allow the researcher to determine whether the variable in question is affected by social factors such as *age, sex, style* and *register, ethnicity, culture, mass media*, among others. For instance, Labov (1972) found higher use of a *negative concord*, a situation where more than one form of negation is used in a sentence, among the working class subgroup than the middle class (example 4).

(4) *That ain't nothing new.*

Also, Sankoff (1974, p. 348), found the omission of *que* very widespread among speakers belonging to the working class (example 6).

(5) *C'est la fille ø j'ai vue*

This is the girl I saw.

As both linguistic and social factors are key factors to the understanding of language change, explanations for the results of this study will be provided in light of sociolinguistic issues as well as various external contexts (Tagliamonte, 2006).

Language Contact

Language contact, a common phenomenon in many speech communities occurs when speakers of two or more languages interact and their languages influence each other in a context where bilingualism takes place. Studies conducted on languages in contact thus examine the modifications or innovations that occur because of contact. In the ensuing paragraphs, there is an emphasis on important concepts in language contact

relevant to this study. Particularly, the outcomes of contact-induced language change, strategies employed by bilinguals in language contact situations, the question of linguistic versus social factors in the prediction of language change, and the question of whether language contact accelerates linguistic change as against the possibility of change occurring due to the internal characteristics of the language in question.

The general consensus in language contact studies is that intensive language contact leads to *contact-induced change*, “any linguistic change that would have been less likely to occur outside a particular contact situation” (Thomason, 2001, p. 62). The outcomes of this change have been traditionally divided into two main categories- *borrowing*, and *interference* from an L1 or other dominant languages (Winford, 2005). *Borrowing* is defined by Thomason and Kaufman (1988) as “the incorporation of foreign features into a group’s native language; although the native language is maintained, it is changed by the addition of the incorporated features” (p. 37). Lexical borrowing is the most common type of borrowing that occurs in language contact situations (Thomason & Kaufman, 1988). English, for instance, has incorporated a large number of borrowed words from French (e.g., garage, mayonnaise, chauffeur), and Arabic (e.g., algebra, alcohol, adobe) into its vocabulary.

Interference, on the other hand, can be defined as “errors in the learner’s use of a foreign language that can be traced back to the mother tongue” (Lott, 1983, p.256). Interference has been characterized by different labels in literature- *substratum influence*, and *transfer* in Second Language Acquisition (SLA) studies (Windford, 2005). *Substratum influence* refers to “a subtype of interference that results from imperfect group learning during a process of language shift” (Thomason & Kauffman, 1988, p. 38).

In the acquisition of the target language, non-native uses produced by the non-native speakers are then shifted to subsequent generations. This interference could be in terms of phonology, vocabulary, morphology, sounds, and syntax.

In contrast to interference, *transfer* broadly refers to the direct and indirect use of forms from the superordinate language, which is the dominant language used in a particular location. It occurs when bilinguals subject the secondary language to characteristics they find in the primary language (Silva-Corvalán, 1994). Thus, a transfer could be a *positive transfer*, when it leads to the rapid acquisition of the second language, or a *negative transfer* when it inhibits the acquisition of the second language.

Moreover, apart from borrowing and interference, contact-induced change could lead to changes in the minority language or the less dominant language, or the creation of new languages, such as *pidgins* and *creoles*. *Pidgins* are simplified language varieties that serve as *lingua franca*. For instance, many pidgins such as Nigerian Pidgin English, and Pichiinglis from Equatorial Guinea arose during the transatlantic slave trade and during colonization. *Creoles*, on the other hand, are developed out of pidgins and as such they have a much more developed and extensive grammar than that of pidgins. Furthermore, because creoles are passed down to children, these linguistic varieties have native speakers (Thomason, 2001). Some examples of creoles are the Louisiana Creole, and the Haitian Creole. Notably, both pidgins and creoles emerge as a result of fulfilling a communication need among speakers of unintelligible languages.

Furthermore, language contact often leads bilinguals to employ strategies to reduce the cognitive load of having to remember distinct language systems. Some of such bilingual strategies are the *simplification of grammatical categories*, *overgeneralization*

of forms, and *code-switching* (Silva-Corvalán, 1994). Gutierrez (1994) defines *simplification* as “a process whereby a form is expanded to a larger number of contexts” (p.112). Silva-Corvalán (1994) for instance found simplification of grammatical structures among English-Spanish bilinguals of East Los Angeles as certain grammatical structures were being lost (e.g., subjunctive, conditional, present perfect and past perfect forms). Similarly, Ocampo (1990) found an absence of the subjunctive in the Spanish spoken in Los Angeles. *Overgeneralization of forms* refers to the extension of the application of grammatical rules to contexts in which they are unconventional. For instance, the extension of *estar* to the context in which *ser* is needed in Spanish (as seen in the Spanish of Los Angeles) is an overgeneralization of *estar* (Silva-Corvalán, 1994).

Notably, *code-switching* is one of the most studied strategies employed by bilinguals. Poplack (2001, p. 2062) referred to it as “the mixing by bilinguals (multilinguals), of two or more languages in discourse, often with no change of interlocutor or discourse.” Gumperz (1982, p. 59) referred to it as “juxtaposition within the same speech, exchange of passages of speech belonging to two different grammatical systems or subsystems.” Three main types of code-switching have been distinguished in literature - *tag switching*, *inter-sentential switching* and *intra-sentential switching* (Poplack, 1980). *Tag switching* refers to the insertion of tags that belong to one language to another language. For instance, inserting English “*you know*” in a conversation in German. *Inter-sentential switching* is when one sentence is in one language and the subsequent sentence is in the other language (e.g. *I left home early. Estoy en la escuela ahora*) while *intra-sentential switching* is when one changes the language in the same sentence. For instance, “*Voy a trabajar because I need the money.*”

In addition to the strategies employed by bilinguals, an important question that is often invoked in the discussion of language contact is whether linguistic factors (such as the typology of the languages and universal constraints) or social factors (such as the extent of contact, sex, social status, education) best predict the outcomes in language contact situations. Thomason and Kaufman (1988) for instance argued that any change is possible with the right amount of social factors, and that linguistic interference is conditioned primarily by social factors and not by structural or linguistic constraints. Moreover, Thomason and Kaufman (1988) stated, “linguistic constraints on linguistic interference . . . are based ultimately on the premise that the structure of a language determines what can happen to it as a result of outside influence. And they all fail” (pp.13-14). Consequently, social constraints affect the direction of the interference, the extent and the type of features that are transferred from one language to the other. In effect, there is no linguistic constraint on the number of features that can be transferred from one language to the other according to Thomason and Kaufman (1988).

Other scholars such as Silva-Corvalán (1994), argue that even though social factors have a role to play, “the structure of the languages involved, to a large extent, constrained by cognitive and interactional processes, governs the introduction and diffusion of innovative elements in the linguistic systems” (p. 6). Consequently, the minority language is revitalized or changed, and the changes attested are constrained by the structure of the minority language. Notwithstanding, as the question of the role of social versus linguistic factors remains an issue, other researchers such as Sankoff (2002) have argued that there is a complex interplay between both factors. Thus, both linguistic

and social factors are of equal relevance in the prediction of the outcomes of language contact.

Moreover, related to the question of the relative strength of linguistic versus social factors is the question of whether the changes that take place in a language are accelerated by language contact vs. the possibility of an autonomous development/internal characteristics that takes place in all languages. Silva-Corvalán (1994) in particular, argues that many linguistic processes, such as simplification, overgeneralization, transfer, analysis, and convergence, often attributed to bilingualism/multilingualism are in fact also found in other monolingual non-contact varieties. For instance, she found that tenses that were lost in the speech of the United States born bilinguals were undergoing simplification in the speech of first-generation immigrants (immigrants who were born in one country but now reside in another), even though both groups had high levels of proficiency. Nonetheless, because of the difficulty in determining whether a change would have occurred with or without the contact situation, various researchers have used social variables to determine whether this change would or would not have occurred. For instance, Otheguy et al. (2007) examined the variable “recency of arrival” of participants into the contact situation and determined that English did have a role to play with respect to high SPE rates found in Spanish in New York (Otheguy et al., 2007).

In sum, language contact can potentially lead to various modifications in the languages that have come into contact. In this section, the focus has been on highlighting some of these changes-the outcomes of contact-induced change, strategies employed by bilinguals in language contact situations, the question of linguistic versus social factors in

the prediction of language change, and the question of whether language contact accelerates linguistic change vs. the possibility of change occurring due to the internal characteristics of the language in question. In the next section, the focus will be on universal and language-internal accounts for non-native linguistic features in bilinguals.

Interface Hypothesis

Sorace (2011) defined an interface as “syntactic structures that are sensitive to conditions of varying nature” (p.6). The Interface Hypothesis (Sorace, 2006; Sorace, 2011; Sorace & Serratrice, 2009) was thus proposed to account for some of the non-target like behavior found in the speech of adult second language learners, with particular emphasis on near-native grammars. White (2007) expanded the Interface Hypothesis (IH from here on) by distinguishing between internal interfaces and external interfaces. Internal interfaces refer to formal features such as interfaces between language and semantics or phonology and morphology whereas external interfaces refer to interfaces involving cognitive domains such as syntax-discourse or semantics and pragmatics. External interfaces pose the greatest instability and present a greater challenge during acquisition than internal interfaces (White, 2007). Taking into account these interfaces, Sorace and Filaci (2006) argued that “narrow syntactic properties are completely acquirable in a second language, even though they may exhibit significant developmental delays, whereas interface properties involving syntax and another cognitive domain may not be fully acquirable” (p. 340).

In light of these differences, various accounts have been proposed to explain the differences between L2 and native speaker end states. These accounts can be grouped into two main categories-the *representational account* and the *processing resources*

account. With respect to the *representational account*, Sorace (2011) argued, “there are differences between bilinguals and monolinguals at the level of knowledge representations ... because one of the grammatical systems affects the other” (p. 13). Particularly, some aspects of L1 universal grammar are no longer accessible to adult learners, thus leading such learners to have a narrower syntactical repertoire (Sorace, 2011; Tsimpli & Dimitrakopoulou, 2007; Tsimpli & Mastropavlou, 2008). Furthermore, there is a cross-linguistic influence at the level of representation resulting in external features, in the syntax-discourse interface becoming underspecified. Sorace (2006) argued that this underspecification could be explained by *residual optionality*, “unsystematic L1 effects surfacing in the L2 speaker’s production” (pp. 111-112).

The *processing account*, on the other hand, argues that the differences between L2 and native speakers have to do with the differences in processing strategies employed in real-time (Sorace, 2011). Particularly, at the syntax-pragmatics interface, the processing is less automatic due to less advanced knowledge representations (Sorace, 2011). Thus, in order for learners to be able to navigate through interfaces, they not only need an adequate grammatical representation but also an advanced processing ability that gives access to extra-linguistic content.

Evidence to support these assertions comes from initial language acquisition studies that tested the IH in near-native speakers and advanced L2 learners. A well-studied phenomenon in this regard is the distribution of pronominal subjects. This distribution distinguishes between languages that allow for null subjects (e.g., Spanish and Italian) and languages that do not allow for null subjects (e.g., English and French). In null subject languages, for instance, the presence or absence of the pronoun is

grammatically correct. Nonetheless, the decision to omit or include the pronoun is based on the interaction between syntax and discourse/pragmatics. Particularly, null subjects occur when there is no change in topic whereas overt subjects occur with a change in topic.

Studies on pronominal subject expression have found that bilingual children, near-native speakers, and speakers undergoing attrition fail to limit overt subjects to contexts where there is a change in discourse topic. For instance, Belletti, Bennati, and Sorace (2007) investigated native English speakers acquiring Italian as an L2 and found developmental delays both with respect to the use of pronominal overt subjects, and postverbal subjects in contexts of focus. Particularly, although the speakers exhibited a native-like or target-like syntactic proficiency with respect to the interpretation of null subjects, they exhibited non-target-like use of postverbal subjects as well as an overuse of overt pronominal subjects. Similar results were found in Tsimpli and Sorace's (2006) study of native Russian speakers acquiring Greek as an L2, where the researchers found the overuse of null subjects by advanced L2 speakers. Sorace and Serratrice (2009) argued that the difficulties attested may be due to "the integration of syntactic and pragmatic information involved in the selection of subject pronouns" (p. 967).

In spite of the solid evidence in support of the IH, recent studies have not been conclusive and as such, have posed challenges to its predictions. For instance, Cerrón-Palomino (2018) found that, contrary to the tenets of the IH, Quechua Spanish bilinguals were able to navigate between the syntax-discourse domains in that there were similar SPE rates between the Quechua Spanish bilinguals (17.1%) and the Spanish monolingual group (15.1%) in his study. Essah's (forthcoming) study on subject pronoun expression

in Equatorial Guinea also found bilinguals having a low SPE rate, which contradicts Sorace (2004) and Sorace and Serratrice (2009) 's prediction of overuse of overt subject pronouns. Similarly, Ivanov's (2009) study of object clitics in L2 Bulgarian and Rothman's (2008) study on subject pronouns in Spanish found that these features can be acquired in a native manner in spite of the prediction of the IH.

In sum, the IH has produced mixed results among various researchers. Linguistic interfaces have not been homogenous as initially proposed by the IH. White (2011, p. 587) argued that "it is not the case that all interfaces lead to difficulties, it is not the case that all phenomena at a particular interface are necessarily problematic, and it is not the case that acquisition failure is inevitable." Nonetheless, the main conclusion that can be drawn from investigations on the IH is that interfaces are unequally vulnerable. Properties at the syntax-discourse/pragmatics interface generally are more difficult to acquire, while properties at the grammar/internal pose fewer difficulties to L2 learners and near-native speakers. Thus, it is categorical that interfaces at the syntax-discourse/pragmatic interface cause problems with ultimate attainment.

Summary

The goal of this chapter has been to discuss the theoretical aspects of this study. Thus, the chapter began with an explanation of the variationist sociolinguistics approach in order to show its suitability to the present study. Next, there was an examination of the effects of language contact and the resulting contact-induced change. Finally, the last section examined the IH, since the objective of the study is to test its tenets as well. The next chapter is dedicated to examining previous studies conducted on SPE.

CHAPTER 4

LITERATURE REVIEW

Linguists have studied SPE in Spanish from three main angles: (a) factors that condition the alternation of the expressed and unexpressed subject pronoun, (b) the rates of overt pronoun use where variation with a null subject is possible and (c) the increased/decreased use of overt subject personal pronouns in some varieties of Spanish. Nonetheless, in spite of the numerous studies conducted on SPE in both monolingual and bilingual varieties, of relevance to this literature review are variationist studies conducted in Spanish that have employed a statistical regression analysis. These studies have been conducted using sociolinguistic interviews as data sources and have examined subject expression as a function of various language-internal and external variables.

To date, the highest SPE rates have been reported in Cameron's (1994) study of Puerto Rican Spanish (a pronominal rate of 45%), while the lowest rates (16.2%) have been reported in Cerrón-Palomino's (2018) study of Huancayo in Peru. A robust pattern found across studies is that higher overt pronoun rates are used in the Caribbean than in the rest of the Hispanic world (e.g., Barrenechea & Alonso, 1977; Bayley & Pease-Alvarez, 1997; Bentivoglio, 1987; Cameron, 1993; Erker & Guy, 2012; Flores-Ferrán, 2004; Morales, 1980; Orozco & Guy, 2008; Otheguy & Zentella, 2007, 2012; Otheguy, Zentella & Livert, 2007; Posio, 2011; Prada Pérez, 2009; Shin & Otheguy, 2013; Silva-Corvalán, 1982, 1994). With respect to both the order of constraint rankings as well as the set of factor groups constraining SPE, regional distributions have been attested. In addition, the increased/decreased overt pronoun use in some varieties has been linked to language contact thus leading researchers to debate whether bilinguals develop two

languages autonomously or there is cross-linguistic influence. In particular, the increase overt SPE in Spanish in contact with English has been linked to the obligatory use of the subject pronoun in English (e.g., Abreu, 2012; Otheguy & Zentella, 2012; Otheguy, Zentella & Livert, 2007; Shin & Otheguy, 2009; Silva-Corvalán, 2014).

The section below is a review of studies conducted on monolingual and bilingual varieties of Spanish.

SPE in Monolingual Varieties of Spanish.

There exists a considerable body of literature on SPE in monolingual varieties of Spanish, particularly, the regional varieties of Latin America and Peninsular Spanish (Carvalho, Orozco, & Shin, 2015). Barrenechea and Alonso (1977), and Morales (1980) in Buenos Aires and Puerto Rico respectively, conducted the earliest studies on subject pronoun expression in Spanish. These were groundbreaking studies that led the way for studies on Latin American Spanish (e.g., Bentivoglio, 1987; Cameron, 1992; Orozco, 2015; Orozco & Guy, 2008; among others), Peninsular Spanish (e.g., Cameron, 1993, 1995; Posio, 2011; among others) and Spanish in the United States (e.g., Otheguy & Zentella, 2007, 2012; Shin & Otheguy, 2013; Silva-Corvalán, 1982, 1994).

Monolingual studies on SPE have well acknowledged the importance of linguistic factors (specifically *grammatical person and number, switch-reference, priming, Tense Mood Aspect [TMA], lexical semantics of verbs, clause types, and reflexivity*) and extra-linguistic factors (*age, sex, genre, and occupation*) in conditioning SPE. With respect to *grammatical person*, most studies have either found the first person *yo* (e.g. Orozco, 2015) or the second person *tú* (e.g. Alfaraz, 2015) as strong predictors of SPE. Regarding

number, singular pronouns favor the expression of overt subject pronouns whereas plural forms disfavor it (e.g., Alfaraz, 2015; Carvalho, Orozco, & Shin, 2015; Lastra & Butragueño, 2015; Orozco, 2015; among others). Orozco (2015) suggests that the disfavoring effect of plural pronouns could be due to the use of fixed expressions such as *vamos* ‘let’s go’ and *nos vemos* ‘see you’ with null subjects. Cameron (1992) on the other hand, suggests that it could be due to the fact that many plural subjects are used when a human referent has already been introduced and thus the omission of plural overt subject pronouns.

For instance, Orozco’s (2015) study on Colombian Costeño Spanish found 44.5% of overt subjects in the first person singular *yo*, 39.6% in the third person singular *él/ella*, and 32.5% in the second person singular *tú*. Along the same lines, Lastra and Butragueño’s (2015) study of Mexican Spanish found 24.7% of overt subjects in the first person singular, 27.1% in the third person singular and 16.6% in the second person singular. In Cerrón-Palomino’s (2018) study of SPE in Andean Spanish, he found 39.6% of overt subject pronouns (SPPs) in the second person singular *tú* followed by the first person singular *yo* (19.9%) and the third person singular *él/ella* (17.7%). In addition, an interaction has been found between pronominal forms and specificity when linguistic factors are examined in terms of the different pronominal forms in a factor group (instead of grammatical person). Particularly, several studies have found that the specific *tú* and *ella* (‘specific’ meaning when *tú* and *ella* refer to the interlocutor) have a strong likelihood of being an overt SPP (e.g., Alba, 2004; Alfaraz, 2015; Martinez-Sanz, 2011). By way of illustration, the sentences, *tú me haces feliz* ‘You make me happy’ and *¿Cómo llegas a la estación?* ‘How do you get to the station?’ are compared. Whereas the first

sentence refers to the interlocutor, the second is a general statement that can refer to anyone.

A change in the referent of the second of two consecutive verbs has also been a strong predictor of SPPs (e.g., Ávila-Jiménez, 1995; Bentivoglio, 1987; Cameron, 1992, 1993; Flores-Ferrán, 2004; Holmquist, 2012; Lastra & Butragueño, 2015; Morales 1980, 1989; Orozco, 2015). Alfaraz (2015), for instance, found overt pronouns occurring 59.2% of the time when there was a change in reference, whereas in same reference contexts the pronoun was overt 40.8% of the time. Alfaraz (2015) also found that overt pronouns were favored by the second person specific in a switch reference context and second person non-specific in same reference contexts. Furthermore, overt forms were disfavored by the third person non-specific in both same and switch reference. Alfaraz (2015) concluded that there is a weakening of constraints in same referent contexts. These results are in line with Orozco's (2015) study of Colombian Costeño Spanish, where a complete change in reference had the highest factor weight, followed by a partial change in reference and finally the same reference.

Concerning *reflexivity*, previous research has emphasized that overt pronouns are less likely to be expressed with a reflexive pronoun than with a non-reflexive pronoun (e.g., Bayley & Péase-Álvarez, 1997; Carvalho, Orozco, & Shin, 2015; Michnowicz, 2015; Otheguy & Zentella, 2012; Shin, 2014). Abreu (2009), for instance, found that each time there was a reflexive verb, overt SPP's were disfavored and slightly favored when the verb was not reflexive. Bayley and Péase-Álvarez (1997) explain that the tendency to have the overt pronoun omitted with the occurrence of a reflexive verb is due to the

already identifying information reflexive verbs possess, thus reducing the likelihood of ambiguity.

Another variable that has been found to favor overt SPE is *verb type* (Orozco, 2015). Interestingly, its classification in SPE studies has not been uniform. Different studies have used different classifications in this respect. For instance, Travis (2005, 2007) divided verbs into verb classes and proposed the following categories; *psychological verbs, speech acts, copula, motion, and other* (with *other* referring to any verb that did not fall within the other four categories). This division was used in other studies such as Cerrón-Palomino (2014), Manjón-Cabeza Cruz, Pose Furest, and Sánchez García (2016). Similarly, Otheguy and Zentella (2012) proposed a division of verbs into four categories of lexical content; mental activity, estimative, external activity and stative verbs. Nonetheless, in spite of the differences in classification, there is a unanimous agreement that cognitive psychological verbs and mental activity verbs (e.g. *pensar* ‘to think’) promote the use of overt pronouns more than other types of verbs (Carvalho, Orozco, & Shin, 2015). By the same token, speech acts (e.g. *decir* ‘to say’), and states (e.g. *estar* ‘to be’) favor overt pronouns whereas external activity verbs disfavor overt pronouns (e.g., Abreu, 2009, 2012; Bentivoglio, 1987; Carvalho, Orozco, & Shin, 2015; Enríquez, 1984; Posio, 2011; Orozco, 2015; Orozco & Guy, 2008; Travis, 2007; among others).

Regarding TMA, previous studies have found uniformity of effects across monolingual speech communities. Ambiguous verb morphology (e.g. when there is great overlap among singular persons *yo* and *él/ella/usted*. in the imperfect forms [e.g. *tenía*], conditional [e.g. *tendría*], and subjunctive [*tenga*]), promote overt pronoun use in contrast

with unambiguous verb forms (when each verb form has a unique ending, e.g., future [tendré/tendrá], indicative [tengo/tiene], and preterit perfect [he/ha tenido]) (Carvalho, Orozco, & Shin, 2015). Specifically, first and third person forms have been linked to the presence of overt SPPs due to the increase of ambiguity in these forms (e.g., Abreu, 2009, 2012; Bentivoglio, 1987; Cameron, 1993, 1994; Claes, 2011; Cerrón-Palomino, 2014; Orozco, 2015; Prada Pérez, 2009; Travis, 2007). Silva-Corvalán (2001) provides an explanation for this tendency. She suggests that it is the attention focusing property of discourse that promotes the use of overt SPP's and not necessarily ambiguity of the verbs in the first and third person. She argues that because verbs in the imperfect forms, conditional and subjunctive usually describe the setting of events, there is more attention on the subject, thereby increasing the tendency to use pronouns that are more overt. On the other hand, verbs in the preterit, future and indicative usually narrate events and as such focus on the predicate.

Apart from grammatical person and number, reference, reflexivity, and TMA, *priming* is another factor often cited in literature to be significant in monolingual varieties of Spanish. In other words, the expression of an overt pronoun has been found to trigger the expression of a subsequent overt pronoun (e.g., Cameron, 1995; Cameron & Flores-Ferrán, 2004; Travis, 2005, 2007; Orozco, 2015). Travis (2007) in particular investigated structural priming in the first-person singular subjects. She specifically looked at New Mexican narratives and Colombian Spanish conversations. The results from an analysis of 2000 verbs occurring with first person singular subjects found that in both data sets, priming was a significant predictor of SPE; a preceding co-referential unexpressed

subject led to a subsequent unexpressed subject and a preceding co-referential expressed subject led to a subsequently expressed subject.

Travis (2007) however made two interesting observations. Firstly, she mentions that although one was likely to assume the New Mexican data would have a higher number of expressed pronouns due to contact with English, the first person singular was more likely to be expressed in the Colombian data than the New Mexican data. She, however, notes that these differences appeared to be related to the type of genre (narratives versus conversations) in which they were contained, thus bringing to light the need to investigate different types of speech. Secondly, the effect of priming was found to last longer in the New Mexican data set (up to ten intervening clauses) than in the Colombian data set. Similar results were found in Travis (2005) and Orozco's (2015) study on Colombian Spanish. Travis (2005) for instance focused on the first person singular subject expression and found that it also underwent a priming effect. However, priming was only statistically significant at a distance of one and two clauses. Orozco (2015) on the other hand, focused on subjects preceding an overt SPP, a noun phrase and a null subject. He found that preceding overt SPP promoted the occurrence of overt pronouns, followed by a preceding noun phrase which had a neutral effect, while a preceding null subject favored the occurrence of another null subject.

In addition to the above-mentioned linguistic factors, a number of studies have shown that SPE rates are conditioned by *clause type*, with an overt pronoun more likely to occur in the main clause than in the subordinate clause (e.g., Abreu, 2009; Carvalho, Orozco, & Shin, 2015; Enriquez, 1984; Orozco & Guy, 2008). Abreu (2009) for instance found that when there was more distance between subjects and their referents, SPPs that

are more overt were produced. For instance, in her study, when there was no intervening clause, the overt pronominal rate was 29%. This increased to 45% when there were one to four intervening clauses. When there were five to ten intervening clauses, the overt SPE shot up to 56%, thus providing evidence in support of distance from the previous reference as being an important factor in determining SPE.

Unlike the aforementioned linguistic factors, to date, there is no consensus regarding the effect of social factors on monolingual varieties of Spanish. Some studies have found social factors do not condition SPE in monolingual varieties (e.g., Bayley & Pease-Alvarez, 1997; Bentivoglio, 1987; Cameron, 1992, 1994; Flores-Ferrán, 2002; Martínez-Sanz, 2011) while others have found a significant effect of social factors (e.g., Ávila-Jiménez, 1995; Alfaraz, 2015; Bayley & Pease Álvarez, 1996; Cameron, 1992; Claes, 2011; Enríquez, 1984; Lastra & Martín Butragueño, 2015; Miró Vera, & De Pineda, 1982; Orozco, 2015). The social factors that have been found to be significant in monolingual studies shall be discussed in the ensuing paragraphs; *sex*, *age*, *genre*, and *occupation*.

Many scholars have investigated the role of *sex* and SPE rates. Some studies have found *sex* to be significant (e.g., Alfaraz, 2015; Bayley & Pease Álvarez, 1996; Cameron, 1992; Claes, 2011; Enríquez, 1984; Miró Vera, & De Pineda, 1982) whereas others have not (e.g., Cerrón-Palomino, 2014; Lastra & Butragueño, 2015). For instance, Alfaraz (2015) found females slightly favoring overt SPP and males favoring null subjects. Also, Orozco (2015) found the combined effect of a speaker's *age* and *sex* significant in his study of Costeño Spanish. The results of this study showed women born before 1960 favoring overt pronouns whereas men born after 1960 favored null subjects. Interestingly,

women born after 1960 and men born before 1960 had a neutral effect. Orozco (2015) mentions that other studies (e.g., Orozco, 2007, 2009, 2010) have found distinct sociolinguistic behaviors among speakers around the age of 40. Likewise, in Orozco's (2015) study on Jalapa, Mexico, sex did significantly constrain SPE. His study differs from other SPE studies conducted in Mexico where sex was not significant. For instance, Lastra and Martín Butragueño (2015) of Mexico City and Michnowicz (2015) study of Yucatan Spanish.

In addition to sex, age is another social variable found significant in some monolingual studies (e.g., Carvalho & Bessett, 2015; Lastra & Martín Butragueño, 2015; Orozco, 2015; Orozco & Guy, 2008) and insignificant in others (e.g., Alfaraz, 2015; Bentivoglio, 1987; Michnowicz, 2015). In the past decade, studies that have found age to condition SPE have found younger speakers to favor consistently linguistic innovations more than older speakers (Labov 2001). For instance, Morales (1986) found that the 16 to 50 age group used significantly more pronouns (58% null pronouns and 42% overt), whereas the 50 plus age group used more null pronouns (72% null and 28% overt). Moreover, Ávila-Jiménez's (1995) study of Puerto Rican Spanish found significant differences among the three age groups; Group A (20-49), Group B (50+) and Group C (10-19) ($p > 0.001$). Both Morales (1986) and Ávila-Jiménez (1995) found significant differences between the age groups under 50 and those over 50. Moreover, there was a statistically significant difference found between the results of Group C ($p < 0.001$), Group A ($p < 0.001$), and Group B ($p < 0.005$). However, contrary to the above-mentioned studies, Lastra and Martín Butragueño's (2015) study of Mexican Spanish found SPE decreasing with age with both frequencies as well as factor weights. In

particular, older speakers (over 55 years) favored overt SPPs, whereas the youngest speakers (20-34 years) favored null subjects. This led the researchers to propose a *floor effect* for subject pronoun expression taking place. They explain that the lowest rates found among the younger generation are a representation of the lowest amount of SPP that adults use in Spanish. The researchers call for the need for additional studies in order to investigate if this was a retrograde change in progress or an instance of age grading.

Notably, very few studies have investigated *genre* and SPE rates in monolingual varieties (e.g., Ávila-Jiménez, 1996; Lastra & Martín Butragueño, 2015; Montes Miró, 1986). Nonetheless, Lastra and Martín Butragueño (2015) found genre as a significant conditioner of SPE in Mexican Spanish. In this study, argumentation favored SPE most strongly while description slightly favored overt SPE. According to the researchers, this tendency could be attributed to the need to establish one's opinion or that of the other when there are divergent opinions. However, dialogic and narrative genres disfavored overt SPP. Similar results were found in Solomon's (1988) study that found conflict narratives narrated by the speaker to have more overt SPE rates than non-conflict narratives. Also, Orozco (2015) found that one-on-one conversations favored overt subjects while the involvement of other subjects of the consultant's social network favored null subjects. Orozco and Guy (2008) explain that the difference in SPE rates could be due to the fact that the presence of other subjects leads to a greater tendency to modify one's speech. Moreover, the need to establish intimacy among speakers during the conversation could be another reason for the high SPE rates.

Finally, very few studies have found *occupation* to significantly condition SPE thus pointing out the need for further research in this area. Ávila-Jiménez's (1996) study

on Puerto Rican speakers divided speakers into two groups; occupational level A comprised of professionals and unknown categories, while occupational level B was comprised of managerial, clerical, skilled and unskilled workers. Each participant took part in interviews with the researcher where they answered questions ranging from academic to personal issues. Both groups differed significantly with respect to SPE rates in that speakers from occupational level A favored more overt pronouns than occupational level B (43% versus 37%). Perhaps the higher overt pronoun rate with occupational level A could be due to them as professionals trying to separate themselves from other groups. Interestingly, when occupational level A and B were cross-tabulated with grammatical person and number, Ávila-Jiménez (1996) found that they had similar rates of overt and null subject pronouns in all persons except the third person singular; level B had significantly more pronouns for *él*, *ella* and *usted* while level A had a more balanced rate between null and overt pronouns. Nonetheless, overall, Ávila-Jiménez (1996) found that the rates of individual grammatical persons by occupation revealed similar patterns of subject pronouns use.

From the above review, it is obvious that over time, an extensive literature has developed on subject pronoun expression in monolingual varieties of Spanish. Most of these studies have focused on explaining what conditions the expression or omission of the subject pronoun. Moreover, using statistical regression analysis, the vast majority of studies place linguistic factors as key determinants of subject pronoun expression in monolingual varieties of Spanish. The linguistic factors that have received the most attention have been grammatical factors, co-reference, verb type, TMA, priming, and clause type. Although there is no consensus regarding the importance of social factors in

monolingual varieties of Spanish, the most commonly studied social variables have been age, sex, occupation, and genre. To sum up this section, four key findings emerge from monolingual varieties of Spanish: (a) Variable linguistic behavior is highly systematic, (b) Studies conducted on Caribbean varieties have higher SPE rates than mainland varieties, (c) Grammatical factors such as co-reference, verb type, reflexivity, TMA, priming, and clause type are significant when included in most studies, and (d) Social factors have had limited relevance.

The following section summarizes SPE in bilingual varieties of Spanish. It examines SPE studies conducted on Spanish in contact with English as well as Spanish in contact with other romance languages.

SPE in Bilingual Varieties of Spanish

Research conducted on bilingual varieties of Spanish has generally explored SPE in contact with English, indigenous languages, romance languages, and creoles. The aim of these studies has been to investigate not only the linguistic and extra-linguistic factors that constrain overt subject pronouns but also to examine whether there is contact-induced change. In doing so, researchers have investigated SPE when Spanish is contact with similar null subject languages such as Portuguese, Catalan, Italian, as well as non-null subject languages such as English and French.

SPE in Spanish in contact with English has been widely studied in the United States. Some studies have compared monolingual varieties to bilingual varieties (e.g. Abreu, 2009), different immigrant groups (e.g. Otheguy, Zentella & Livert, 2007), Spanish dominant speakers, English dominant speakers and balanced bilingual speakers

(e.g. Cerrón-Palomino, 2016), and different generations of bilingual speakers (e.g. Silva-Corvalán, 1994). Nonetheless, to date, there is no consensus regarding whether or not contact with English has an effect on the production of overt SPP. On one hand, some studies have found contact with English leading to higher overt SPPs (e.g., Abreu, 2012; Bayley & Pease-Álvarez, 1997; Lapidus & Otheguy, 2005; Michnowicz, 2015; Otheguy, Zentella, & Livert 2007; Otheguy & Zentella, 2012; Shin, 2014; Shin & Otheguy, 2009; Silva-Corvalán, 2014), while on the other hand, there are studies that have found contact with English having no differential effect (e.g., Bayley-Cardenas, Treviño Schouten & Velas Salas, 2012; Besset, 2018; Cerrón-Palomino, 2016; Flores-Ferrán, 2004; Silva-Corvalán, 1994; Torres Cacoullous & Travis, 2010, 2011; Travis, 2005, 2007).

Studies that have found contact with English leading to high SPE rates in Spanish have attributed this to the obligatory use of the subject pronoun in English being transferred to Spanish. For instance, Otheguy and Zentella (2012) found that overt SPPs were especially elevated among speakers born in New York or arrived in New York at the age of three, immigrated to the city at an early age, or had lived in New York City for a long time. These speakers had an overt pronoun rate of 38%, whereas the Latin American raised had a rate of 33%. Otheguy and Zentella's (2012) results are also in line with that of Abreu (2012) where bilinguals of English and Spanish were found to have higher overt SPPs than their monolingual counterparts (49% versus 38%, respectively). Likewise, Shin (2014) examined the impact of English on the Spanish of New York by studying divergent contexts where pronoun omission was common in both Spanish and English and convergent contexts where omission was common in both Spanish and English. After analyzing 25,000 verbs, the results indicated that English not only acted as

a promoter of pronoun use in Spanish, but it also inhibited pronoun use in contexts where both languages tended to omit pronouns. Consequently, participants who were raised in New York used more pronouns than recently arrived speakers (37.6% versus 28.8%, respectively).

Studies that do not lend support to the English influence have found low pronoun rates in Spanish varieties in contact with English. Besset (2018) found a pronominal rate of 16.7% and 19.3% respectively in his comparison of SPE in monolinguals from Sonora versus bilinguals from Arizona, while Cerrón-Palomino (2016) found a low pronominal rate of 17.8% in Spanish spoken in Arizona. In Cerrón-Palomino's (2016) study, participants took part in interviews where they narrated events regarding their childhood, schooling, and leisure activities, among others. The results of this study showed that participants with the most contact with English produced the least amount of subject pronouns. By way of illustration, Spanish dominant speakers produced an overt pronominal rate of 20.7%, English dominant speakers produced an overt rate of 14.8% and balanced bilinguals produced an overt rate of 16.2%. These results are in line with Silva-Corvalán (1994) study on Mexicans living in East Los Angeles where the first generation (which was more Spanish dominant) produced more overt SPP than the second and third-generation (33%, 28%, and 26%, respectively) which were more English dominant.

Some explanations have provided an account for the discrepancy in results among researchers who find an English influence and those that do not. On one hand, Otheguy, Zentella, and Livert (2007) suggest the different results found among studies

investigating Spanish in contact with English could be attributed to the sample size as well as the different methods of analysis. On the other hand, Cerrón-Palomino (2016) provides three possible explanations that go beyond methodological differences. Firstly, he explains that the particular groups involved in this study are underprivileged groups with low socioeconomic status. These groups tend to have tighter networks, which in effect reinforce grammatical patterns. The second explanation Cerrón-Palomino (2016) provides deals with Hoffman and Walker's (2010) ethnolinguistic variation framework, which argues that distinctive physical characteristics make it difficult for some groups to assimilate to others thus preventing complete linguistic assimilation. In the case of Phoenix, Cerrón-Palomino (2016) argues that the awareness of the different phenotypes among the Mexican Americans and Anglos makes it difficult for the Mexican Americans to assimilate to Anglo-American English. Thirdly, Cerrón-Palomino (2016) argues that the location of the studies accounts for the discrepancy in results. He states that because Arizona, New Mexico, and California share a border with a Spanish speaking country, this ensures continuous fluent contact with a monolingual Spanish speaking country. New York and Florida however do not have this contact thus allowing for easy assimilation to Anglo-American English.

Although the explanations provided by Otheguy, Zentella, and Livert (2007) and Cerrón-Palomino (2016) provide a more profound understanding of the discrepancy in SPE results, further research still needs to be conducted in order to have more conclusive explanations. For instance, Cerrón-Palomino (2016) attributes this discrepancy to continuous contact Arizonans, Californians, and New Mexican speakers have with Mexico. However, the question now arises as to how different this contact is from the

constant contact New York speakers have with Puerto Rican and other Caribbean speakers who constantly immigrate to New York from their islands. Moreover, a study comparing SPE rates among Mexican English-Spanish bilinguals in Arizona with those in New York City will shed more light on the differences between these speakers. Furthermore, most prior research has failed to take into account the genre in the analysis. This is of utmost importance because certain types of texts are more likely to promote SPE rates much more than others. For instance, as seen in Travis's (2007) study mentioned above, argumentations favored SPE much more than narrations in her study.

In addition to studies focusing on SPE in contact with English, many other studies have focused on Spanish in contact with other romance languages such as Catalan (e.g. Prada Pérez, 2015), Italian (e.g. Flege, MacKay, & Piske, 2002), Greek (e.g. Argyri & Sorace, 2007) and Portuguese (e.g. Carvalho & Besset, 2015). Prada Pérez (2015), for instance, compared Spanish to Catalan and bilingual to monolingual varieties of Spanish. She strictly focused on the first person singular subject pronouns in Spanish in contact with Catalan. Participants were divided into four groups; two control groups and two bilingual groups. The control groups were made up of 12 monolingual speakers from Valladolid whereas the 12 Catalan controls were Catalan dominant speakers from villages around Minorca. She found an overall overt pronominal rate of 20.5%. In addition, the four groups used the overt first person singular pronominal rates in a similar fashion (19.8% among the Spanish controls, 20.7% among the Catalan controls, 19.9% among the Spanish bilinguals and 21.3% among the Catalan bilinguals). The language group in this study thus was not a significant variable.

Carvalho and Besset (2015) examined SPE in Spanish in contact with Portuguese and argued that SPE in this variety did not indicate completely divergent cross-linguistic behavior. They find contrary to the expectation that both grammars would have merged due to their similarities, this study showed no convergence towards Portuguese. Moreover, these results countered the hypothesis that bilinguals will express more subjects due to simplification (Sorace, 2011). Additionally, Carvalho and Besset (2015) mention that the differences among constraint rankings with respect to grammatical person are in line with that of their monolingual counterparts.

SPE in Spanish in contact with non-European languages has also received considerable attention (e.g., Barnes, 2010; Cerrón-Palomino, 2018; Essah, forthcoming; Michnowicz, 2015, Ortiz, 2011). For instance, Michnowicz (2015) investigated SPE in Yucatan Spanish in contact with Maya and found that Yucatan Spanish did not differ from that of Mexican speakers in other parts of the country and New York City. He found an overall pronominal rate of 19.7%, which is not only very similar to Solomon's (1999) study of Yucatan Spanish (a pronominal rate of 20%) but also to Otheguy, Zentella and Livert's (2007) study of Mexicans in New York (a pronominal rate of 19%). Moreover, in this study, Maya-speakers produced significantly more overt pronouns than did monolingual Spanish speakers (23.5% versus 16%, respectively). Conversely, Cerrón-Palomino's (2018) study of Spanish in contact with Quechua found no statistically significant difference in pronoun rates between Quechua-Spanish bilinguals and Spanish monolinguals (17.6% for bilinguals and 15.1% for monolinguals).

It is noteworthy that several theories have been proposed to account for the source of contact-induced change. Sorace (2011) for instance with the Interface hypothesis (already discussed in chapter 3) suggests that processing is the reason for the contact-induced change. She predicts that irrespective of the language pair, bilinguals will overuse subject pronouns in order to reduce the cognitive load in the processing of pragmatic content. Prada Pérez (2015) suggests that bilingualism affects lower-ranked internal variables. Thus, bilinguals tend to go with the language that has a more categorical distribution (in other words, the language that provides a fixed number of options to the speaker). Prada Pérez (2009) finds evidence for convergence in both Spanish L1 and Catalan bilingual's preference for overt pronominal subjects. According to Prada Pérez (2009), this is in spite of the fact that non-co-referential subjects contexts tend to favor overt lexical subjects in Spanish and overt pronominal subjects in Catalan.

To date, very few studies have tested the interface hypothesis in order to account for differences between bilingual and monolingual use of SPP. Some of these studies give support to this hypothesis (e.g., Barns, 2010; Michnowicz, 2015), while others have found no evidence in support of this hypothesis (e.g., Cerrón-Palomino, 2018; Carvalho & Besset, 2018). For instance, Cerrón-Palomino (2018) found (contrary to the tenets of the interface hypothesis) that there were similar SPE rates between the Quechua Spanish bilinguals (17.1%) and the Spanish monolingual group (15.1%) in his study. Carvalho and Besset (2015) similarly provide results that counter Sorace's (2011) prediction that bilinguals will show higher rates of SPE due to simplification. In this study among the 2,641 tokens in Uruguayan border Spanish, 25% were verbs with the overt SPP, whereas, among the 1,181 verbs expressed in Uruguayan Portuguese, 46% were overt SPP. In

addition, with the exception of Clause type that was significant in Uruguayan Portuguese but insignificant in Uruguayan Border Spanish, the same factors groups operated in both languages in the same hierarchy or order of importance. In order of strength, these factor groups were grammatical person, TMA, co-referentiality, Clause type (only significant in Uruguayan Portuguese), age group, socio-economic group, and sex.

Despite the ongoing debate regarding SPE rates and bilingualism, the above-mentioned studies prove the highly systematic nature of variable linguistic behavior. Linguistic factors that condition SPE in bilingual speech communities have been found not to differ from that of their monolingual counterparts: grammatical person and number switch reference, priming, TMA, lexical semantics of the verb, clause type, and reflexivity. All the studies have found person and number to be the strongest predictor of SPE (first and second person singular pronouns have the highest SPE rate), followed by reference or TMA. Some bilingual studies nonetheless have noticed minor differences in variable and constraint rankings when monolingual and bilingual groups are compared. Cerrón-Palomino (2018) for instance found that the constraints for monolingual Spanish and bilingual Spanish-English groups looked identical except for the grammatical person and number group. Whereas the overt SPP for the Spanish dominant group was the first person, for the English dominant group it was the third person. Similar results are also attested in Shin (2014) where participants used narratives to speak about life in New York City, visits to the homeland, differences between life in the homeland versus New York City, politics and governance in the USA and Latin America. At the end of the analysis, Shin (2014) found that TMA favored SPP in the New York group but not among Spanish speakers raised in Latin American. In terms of constraints. Besset (2018)

found that although reflexive verbs were not a significant predictor of SPE in the monolingual group, they were significant in the bilingual group.

In sum, there is no consensus regarding the effect of social or language-contact factors on SPE rates. Researchers have attributed varieties with high SPE rates to either contact with English or contact with Caribbean varieties of Spanish. Moreover, unlike monolingual studies, bilingual studies have found several social factors as significant predictors of SPE: level of bilingualism, age, sex, language group, and time in the bilingual environment. With respect to level of bilingualism, there is still no consensus regarding its effect on SPE rates. Some studies have found higher SPE rates in bilingual varieties in comparison with monolingual varieties whereas others have found no differential effect. Regarding age, differences within age groups and SPE have been attested. Some studies have found younger speakers favoring more overt pronouns (e.g. Flores-Ferrán, 2007) whereas others have found older speakers favoring overt pronouns (e.g., Lastra & Butragueño, 2015; Orozco & Guy, 2008; Prada Pérez, 2015). With regards to sex, whereas some studies have found no effect (e.g., Holmquist, 2012; Orozco & Guy, 2008; Otheguy, Zentella, & Livert, 2007) others have found an effect for sex (e.g., Bayley & Pease-Alvarez, 1996; Carvalho & Child, 2011; Otheguy & Zentella, 2012; Shin, 2013; Shin & Otheguy, 2013). In the studies that have found an effect for sex, it has been women/female participants who use more SPE than males. Also, language group was found to be significant in studies such as Essah (forthcoming) which looked at African languages in contact with Spanish (Fang, Bube, Fá d'Ambu, Kombe, and Balengue) and Michnowicz (2015) for Spanish-indigenous language bilinguals (Maya-

Spanish versus Spanish only). Nonetheless, in Prada Pérez's (2015) study of Spanish-Catalan bilingual, language group was insignificant.

From the above review of literature, it follows that findings in bilingual varieties have been less consistent, in comparison to those of monolingual varieties. There are key questions and notions that have still not been discussed in the literature. Moreover, in spite of this immense research with respect to SPE, to date, there are almost no variationist studies regarding the effect on SPE when Spanish is in contact with African languages. In view of this, the following section addresses previous studies conducted on the only African country with Spanish as an official language, Equatorial Guinea.

Previous studies conducted on Equatoguinean Spanish. Previous studies conducted on Equatoguinean Spanish have almost exclusively focused on dialectology, for instance, the works of Castillo Barril (1964), Echeragay (1951), De Granda (1984, 1990), Lipski (1984, 1995, 2002, 2004, 2008), and Quilis and Fresnillo (1995). Others have focused on phonology (e.g. Granda, 1984; Quilis, 1992), semantics (Nistal, 2009; Nguen, 2009) and syntax (Vuskovich, 2013). To date, Essah (forthcoming) is the only study that has examined Spanish in Equatorial Guinea using the variationist approach and as such, this study will be the focus of this section.

The goal of Essah (forthcoming) was to answer three research questions. Firstly, to investigate which linguistic and social constraints favor the use of SPE in Equatoguinean Spanish. Secondly, to find out how this variety compares with other varieties of Spanish regarding SPE. Thirdly, to examine if the pattern of SPE found in Equatoguinean Spanish could be explained by specific contact or by universal bilingual

properties such as the Interface Hypothesis (IH). Data for this study came from interviews conducted by Quilis and Casado-Fresnillo (1995) over two decades ago. The participants from Quilis and Casado-Fresnillo's (1995) study were native speakers of Fang, Bubi, Kombe, and Annobonese, with the majority being natives of Fang. In these recordings, speakers narrated various stories, events, traditions, songs, and poetry in approximately twenty minutes to one hour. These recordings of 21 Equatorial Guineans (16 males, five females) were used by Essah for her forthcoming study. With respect to the native language of these speakers, 11 speakers were speakers of Fang, three were speakers of Bube, two spoke both Fang and Bube as their first language, and three were speakers of Kombe and one of Fá d'Ambu.

Essah (forthcoming) transcribed the recordings, coded them and organized them into linguistic and social variables. The linguistic factor groups looked at were grammatical person and number, reference, ambiguity, verb class, ambiguity, and reflexivity. The extra-linguistic factors were age, sex, occupation, educational level, and L1. The data was then analyzed with inferential statistical software Goldvarb X in order to verify the impact of the factor groups (social and linguistic) on the subject pronoun expression rate. The results from this study revealed that the overt personal pronoun was expressed in 144 cases (22.4%) and unexpressed in 499 (77.6%) cases. These results place Equatoguinean Spanish as one with a low SPE rate, lower, in fact than all South American varieties, except for Peru.

Five significant factors were found in this study. The first conditioning factor was reflexivity. In this category, non-reflexive verbs were stronger predictors than reflexive verbs. Since the latter already contain identifying information regarding the speaker, it

hampers the use of the subject pronoun. The second conditioning factor was grammatical person and number. In this category, *Yo* had a higher factor weight than any other pronoun (*tú, él/ella, nosotros* and *ellos*). The third conditioning factor was reference. Same reference was a stronger predictor of SPE than switch reference and this had not been found in prior studies. Essah (forthcoming) argued that although the results regarding reference contrast with that of previous studies, these results were not only in line with the interface hypothesis (Sorace & Filiaci, 2006) but also they were in line with previous studies that showed bilinguals unpragmatic use of switch reference (Shin & Otheguy, 2009; Michnowicz, 2015; Shin, 2013; Shin & Otheguy, 2009). She added that a possible explanation could be the typology of the L1 spoken in Equatorial Guinea. Speakers of Equatoguinean Spanish possess a native language in which pronouns act differently from Spanish, in that they are prefixes amalgamated to the verb, the low pronoun rate could be due to these L2 speakers transferring their knowledge of the L1 to the L2.

The fourth conditioning factor Essah (forthcoming) found was L1. In order of hierarchy, Fá d'Ambu, Kombe, Fang, Mestizo (Fang-Bube) and Bube. Essah (forthcoming) mentions that this significant factor indicates the dissimilarity in the grammar of the local languages. She states that according to Bibang Oyee (1990), Fang differs from Spanish in that there are two types of pronouns; pronouns that designate persons (first and second persons) and pronouns that designate names of classes (people or things spoken about) or the third person. Moreover, according to Bolekia (1991), in Bube, there are pronouns designated as tonic pronouns act as subject pronouns while those designated as atonic pronouns act as objects. This dissimilarity in the local

languages could account for the significance of the L1 as an indicator of SPE rates. The fifth conditioning factor was age. Like previous studies, the younger generation used more subject pronouns than the older generation. The insignificant factors in this study were occupation, verb class, and ambiguity.

Although this study can be considered the first step towards a more profound understanding of subject pronoun expression in Equatorial Guinea, there is a need for a more up-to-date study since data from this study dates back 25 years. The present study aims at filling this gap in previous literature by investigating SPE in Equatoguinean Spanish.

Justification for the Present Research

A closer look at the above-mentioned studies reveals a number of gaps and shortcomings. For instance, despite decades of research, there is still no consensus regarding the relationship between SPE and bilingualism and as such, a number of questions still need to be answered. Moreover, whereas some studies have argued that SPE rates are higher in bilingual varieties than in monolingual varieties, others have found no differential effect between the two. Several explanations have been given to account for the increase in overt SPP in bilingual varieties, e.g., simplification, convergence, and of interest to this essay, the interface hypothesis. This dissertation thus will be one of the first variationist studies to examine the IH predictions in bilinguals speaking African languages alongside Spanish. Moreover, Equatorial Guinea being the only country where Spanish is strictly an L2 in a stable bilingual environment makes it possible to test this hypothesis.

Furthermore, in the analysis of linguistic and social variables, several methods have been employed in the SPE literature: ANOVA, correlations, cross-tabulations, and multivariate regression analysis such as GoldVarb. However, very few studies have used a mixed effects model like Rbrul, which allows the researcher to identify the sources of variation and correlation that arise from the group data. Moreover, in the analysis of unbalanced data, mixed effects models are more robust than analysis with General Linear Models (Pinheiro & Bates, 2000). Mixed effects models also allow the researcher to analyze inter-speaker variation, which has not been done in the majority of previous SPE studies. A mixed-effects model is therefore needed to contribute to this discussion on SPE.

Finally, previous research has almost exclusively focused on Spanish in contact with Indo-European languages, Portuguese, Catalan, and creoles. To date, except for Essah (forthcoming), no variationist study has examined Spanish in contact with African languages. Thus, there is a need for a study with speakers of Equatorial Guinea in order to shed light on the processes at work in bilinguals/multilinguals of Spanish and African languages. To fill this gap in the literature, this dissertation examines factors that may interact with subject pronoun expression in Equatoguinean Spanish.

Summary

The goal of this chapter has been to review previous studies conducted on Subject Pronoun Expression using the variationist approach. This chapter began with a review of studies on monolingual varieties of Spanish before progressing to bilingual varieties of Spanish. Previous studies show that linguistic factors such as grammatical person and number, reference, ambiguity, verb type, and priming are consistently significant factors

in studies in which they have been analyzed. Extra-linguistic factors on the other hand, such as age, sex, education, genre, bilingualism, among others, have not been that consistent across varieties. Thus, linguistic factors seem to condition Subject Pronoun Expression much more strongly than extra-linguistic factors. In the next chapter, the methodology used in the present study is examined.

CHAPTER 5

METHODOLOGY

The goal of this chapter is to discuss the methodology used in this study. As stated earlier (see Chapter 4), although there is extensive work published on SPE, to the best of my knowledge, the variables that constrain SPE in Spanish in contact with African languages have been largely unexplored (eg. Essah, forthcoming). To fill this gap in the existing literature, the present study draws on a naturalistic corpus via sociolinguistic interviews as proposed by Labov (1972). Even though this method is not free from limitations, in comparison to other experimental methods such as grammaticality judgment tasks, reading exercises, and matched guise techniques, sociolinguistic interviews are ideal for soliciting natural and spontaneous speech.

The island of Bioko, specifically Malabo, was selected as the site for this research and collection of data. It was selected because it is more cosmopolitan in comparison to the other regions of the country thus providing access to people from different backgrounds living in close proximity to each other. Malabo is the capital of Equatorial Guinea and it is the oldest city in the country. It lies on the northern edge of the island of Bioko. It is the center for trade and commercial activities for Equatorial Guinea and it houses the only international airport, the embassies, and a harbor where cocoa, coffee, and timber are exported. It is also a large urban area that has attracted people from different parts of the country.

Being an outsider to the community whose speech was under investigation, I had to establish contacts in Malabo in order to have the investigation carried out. The Ghana

embassy was contacted and through this connection, a research assistant with sociolinguistic training was hired to conduct the interviews. We communicated by phone on numerous occasions and she was made aware of the type of study, the goal, and how the interviews were to be conducted. She was also sent videos of me conducting similar interviews so she had a good idea of how the interviews were to be conducted. Moreover, the interview questions and a recorder that could record about 200 hours of interviews were sent to her. The recorder had to be sent because it was almost impossible to get a good quality recorder in Equatorial Guinea. On the day the research assistant conducted the first interview, she sent me an audio file of a recording she had carried out before further interviews were conducted.

The rest of this chapter is organized into four main sections. I begin with an elaboration of the research questions that guided the study, next I describe the participants, data collection, and finally, I provide a summary of the chapter.

Research Questions

As noted in previous chapters, SPE reveals the distinction between groups of languages that allow null subjects (pro-drop) and languages that do not allow them (e.g., Cerrón-Palomino, 2016; Flores-Ferrán, 2007; Orozco & Guy, 2008; Otheguy & Zentella, 2012). Although over time an extensive literature has developed on SPE, the question of SPE in Spanish spoken in Africa has not been addressed. While Essah's (forthcoming) study on SPE in Equatorial Guinea is the first step towards a more profound understanding of SPE in this variety, the fact that the data used for this study were

collected more than 25 years ago calls for the need of the present study. To fill this gap in the literature, the present study poses the following research questions:

- (a) What is the null and overt pronoun rate¹ in Equatoguinean Spanish? How does this rate compare with pronoun rates in other varieties of Spanish?

The literature on SPE strongly suggests that there are differences in terms of pronoun rates between regional varieties of Latin America and Peninsular Spanish varieties (Carvalho, Orozco, & Shin, 2015). As stated earlier (see Chapter 4), the highest SPE rate was reported in Cameron's (1994) study of Puerto Rican Spanish (a pronominal rate of 45%) while the lowest rate (16.2%) was reported in Cerrón-Palomino's (2018) study of Huancayo in Peru. Although Essah's (forthcoming) study placed Equatoguinean Spanish as one with a low SPE rate (22.4%), because Essah's (forthcoming) study was based on data collected more than 25 years ago, there is a need for a more up to date study. This study fills this gap in the literature by examining overt and null pronoun rates using the variationist approach.

- (b) What linguistic and social constraints favor the use of SPE in Equatoguinean Spanish?

Previous variationist research on SPE has found several linguistic and social constraints favoring the use of SPE in Spanish (e.g., Barrenechea & Alonso, 1977; Bayley & Pease-Alvarez, 1997; Bentivoglio, 1977; Cameron, 1993, 1995; Enríquez, 1984; Erker & Guy, 2012; Flores-Ferrán, 2004; Morales, 1980; Orozco & Guy, 2008; Otheguy & Zentella, 2007, 2012; Otheguy, Zentella & Livert, 2007; Posio, 2011; Prada

¹ The term rate used here refers to the percentage of null and overt subjects that are actually expressed in comparison to how often they could have been expressed but were not.

Pérez, 2009; Shin & Otheguy, 2013; Silva-Corvalán, 1982, 1994; among many others). Although the linguistic and social variables vary from author to author, these studies have brought to light the fact that in monolingual varieties, linguistic factors appear to condition SPE more strongly and consistently than social factors. However, in bilingual varieties, bilingualism in itself leads to subtle changes regarding the linguistic and social constraints that favor the use of SPE (Carvalho, Orozco, & Shin, 2015). This study aims at shedding more light on how these constraints work.

(c) Does SPE in Equatoguinean Spanish support or contradict universal trends of SPE in bilinguals, as proposed by the Interface Hypothesis?

As previously discussed in Chapter 3, the IH (Sorace, 2006; Sorace & Serratrice, 2009; Sorace, 2011) was proposed to account for some of the non-target like behavior found in the speech of adult second language learners. Its main tenet is that not all linguistic interfaces share the same properties. “External” interfaces, such as subject reference shift present a greater challenge during acquisition than “internal” interfaces (Tsimpili & Sorace, 2006).

However, the IH has produced mixed results among various researchers (see Chapter 3). Some researchers have provided evidence in support of its claims and others have not. For this reason, additional studies are needed to understand more completely the key tenets of the IH. Consequently, with Equatorial Guinea being a bilingual population/multilingual population and having an L2-only status of Spanish, this study presents an ideal testing ground for universal trends as proposed by the IH.

(d) What does SPE in Equatoguinean Spanish reveal about language contact between Spanish and the local languages that are spoken in Equatorial Guinea?

Previous research has shown that there are differences between contact and non-contact varieties of Spanish (Carvalho, Orozco, & Shin, 2015) and bilingualism has been associated with simplification, convergence, transfer, among others (see Chapter 3). Sorace (2004) explained that simplification, for instance, occurs in contact varieties because of the need to simplify the cognitive load of having to remember two different systems. Bullock and Toribio (2004) also suggested that languages in contact led to the languages converging towards each other. Granted that Spanish is a Romance language in contact with languages from the Niger-Congo group, creoles, and languages from the Indo-European family (English, French and Portuguese), the goal of this study is to examine what the use of SPE in Equatoguinean Spanish reveals about language contact between Spanish and the other languages that are spoken in Equatorial Guinea.

Participants

In total, 30 participants took part in individual sociolinguistic interviews. These consisted of 18 males and 12 females. Concerning age, the youngest participant in the study was 18 years of age whereas the oldest was 64 years of age. It was difficult to get participants for the study, even in spite of the financial compensation they were going to receive. This is because of the fear most people had of the government. Interviews of this nature are not common in Equatorial Guinea and as such, none of the participants had ever participated in interviews before. For this reason, many were curious and apprehensive as to what they entailed.

As the country is a bilingual/multilingual speech community, my goal was to get a representative sample of speakers of different native languages spoken in Equatorial Guinea. Hence, there were 12 native speakers of Fang, 14 native speakers of Bube, three native speakers of Kombe, but only one of Fá d'Ambu. At the beginning of the interview, the participants filled out a language background questionnaire (see Appendix A) to collect demographic data. In the questionnaire, participants were asked where they were born, the background of their parents, what languages they speak and their proficiency in these languages (basic, advanced, or intermediate).

The information obtained from the language background questionnaire showed that every participant spoke at least two languages- for example, Spanish and a native language, or Spanish, a native language and English or French. All participants stated that they felt comfortable speaking Spanish. They mentioned that speaking Spanish (as an acquired *lingua franca*) facilitated conversations with people of different native languages and as a result, Spanish was used very frequently in their day-to-day activities. Several participants also expressed their pride in belonging to the only Spanish-speaking African country. They also stressed that it was important to speak Spanish in Equatorial Guinea because it was a means to move up the social and economic ladder due to its status as the official language of government and industry.

Concerning education, educational backgrounds ranged from elementary education to college education. However, even within each group, participants were not completely homogenous. Participants differed regarding the number of elementary school education, high school education or college education they had received. For instance,

some participants who described themselves as college-educated had just started taking college classes at the time of the interview, and some of those who described themselves as high school educated had only taken a few years of high school classes before they dropped out.

Moreover, all participants were born in Equatorial Guinea and of Equatoguinean parents. With the exception of two participants (one who had spent six years in Gabon, and another who had spent a short vacation in Spain), none of the participants had ever left the country.

To summarize this section, details regarding the individual participants have been provided in the table below.

Table 1. Participant information (n=30)

Speaker	Sex	Age	Education	L1	Other language
1	male	64	elementary	Kombe	Spanish, French
2	male	35	college	Bube	Spanish
3	female	25	college	Fang	Spanish, French
4	male	30	high school	Fang	Spanish, English
5	male	29	college	Kombe	Spanish, English
6	female	36	college	Fang	Spanish, English, French
7	male	35	college	Fang	Spanish, French
8	male	32	college	Bube	Spanish, English
9	female	67	elementary	Bube	Spanish, French
10	male	69	high school	Bube	Spanish, French, Pichi
11	female	26	college	Bube	Spanish, French
12	male	29	high school	Fang	Spanish
13	male	32	high school	Bube	Spanish. English, French
14	female	24	high school	Fang	Spanish, French, Bube
15	male	18	college	Bube	Spanish, French, English
16	male	44	high school	Fá d'Ambu	Spanish, French, Bube
17	male	21	high school	Bube	Spanish

18	female	31	college	Fang	English, Pichi
19	female	32	high school	Bube	Spanish, English
20	male	22	college	Kombe	Spanish, English
21	female	35	college	Fang	Spanish
22	male	24	college	Fang	Spanish, English
23	male	32	college	Bube	Spanish, English
24	female	21	college	Fang	Spanish, English
25	female	28	college	Fang	Spanish, French, English
26	male	26	college	Fang	Spanish
27	female	20	college	Fang	Spanish
28	male	21	college	Bube	Spanish
29	female	20	college	Fang	Spanish
30	male	20	college	Bube	Spanish

Data Collection

Two main methods of recruitment were used for data collection; the *snowball* and the *random sampling method*. With the *snowball method*, initial research participants recruit future participants from among their acquaintances. Thus, not every participant has the chance to be included in the sample. In the present study, the interviews using the snowball technique took place at the various homes and workplaces of friends and acquaintances who had agreed to be interviewed. Although the snowball method may introduce some type of bias regarding the participants for the interview, it was the most convenient way to get the required number of participants needed for the study. Moreover, because Equatorial Guinea is a small country, most Equatoguineans have tight social networks, know people through friends, and extended family thus facilitating the use of the snowball method.

In contrast to the snowball method, *random sampling* gives every member of the community an equal chance of being included in the sample. Interviews using this

method took place at a shopping mall in the heart of Malabo. The random sampling method gave access to participants from different language groups and social classes, who would not have been represented otherwise through the snowballing technique. The shopping mall was chosen because it is a busy location and it is the main destination for buyers and sellers in Equatorial Guinea. The interviewer introduced herself to potential interviewees stating that she was conducting interviews for a dissertation project directed at finding out the life experiences of residents of Malabo. The reason for not revealing the exact nature of the research is in line with the guidelines for conducting sociolinguistic fieldwork, which is to prevent the interviewee from modifying his/her, speech to produce a particular result (Labov, 1972).

After participants agreed to take part in the interview, they were informed that their identity was going to be hidden to protect their privacy. They were also informed that the use of the recorder was simply to remember the information participants provided because writing everything down would be impossible. Moreover, participants were made aware that they were not obligated to take part in the interview. For this reason, if they felt uncomfortable answering any questions during any part of the interview, they had the option to opt-out. Many participants felt at ease after receiving this information.

The first part of the interview consisted of participants filling out a language background questionnaire, and demographic information (see Appendix A). For instance, all participants had to state what their native language(s) was/were, what language they used at home, the background of their parents, their place of birth and if they had ever

lived in a Spanish speaking country. These questions were solicited to ensure that participants were eligible to take part in the study.

Participants took an average of 10-15 minutes to complete this assignment. The interviews were recorded with a Panasonic recorder with a sound equalizer and they ranged from 35 minutes to one hour. The interviewer began the interviews by introducing herself. She then asked the name, family origin, place of birth, age and educational level of each participant.

Labov (1984) argues that the order of questions in a sociolinguistic interview should be in the form of conversation modules. Thus, the modules in the present study revolved around the topic of schools, music, food, family gatherings, among others. Moreover, according to Labov (1984), questions within each module should be hierarchically structured. In other words, one moves from more general and impersonal questions to specific and personal ones in each module. Thus in the present study for instance, the module on schooling began with questions relating the educational system in Equatorial Guinea before progressing to questions regarding a typical day at school, teachers, and punishments (See Appendix B).

Labov (1972) found that participants paid less attention to their speech whenever they spoke about emotional topics. For this reason, many of the interview questions were questions that solicited the description of emotional circumstances. For instance, a description of the most unfair punishment one has ever experienced in school, a description of one's favorite food, among others. Nonetheless, although there were planned interview questions, participants were allowed to freely elaborate on topics they

wanted to and as such, the interviews did not produce the same information. However, the same techniques were used in each interview and participants narrated their experiences in a casual style (Labov, 1984).

The most frequent topics participants brought up were their dissatisfaction with their current economic situation, the non-fulfillment of their dreams and their hope that God would resolve their economic hardship. I also noticed that whereas some speakers were very eager to narrate certain experiences to the interviewer, others completely avoided these topics or answered abruptly and then moved on to another. I later realized that this was because some of the questions I had prepared were not culturally suitable to this community. For instance, most participants avoided the topic of romantic relationships or describing an ideal partner. Also, for the prompt regarding the most unjust punishment they had received in school, almost every participant said that their punishments were well deserved. They believed that punishments were a means to ensure that they became good citizens of the country. Furthermore, I noticed that for the questions on education, almost every participant who did not have a college degree when asked if they had been to college, said they had plans of entering college that coming year. I deduced that this answer could be due to a college education being linked to one's status in the country. At the end of the interview, the interviewer thanked participants and wished them the best in their future endeavors.

Following the interview session, each interview was saved in an mp3 format with a code name for participants to guarantee their anonymity. The recordings were then transcribed a few weeks after all interviews were completed. In total, there were about 28

hours of transcribed speech. In transcribing the interviews, I wrote exactly what participants said, without any edits or corrections. For this reason, some of the examples I might provide in subsequent sections may contain some lack of grammatical agreement and other features of Equatoguinean Spanish. The data analysis procedures will be described in the next section.

Data Analysis Procedures

Following variationist methodology (Tagliamonte, 2006), it is important to analyze all cases in which the phenomena appeared versus where it could have appeared but it did not, as seen in example 1.

(1) (*Yo*) *Tengo 35 años* (Speaker 2)

I am 35 years old

The above sentence has two variants. “*Yo tengo*” or “*Tengo*”. Thus, the number of cases in which the pronoun could have been produced but was not produced is compared to the number of cases in which it was produced.

Furthermore, in line with previous studies (e.g., Cerrón-Palomino, 2016; Silva-Corvalán, 1982) only tensed verbal forms with human referents were included in the study. Thus, a variety of sentences fell outside the envelope of variation. For instance, impersonal sentences or sentences that lacked a human referent such as (2a), (2b) and (2c) which is a non-personal pronoun subject, flip verbs like *gustar* which tend to have postverbal subjects and rarely SPPs, as seen in (3a), discourse markers (4a) and (4b), imperatives (5a) and subject pronouns with emphatic *mismo* (6a) and (6b). In each of these cases, the null subject fails to alternate with the expressed subject.

(2) Outside the envelope of variation

(a) *Es el español* (speaker 4)

It is Spanish.

(b) *Sí, se habla el francés, el inglés, el chino, el portugués* (speaker 5)

Yes, French, English, Chinese and Portuguese are spoken.

(c) *Algunos pasan por su propio esfuerzo, otros porque tienen dinero*
(speaker 6)

Some pass because of their own effort, others, because they have money.

(3) Flip verbs like *gustar* which tend to have post verbal subjects

(a) *Me gustaría ser abogado* (Speaker 3)

I would like to be a lawyer.

(4) Discourse markers

(a) *Pues vamos que la comida congelada, no es buena porque nos produce enfermedades.* (Speaker 19)

Well frozen food is not good because it gives us diseases.

(b) *yo iba a clase en la secundaria desde la mañana que uno sale de su casa para tomar buena clase hasta digamos la tarde.* (Speaker 20)

I used to go to class in high school from the morning when one leaves the house to have a good class until **let's say** the evening.

(5) Imperatives

(a) *Aprende español* (Speaker 19)

Learn Spanish

(6) Subject pronouns with emphatic *mismo*

(a) *Yo misma voy a hacerlo* (Speaker 22)

I myself will do it

(b) *Él mismo lo hizo* (Speaker 24)

He himself did it.

Except for the above-mentioned cases, all other conjugated verbs were coded into overt and null subject pronouns resulting in 3980 tokens. In the next section, the factor groups analyzed and their predictions will be explained.

Factor Groups/Independent Variables Analyzed and their Predictions

In this section, the linguistic and extra-linguistic factors coded in this study are described. The analysis of these factors was informed by previous studies on SPE (see Chapter 4). Each of these factors along with their hypotheses will be described in detail in the following paragraphs.

Linguistic factors. The linguistic factors coded in this study were *grammatical person and number, reference, ambiguity, verb class, and reflexivity*. Previous studies have shown that *grammatical person and number* is one of the strongest and most consistent predictors of SPE (Otheguy, Zentella & Livert, 2007) and the different grammatical persons have different effects on SPE (see Chapter 4). For instance, the first and second person have been found to condition SPE much more strongly than other grammatical persons. For example, Orozco's (2015) study on Colombian Costeño Spanish found 44.5% of overt subjects in the first person singular *yo*, 39.6% in the third person singular *él/ella*, and 32.5% in the second person singular *tú*. Furthermore, singular subjects have been found to more likely to favor SPE rates than plural subjects. For instance, Lastra and Butragueño (2015) found the pronouns in the first person, third

person and second person singular forms were the strongest predictors of SPE.

Consequently, the data were coded according to the first person, second, and third person in both singular and plural forms. The prediction for this factor group is that singular verb forms will favor the overt variant more than plural verb forms and the second and first-person singular pronouns will be the strongest predictors of SPE.

Moreover, in all SPE studies, *reference* has been found to be a significant factor with a *change in reference/switch reference* being a stronger predictor of overt SPE whereas *same reference*, the predictor of null subjects (e.g., \emptyset *Hablo* 'I speak'). Cameron (1995) defined *switch-reference* as the "relationship of same or different reference between two sequentially ordered subjects" (p. 2). The following examples from the present study illustrate the concept of same reference and switch reference, respectively.

(7) *Si \emptyset podré contestarte, porque no solo \emptyset estuve en una escuela, \emptyset estuve en varias.* (speaker 23)

Yes (**I**) will be able to answer you, because not only was (**I**) in one school, (**I**) was in many.

(8) *\emptyset Me bañé porque ellos querían visitarme.* (speaker 23)

(**I**) bathed because **they** wanted to visit me.

For example (8), the same referent is being maintained throughout each of the clauses in the sentence. However, for example (8), the subject of the first clause is the first person singular while the second clause is the third person plural.

Previous studies have also found that subject pronouns tend to be expressed in switch reference situations, as seen in example (8), considerably more than those of same

reference as in (7), because references can easily be inferred from previous utterances in the same referent contexts. However, considering that the Equatoguinean variety is spoken by bilinguals, if the predictions of the IH are correct, switch reference will be a weak predictor of SPE.

In addition, *ambivalence* created by Tense Aspect and Mood endings has been found to favor the use of subject personal pronouns (e.g., Abreu, 2012; Carvalho, Orozco & Shin, 2015; Lastra & Butragueño, 2015; Travis, 2007). For instance, verbs in the first and third-person singular forms of the imperfect tense, present subjunctive and conditional have identical morphological markings thus making it difficult to deduce the identity of the referent (Carvalho, Orozco & Shin, 2015). In effect, an increase in ambiguity leads to an increase in SPE rates because subject pronouns are used to disambiguate these utterances. By way of illustration, example (9) is provided.

(9) *yo veía que los, los compañeros decían que ese maestro era malo, porque cuando yo había venido a esa escuela, me decían hay un maestro que pega fuerte.*

(Speaker 1)

I saw that the colleagues said that teacher was bad, because when **I** had come to that school, they told me that he beats strongly.

In the above example, the subject pronoun used with the imperfect tense provides more clarity as to who the referent is in the sentence. Thus in this study, verbs in the first and third person of the imperfect, conditional and the present subjunctive tense were coded as ambiguous verbs and other verbs that did not have identical forms were coded as non-

ambiguous. The prediction is morphological ambiguous forms will have their subject's expressed while non-ambiguous forms will have unexpressed subjects.

Furthermore, most studies have found that certain classes of verbs, e.g. *psychological verbs* (verbs that express a mental state), *copulas* (verbs that link subjects to complements or adjectives), *statives* (verbs describing an existence), and *speech act verbs* (communicative acts that express actions) are more likely to favor the expression of the personal pronoun whereas other verb types do not (e.g., Bentivoglio, 1987; Cerrón-Palomino, 2016; Orozco, 2015; Otheguy & Zentella, 2012; Prada Pérez, 2009; Silva-Corvalán 1994; Morales 1997; Travis, 2005, 2007; Travis & Torres Cacoullós; 2012). Thus in this study, I follow the classification of Travis (2007) and divide verbs into five classes: *psychological*, such as *saber* (know), *creer* (believe), *pensar* (think), *acordarse* (to remember), *contar* (to tell); *speech acts*, such as *decir* (say), *llamar* (call); *statives* such as *ser*, *estar* (be), *motion*, such as *ir* (go), *venir* (come), and *other*, which includes verbs not fitting into the above categories. Studies that have used this classification have found psychological verbs to be the most significant predictor of SPE rates (e.g., Cerrón-Palomino, 2017; Silva-Corvalán, 1994; Travis, 2007; Torres Cacoullós et al., 2010). According to Silva-Corvalán (1994), psychological verbs tend to be expressed with the subject pronoun because of the contrastive situation in which they are found. Hence, the prediction regarding this factor group is that, just like previous studies, psychological verbs will significantly favor the presence of SPPs, whereas the rest will disfavor it.

Finally, *reflexive verbs* (verbs whose direct object is the same as its subject) have been found to generally disfavor the use of SPE whereas *non-reflexive verbs* favor the use of SPE (Bayley & Pease-Alvarez, 1996). Because the clitic pronoun of reflexive verbs is

an indicator of the subject, personal pronouns tend to be omitted with reflexive verbs. Non-reflexive verbs on the other hand favor SPE because the listener needs information regarding the referent of the conversation. By way of illustration, the following example is provided from the present study:

(10) *Se quedó así por mucho tiempo* (speaker 4)

It stayed like that for a long time.

For example (10), the clitic pronoun of the verb provides information regarding who the subject of the sentence is. For this reason, placing a subject pronoun in front of this verb will make it redundant, unless it is for the sake of contrast. The hypothesis is that, along the same lines as what has been found in previous studies, reflexive verbs will disfavor the presence of the subject pronoun because they disambiguate subjects while non-reflexive verbs will favor the presence of the subject pronoun.

Extra-linguistic factors. The extra-linguistic factors coded were *age*, *educational level*, *L1* and *sex*. *Age* has been found to be a significant extra-linguistic factor in some SPE studies (e.g., Lastra & Butragueño, 2015, Orozco, 2015; Orozco & Guy, 2008) and insignificant in others (e.g., Alfaraz, 2015; Michnowicz, 2015). In the studies where age was significant, the younger generation tended to use more pronouns than the older generation who tended to adhere to more conservative and prescriptive grammar rules. In this study, age will be coded and run as a continuous variable.

Another extra-linguistic factor examined in this study is *L1*. As already stated, Spanish is not the first language of most Equatorial Guineans. Thus, there were 12 native speakers of Fang, 14 native speakers of Bube, three native speakers of Kombe, and one of

Fá d’Ambu. The L1 of speakers with different native languages was coded to determine if the native language of participants made a difference regarding SPE rates. Fang, Bube, and Kombe are Bantu languages that use prefixes in place of personal pronouns (Zamponi, 2009). These prefixes are amalgamated to the verb as seen in the following examples in Fang from Bibang Oyee (1990, p.72).

(11) *ma-dzin*

I -love

(12) *wa-dzin*

You-love

(13) *aa-dzin*

He/She –loves

Fá d’Ambu on the other hand is a language that maintains a fixed Subject Verb Object order and the sentence is organized into Noun Phrase-Verb-Adjective-Adverb (Sergobe, 2010). By way of illustration, the following examples from Sergobe (2010, p.112) are provided.

(14) *Antoñi* *bias* *mundadu pol uã jasôlô*

Noun Phrase Verb Prepositional phrase

Antony was bitten by a dog

I hypothesize that the native language will be a significant predictor of subject pronoun expression. Specifically, the speaker with Fá d’Ambu as a native language will use more subject pronouns than the speakers with Fang, Kombe, and Bube as native languages.

Furthermore, previous Variationist studies have coded for *sex* using the categories of male and female. According to Labov (1990), “of all the sociolinguistic principles, the clearest and most consistent one is the contrast between women and men” (p. 205). This has been explained as being caused by a variety of factors; biology (Chambers, 2003), a cultural pattern (Labov, 2001) and covert prestige (Trudgill, 1974). Nonetheless, studies looking at the effect of sex on SPE have not been consistent in their findings. Some studies have found sex to be a significant social factor (e.g., Alfaraz, 2015; Bayley & Pease-Álvarez, 1997; Carvalho & Child, 2011; Orozco, 2015) and in others it is non-significant (e.g., Ávila-Jiménez, 1995; Jia & Bayley, 2002). For instance, in Alfaraz’s (2015) study, females were seen to slightly favor overt pronouns, whereas males favored null pronouns. Similarly, in Otheguy and Zentella (2012), and Shin and Otheguy’s (2013) study of New York, women were the leading users of overt pronouns. Solomon’s (1999) study of Valladolid also found similar results. Although in the present study, there is a lack of balance between males and females (18 males and 12 females), sex will be included for exploratory reasons.

Education has also been used in previous studies as an indicator for social status (e.g., Otheguy & Zentella, 2012; Shin & Otheguy, 2013), since studies such as Shin and Otheguy (2013) found that the more affluent in New York City experienced the most increase in pronoun use while the least affluent experienced no change. Ávila-Jimenez (1996) found that unskilled workers favored the use of overt pronoun forms much more than college-educated speakers. Other studies, such as Yolanda Lastra and Butragueño found that education was not a significant predictor of SPE. In this study, there are two speakers with elementary education as the highest educational degree, eight with high

school education and 20 with a college education. I will thus divide speakers into three groups; college-educated, high school educated and elementary educated. Because Spanish is acquired through formal education in Equatorial Guinea, I hypothesize that education will be a significant factor and the elementary school educated will have the highest number of overt personal pronouns, followed by the high school educated. This will probably be because the high school and elementary educated have less exposure to the standard dialect in comparison to the college-educated who by exposure may be aware of the disambiguating and emphatic uses of the subject pronoun.

Summary

The goal of this chapter has been to describe the methodology used to examine SPE in Equatoguinean Spanish. As the present study adopts the variationist approach, sociolinguistic interviews were conducted with participants recruited from Malabo in Equatorial Guinea. Participants were recruited using two main methods; random sampling and the snowball technique. In total, 30 participants took part in the study. These participants are both male and female (18 males and 12 females) from different age groups (between 18-64 years), native languages (Fang, Bube, Kombe, and Fá d'Ambu) and with educational backgrounds ranging from elementary to college education.

After the interviews, the recordings were transcribed and analyzed. The envelope of variation included all tensed forms with a human referent. Impersonal sentences, discourse markers, imperatives, noun phrases, and emphatic *mismo* were excluded. The linguistic factor groups predicted to influence SPE were person/number, ambiguity, reference, verb type, reflexivity, and extra-linguistic factors were age, sex, L1, education

and individual speaker. In the next chapter, the analysis with Rbrul is described and the results of the study are presented.

CHAPTER 6

RESULTS AND DISCUSSION

This empirical study was aimed at analyzing the linguistic and extra-linguistic factors that constrain Subject Pronoun Expression in the Equatoguinean variety of Spanish. Equatorial Guinea was deemed an ideal testing ground for Subject Pronoun Expression due to its status as a Spanish L2-only country. Also, of particular interest to the study was whether the results from this L2-only variety refutes or supports the universal trends proposed by the IH.

The chapter is thus divided into three main sections. The first section presents the reader with a description of how analysis with Rbrul is conducted and interpreted. The second section addresses the research questions posed for the study. Finally, in the third section, there is a discussion of how the results obtained fit into the larger context of bilingualism and second language acquisition (L2) established by prior literature.

Analyzing with Rbrul

The data for this study were entered into the mixed effects models software Rbrul for quantitative analysis. Through regression analysis, Rbrul identifies which factor groups, both linguistic and extra-linguistic, affect the variable under study, in what direction and to what degree (Johnson, 2008). More explicitly, the software requests for the researcher to select one of the variants of the variable as an application value, which is the variant that the software will be yielding results on.

Following the SPE research tradition, the application value selected for this study was overt subject pronoun, and the linguistic factor groups analyzed were

person/number, reference, ambiguity, and verb type. By the same token, extra-linguistic factor groups were *age, education, sex, and LI*.

In a number of respects, Rbrul is an improvement from fixed-effects-only software like Goldvarb, the previous variationist standard analytical tool. Firstly, Rbrul allows for the study of *continuous variables*, not only *discrete variables*. A *continuous variable* is a variable that can take an uncountable number of values. An example in the present study is *age*. By selecting age as a continuous variable, the study avoided the division of participants into age groups. A *discrete variable*, on the other hand, is a variable that takes values countable in a finite amount of time. Examples in the present study are sex (male/female) and reference (same/switch reference).

Secondly, and most importantly, Rbrul allows for the distinction between *fixed effects* and *random effects*. Fixed effects are factors that have limited possible levels and as such, their levels are replicable in other studies (Johnson, 2008). For example, the fixed effects used in this study were *male/female, college/high school/elementary educated, switch reference/same reference, ambiguous/non-ambiguous, and the personal subject pronoun forms*. *Random effects*, on the other hand, are factors that are drawn from a larger population (Johnson, 2008). Unlike fixed effects, random effects are not replicable. For instance, speakers drawn from a specific population cannot be replicated in another study. In particular, Rbrul can account for interspeaker variation when individual speakers are analyzed as a random effect. Accounting for inter-speaker variation makes it possible to determine which speakers influence the most/least the variable in question. It also provides validation for the study in that it shows that the

linguistic and extra-linguistic factors are significant in spite of the individual differences (Tagliamonte, 2011).

Like previous regression analytical tools, in analyzing the linguistic and extra-linguistic factors, Rbrul follows a *step up/step down* regression. In the step-up regression, Rbrul adds factor groups that significantly affect the variable in question. It starts with zero factor groups and repeats this until there are no more factors to add. After the completion of the step-up regression, Rbrul begins with the step-down regression, starting with all of the factor groups and then systematically subtracting them. In this regression, Rbrul eliminates predictors that are not significant and it repeats this until there are no more factors to eliminate. If the best step-up and step down match, both runs are selected as the best model and Rbrul gives a summary of the factor groups and the probability of each factor within the factor group (Johnson, 2008). After obtaining the results, the researcher then explains them in terms of factor group ranking (which factor groups are stronger in affecting the linguistic variable) and in terms of constraint ranking, which constraints, within each factor group have the strongest probabilities to predict the application value. However, if the best step-up and best step-down do not match, the researcher will have to re-examine the results in more detail.

The results presented by Rbrul are in the form of probabilities, in both *factor weights* and *log-odds*. A *factor weight* is a numerical value that describes the extent to which a combination of variables favors or disfavors the choice of a particular form. Factor weights are reported within a range of 0 and 1. A factor weight higher than 0.5 generally favors the rate of occurrence whereas a factor weight lower than 0.5 disfavors the rate of occurrence. Factor weights of 0.5 are relatively neutral. *Log-odds*, like factor

weights, depict the strength of the factor and the dependent variable. Thus, a positive log-odds value indicates that the factor favors the dependent variable whereas a negative log-odds value indicates that the factor disfavors the dependent variable. The higher the log-odds value, the stronger the relationship between the factor and the dependent variable.

The next section addresses the research questions posed for the study.

Research Question 1

The first research question asked what the null and overt subject personal pronoun (SPP) rate was in Equatoguinean Spanish, and how this rate compared with pronoun rates in other varieties of Spanish. In total, there were 3980 tokens coded from the 30 speakers who took part in this study. The overall distribution of the variable under scrutiny was 19.1% overt subject pronouns, and 80.9% null subjects. These results not only confirm the variability of SPE but also, they show that the Equatoguinean variety is a variety with more null subjects than overt subject pronouns. Table 2 displays the overall distribution of null and overt SPP's in Equatoguinean Spanish.

Table 2. The overall distribution of null and overt SPPs in Equatoguinean Spanish

Variant	Number	Percentage
Overt pronoun	760	19.1
Null pronoun	3220	80.9
Total	3980	100

The overt SPP rate in Equatoguinean Spanish contrasts strongly when compared to the 45% rate found in Cameron's (1994) study of Puerto Rican Spanish, the 42.3% found in Alfaraz (2015) study of Spanish spoken in Santo Domingo and with Otheguy and Zentella's (2012) study of Spanish spoken in New York by the Dominicans, Puerto Ricans, and Cubans (41%, 39%, and 38% respectively). Thus, compared with these

Caribbean varieties of Spanish, Equatorial Guinea shows a clearly low overt SPP rate (19.1%) as seen in table 9 below.

On the other hand, the Equatoguinean Spanish rate is similar to the 17.5% Manjón-Cabeza, Pose, and Sánchez García (2016) found in Granada Spain, the 17.8% Cerrón-Palomino (2016) found among the Mexican-Americans in Phoenix, the 19.7% Michnowicz (2015) found in Yucatan in Mexico, the 20.9% found in Cameron (1993) study of Madrid, Spain and the 16.8 % and 16.2% Cerrón-Palomino (2014) found in Lima and Huancayo respectively. Moreover, this pronoun rate does not depict a sharp divergence from the 22.4% Essah (forthcoming) found in Equatoguinean Spanish even though these interviews were conducted more than 25 years ago. This similarity between the pronoun rates in Equatoguinean Spanish, and that of Peru and Spain could be explained in terms of the varieties that teachers use in Equatorial Guinea: mainly Northern peninsular.

Cerrón-Palomino (2014) points out that Mexican and Peruvian varieties pattern with northern peninsular varieties, apparently due to the very fluent relationship between the Kingdom of Spain and their two former –and most important- viceroyalties: New Spain (Mexico) and Peru. By the same token, EG Spanish has a northern peninsular base, which would explain its low overt SPP rate. Overall, these findings are consistent with research showing that Spanish tends to have more unexpressed subjects than expressed subject pronouns. The discussion of the significant factor groups and factors follows.

Research Question 2

The second research question asked what linguistic and social constraints favored the expression of SPE in Equatoguinean Spanish. The linguistic factors and the extra-linguistic factors were analyzed separately in Rbrul in order to determine which of the proposed factor groups and the constraints within them significantly conditioned the expression of an overt pronoun in Equatoguinean Spanish. As mentioned before, besides social factor groups, individual speaker was also analyzed as a random factor group. Table 3 and Table 4 below present the results of the linguistic and social factors' analyses, respectively. In the tables, I first state the factor groups and their constraints in the leftmost column, followed by the factor weight, the overt pronoun SPP rate, the number of tokens, the p-value and the range.

A factor weight above 0.5 favors the occurrence of the application value (overt SPP), whereas a factor weight below 0.5 disfavors its occurrence. The concept *favoring* and *disfavoring* refers to the strength of the constraint in predicting the occurrence of the application value. The p-value depicts the significance of the results. A p-value equal to or lower than 0.05 is statistically significant whereas a p-value higher than 0.05 is not statistically significant. The former proves that there is less than 5% probability that the results are random whereas the reverse is true for the latter. With respect to the range, it depicts the relative strength of each statistically significant factor group. The range is calculated by subtracting the highest factor weight from the lowest factor weight within a given factor group. These results are presented in order of decreasing magnitude of effect.

SPE is examined separately, beginning with the strongest group to the weakest.

Person and Number. The factor group *person and number* ranked first ($p < 0$), with a range of 45. In line with previous studies, subjects with a singular referent favor the expression of the overt subject pronouns whereas plural pronouns disfavor it. In this study, the first person singular (I ‘yo’) is the strongest predictor of SPE (a factor weight of 0.68). This trend has been found in several Spanish varieties (e.g., Abreu, 2012; Cameron, 1992; Enríquez 1984; Flores-Ferrán, 2002; Otheguy & Zentella, 2012; Prada Pérez, 2009; Silva-Corvalán 1994; Solomon, 1999). Silva-Corvalán (1994) attributes this tendency to the egocentric nature of verbal communication prompting the expression of more first-person singular subjects.

The next strongest predictor of SPE is the second person singular (you ‘tú’) with a factor weight of 0.64. This is followed by the third person singular (he/she ‘él/ella’) with a factor weight of 0.57. In contrast, there is a disfavoring effect with plural pronouns, as seen in the third person plural (them ‘ellos/ellas’) with a factor weight of 0.38 and the first person plural (we ‘nosotros’) with a factor weight of 0.23. These results are in line with this study’s hypothesis; singular pronouns will favor the expression of the overt subject pronoun and plural pronouns will favor null subjects due to their unambiguous verbal morphology (*-mos* or *-n*).

Ambiguity. The factor group ambiguity was ranked second ($p < 0$), with a range of 33. The expression of an *ambiguous verb form* (a factor weight of 0.66) favored the expression of an overt subject pronoun whereas unambiguous verbs (0.33) disfavored it. This tendency confirms the initial hypothesis proposed in this study

and is in line with previous studies that have found ambiguity favoring the expression of overt subject pronouns (e.g., Cameron, 1993; Lastra & Martín-Butragueño, 2015; Shin, 2014).

Verb class. The factor group verb class ranked third ($p < 0$), with a range of 10. The presence of a *psychological verb* such as *saber* ‘to know’ and *pensar* ‘to think’ favored the expression of the overt pronoun (a factor weight of 0.54) more than any other verb form. The favorable effect of psychological verbs has been found in other studies such as Bentivoglio (1987), Cerrón-Palomino (2016), Enríquez (1984), Silva-Corvalán (1994), and Travis (2007). In this respect, Silva-Corvalán (1994) points out that the contrastive situation in which these types of verbs are found favor overt SPPs. After psychological verbs, the next strong predictor was that of *stative verbs* (a factor weight of 0.53). Examples of such verbs are *ser* ‘to be’, *estar* ‘to be’, and *existir* ‘to exist’. On the other hand, *speech acts* (a factor weight of 0.49) such as *decir* ‘to say’, *motion verbs* (a factor weight of 0.48) such as *venir* ‘to come’ and the group ‘other’ (a factor weight of 0.44) disfavored the expression of the overt subject pronoun.

Reference. The factor group *reference* ranked last ($p < 0$), with a range of 9. A change in reference favored the expression of the overt pronoun (a factor weight of 0.54) whereas the subject pronoun was disfavored when the same reference was maintained (a factor weight of 0.45). These results are in line with the initial hypothesis as well as previous studies that have found bilinguals and L2 speakers to exhibit a decreased sensitivity to switch reference (Michnowicz, 2015; Shin, 2013;

Shin & Otheguy, 2009). However, in all of the aforementioned studies, switch reference was a much stronger predictor of overt SPPs, whereas in Equatoguinean Spanish its effect is notably diminished. I will return to the implications of this finding when addressing Research Question 3.

The results regarding the significant extra-linguistic factors are now presented in table 4 below.

Table 4. Rbrul analysis of extra-linguistic factors favoring the expression of overt SPPs in Equatoguinean Spanish. Speaker as a random factor.

Factor	Factor weights	% Overt Subject Pronoun	#Tokens	p-value	Range (fw)
Education				0.0121	11
High school	0.60	24.2	2005		
Elementary	0.46	14.2	219		
College	0.42	13.9	1755		
Speaker (random) Std	0.44				

Education. The factor group education was the only factor group that was significant among all the proposed social factors (age, sex, L1 and education). This result is in line with the initial hypothesis that education will be a significant factor group. Specifically, speakers with a high school education favored the use of overt subject pronouns (a factor weight of 0.60), whereas speakers with elementary education (a factor weight of 0.46)

and college education (a factor weight of 0.42) disfavored it (more on education is presented in the discussion of significant factor groups below).

Speaker (random). To determine how much variability exists between speakers, individual speaker was accounted for as a random variable and it was found as a significant factor of SPE. In other words, this sample of Equatoguinean speakers is very heterogeneous, even within each one of the social groups analyzed. Moreover, including speaker as a random factor made it possible to recognize which speakers contributed the most or the least to the variation of the overt subject pronoun. In Table 5 below, the overt SPE by speaker is presented in a hierarchical order. It begins with the five speakers that most contributed to the variation of the overt subject, and then the five speakers that least contributed to the variation of the overt Subject Pronoun.

Table 5. Rbrul analysis of Overt SPE by Speakers (p-value = 0.19)

Speaker	factor weight	Tokens	Sex	Age	Education	L1	Other language
Highest SPE rate							
7	0.656	129	male	35	college	Fang	Spanish French
10	0.65	298	male	69	high school	Bube	Spanish French Pichi
11	0.585	199	female	26	college	Bube	Spanish French
29	0.581	152	female	20	college	Fang	Spanish
30	0.58	125	female	20	college	Bube	Spanish
Lowest SPE rate							
4	0.374	99	male	30	high school	Fang	Spanish French English
1	0.395	219	male	64	elementary	Kombe	Spanish French
23	0.398	199	female	26	college	Bube	Spanish, English
9	0.401	123	female	67	elementary	Bube	Spanish

6	0.58	54	female	20	college	Bube	French Spanish
---	------	----	--------	----	---------	------	-------------------

From the above, the speakers contributing most to the attested variation in order of hierarchy are Speaker 7 (factor weight of 0.656), Speaker 10 (factor weight of 0.65), Speaker 11 (factor weight of 0.585), Speaker 29 (factor weight of 0.58) and Speaker 30 (factor weight of 0.58). On the other hand, the five speakers contributing the least towards this variation are Speaker 6 (factor weight of 0.404), Speaker 9 (factor weight of 0.401), Speaker 23 (factor weight of 0.398), Speaker 1 (factor weight of 0.395), and Speaker 4 (factor weight of 0.374). These results show that there is no consistent pattern of SPE within any of the social classes analyzed in this study except for education.

The next section presents the non-significant predictors of SPE in Equatoguinean Spanish.

Non-significant predictors. Four of the proposed factor groups turned out to be not statistically in constraining SPE in Equatoguinean Spanish; *reflexivity* (linguistic factor), *age*, *L1* and *sex* (extra-linguistic factors). In Table 6, the distribution of the non-significant linguistic factor analyzed in this study is provided, followed by the non-significant extra-linguistic factors in table 7.

Table 6. Non-significant linguistic predictor of SPE in Equatoguinean Spanish.

Factor	% Overt Subject Pronoun	#Tokens	Factor weight	Range
Reflexivity				3
Reflexive verbs	15.2%	302	[.48]	
Non-reflexive verbs	19.4%	3677	[.51]	

In table 6, the linguistic factor, *Reflexivity* of the verb used was found not to influence SPE in Equatoguinean Spanish. This is similar to Abreu (2012) who also found reflexivity not significant among the bilingual group, although it was significant among the monolingual and learner group in her study. Thus, bilingual Equatoguinean speakers do not use the identifying information found with reflexive verbs to determine the expression or omission of the subject pronoun.

Table 7. Non-significant extra-linguistic predictors of SPE in Equatoguinean Spanish.

Factor	% Overt Subject Pronoun	#Tokens	Factor weight	Range
L1				20
Fá d'Ambu	34.6%	338	[.59]	
Kombe	23%	462	[.59]	
Fang	17.3%	1234	[.41]	
Bube	16.5%	1945	[.39]	
Sex				3
Male	19.5%	1844	[.51]	
Female	18.8%	2135	[.48]	
Age				
Continuous+1	[ns]			

In table 7, the extra-linguistic factors found not to influence SPE are L1, sex and age. The single speaker with Fá d'Ambu as an L1 expressed the highest rate of SPE followed by speakers with Kombe, Fang, and Bube as L1. However, these results contrast with Essah (forthcoming) who found L1 as a significant predictor of SPE in the Equatoguinean variety of Spanish. In Essah's (forthcoming) study, Fá d'Ambu, Kombe and Fang favored SPE whereas Bube disfavored SPE. A possible explanation for the

mismatch could be the different number of participants, the older speakers who took part in Essah (forthcoming) versus the younger speakers in the present study, or the unequal distribution by sex and age in both studies.

Similarly, the factor group *sex* was deemed not significant in the present study. This factor group was found to be non-significant for SPE expression in other communities such as Mexico City (Lastra & Martín Butragueño, 2015), Yucatan (Michnowicz, 2015), Jalapa (Orozco, 2016), and San Juan in Puerto Rican Spanish (Cameron, 1992).

Finally, age was also found to be non-significant and therefore not affecting SPE in the community studied. This finding has been attested in some previous studies such as Alfaraz (2015), and Michnowicz (2015).

In the next section, the third research question is addressed.

Research Question 3

The third research question asked whether SPE in Equatoguinean Spanish supports or contradicts the universal trends of SPE in bilinguals, as proposed by the IH. As earlier stated (see Chapter 3), the IH has two main predictions regarding SPE. Firstly, it predicts that bilinguals would have a challenge in acquiring syntax and discourse/pragmatics interfaces such as the switch reference constraint, because of the processing cost it entails. Secondly, it states that bilinguals will have high SPE rates in order to ease the cognitive burden of having to keep track of a subject's referent. The first prediction of the IH is supported by this study's results. Equatoguinean speakers do have a decreased sensitivity to switch reference and this is evident by reference being ranked as the lowest significant predictor of SPE. This decreased sensitivity to switch reference

was similarly found in Essah (forthcoming), Shin and Otheguy (2009), Shin (2003), and Toribio (2004) in which bilingual communities were studied.

The second prediction of the IH, of high pronoun rates in bilingual varieties irrespective of the language pair is not supported with this study. An overt SPP rate of 19.1% is an indicator of a very low pronominal use in Equatoguinean Spanish when we compare it with other varieties, as shown in Table 8. It seems clear that the weakening of the referent constraint does not necessarily involve an increase of overt SPPs

The table below presents a comparison of overt SPE's by country.

Table 8. Overt SPE Rates by City/Country. Adapted from Cerrón Palomino (2018)

Location/variety	Author	SPE rate
San Juan, Puerto Rico	Cameron (1993)	44.7%
Santo Domingo	Alfaraz (2015)	42.3%
Dominicans in New York	Otheguy and Zentella (2012)	41%
Puerto Ricans in New York	Otheguy and Zentella (2012)	39%
Isabela, Puerto Rico	Abreu (2012)	39%
Cubans in New York	Otheguy and Zentella (2012)	38%
Santiago, Chile	Cifuentes (1980), quoted by Silva-Corvalán (2001)	38%
Barranquilla, Colombia	Orozco and Guy (2008)	35.7%
Colombians in New York	Otheguy and Zentella (2012)	32%
Buenos Aires	Soares de Silva (2006), quoted by Carvalho and Bessett (2015)	29%
Ecuadorians in New York	Otheguy and Zentella (2012)	28%
Castañer, Puerto Rico	Holmquist (2012)	28%
Mexican-Americans in San Antonio	Bayley et al. (2012)	27%
Chipilo, Mexico	Barnes (2010)	26%
Mexican-Americans in Los Angeles	Silva Corvalán (2001)	25.1%
Rivera, Uruguay	Carvalho and Bessett (2015)	25%
Jalapa, Mexico	Orozco (2016)	24.8%
Puente Genil, Spain	Ranson (1991)	24%
Equatorial Guinea	Essah (forthcoming)	22.4%
Mexicans in New York	Otheguy and Zentella (2012)	22%
Mexico City	Lastra and Martín Butragueño (2015)	21.7%
Castilla, Spain	Rosengren (1974), quoted by	21%

	Holmquist (2012)	
Madrid, Spain	Cameron (1993)	20.9%
Yucatan, Mexico	Michnowicz (2015)	19.7%
Equatorial Guinea	Padilla (2020)	19.1%
Mexican-Americans in Phoenix	Cerrón-Palomino (2016)	17.8%
Granada, Spain	Manjón-Cabeza, Pose, and Sánchez García (2016)	17.5%
Lima, Perú	Cerrón-Palomino (2014)	16.8%
Huancayo, Perú	Cerrón-Palomino (2018)	16.2%

From the above table, it becomes evident that Equatoguinean Spanish exhibits an overt SPP rate even lower than all Latin American varieties except Peru. It is nonetheless higher than that of the Mexican Americans in Phoenix, Granada in Spain, Lima, and Huancayo in Peru.

In the next section, the fourth research question is answered.

Research Question 4

The fourth research question asked what the use of SPE reveals about language contact between Spanish and the local languages that are spoken in Equatorial Guinea. In this study, the lack of significance of L1 factor group suggests that there is no impact of the L1 on SPE in the Spanish of Equatorial Guinea.

Based on the results of prior research carried out by the author (Essah, forthcoming) the initial hypothesis expected L1 to be a significant predictor of SPE, given that, Fang and Bube use verbal prefixes instead of subject pronouns, and Fád'Ambu uses a fixed subject-verb order. In other words, whereas Fang and Bube do not use subject pronouns, Fád'Ambu does use subject pronouns and they are always overt (see Chapter 5). Nonetheless, the results of the present study do not support the initial hypothesis and show no impact of the native languages on SPE in their L2 Spanish. In

other words, the underlying Spanish grammars of the Equatoguinean speakers are comparable irrespective of the native languages they speak.

A possible explanation for this tendency could be the fact that Spanish in Equatorial Guinea is a *lingua franca*, i.e., a means of communication used by people with different native languages (House, 1999; Ostler, 2010). According to Lipski (2004), in Equatorial Guinea, Spanish is acquired in school and it used for a wide number of communicative situations in various settings, e.g., in schools as the medium of instruction, during all official functions, and whenever people of unintelligible native languages need to communicate. In effect, Spanish not only bridges the gap between the numerous native languages the Equatoguinean population speaks, but also, it narrows the possible language contact effects. This could be an explanation for the non-significance of the local languages on SPE expression in their L2 Spanish.

Interestingly, the native languages were significant predictors of SPE in Essah's (forthcoming) study but not significant in the present study. This difference could be based on the location of the studies. The present study took place on the island of Bioko, which is a cosmopolitan region that brings together speakers with different L1's. In contrast, the interviews in Essah's (forthcoming) took place in three different locations in Equatorial Guinea; the island of Bioko, the island of Annobon and the mainland region Rio Muni. According to Quilis (1995), as one moves away from the island of Bioko, there is a decrease in Spanish use, in favor of the local languages or L1's. Thus, the island of Annobon is likely to have more participants using Annobon in their daily activities while the mainland Rio Muni is also likely to have more speakers using Fang than Spanish in their daily activities. This could be the reason why the native languages were

significant predictors of SPE in Essah's (forthcoming) study but not significant in the present study. The following section contains a discussion of the results of the present study.

A Comparison of Pronoun Rates: L2 Speakers and Bilingual Speakers.

The overall overt SPE rate of 19.1% places Equatoguinean Spanish as a variety with one of the lowest SPE rates found in the Hispanic world. It is only higher than Mexican Americans in Phoenix, Granada in Spain and Peru (see table 8 above). A data set of 3980 tokens produced by 30 speakers is fairly reliable to assess the pronoun rate of a speech community. Thus, from these results, it is clear that Equatoguinean speakers do prefer null subject pronouns (81.9%) over overt subject pronouns (19.1%).

Studies that have investigated SPE in L2 speakers', although few, have found similar patterns of low overt SPP rates among advanced L2 speakers. For instance, Abreu (2012) found a rate of 26% among L2 learners (native speakers of English acquiring Spanish as a second language in the classroom) in comparison to the 49% found among the bilinguals of Spanish and English. Similarly, Geeslin and Gudmestad (2008) found that 70% of the forms L2 speakers produced with finite (conjugated) verbs were null pronouns in comparison to 67% of the native speakers produced in the same study. By the same token, Essah's (forthcoming) study of Equatoguinean Spanish found an overt SPP rate of 22.4%, albeit with data over two decades ago. From these studies, it is evident that SPE in Equatoguinean Spanish is not only lower than that of many monolinguals and bilinguals but also, it is even lower than that of other L2 speakers of Spanish.

Two possible explanations could account for this low pronoun rate in Equatoguinean Spanish. The first is the context of acquisition. Like most second-language speakers, the context of acquisition is through formal education. This might suggest that these L2 speakers are using more null than overt pronouns because they have been taught through formal education that in Spanish, overt SPP's are mainly used in cases of referential ambiguity or emphatic/contrastive contexts since this is the information propagated by many educational and curricular materials. In consequence, these L2 speakers use null subjects as a default strategy.

Although L1 turned out to be a non-significant factor group, the second plausible explanation has to do with the native languages, specifically Fang and Bube, which is spoken by the majority of speakers in this study. Except for one speaker (the Fá d'Ambu speaker), Equatoguinean speakers are native speakers of languages generally lacking subject personal pronouns (see Chapter 5). According to Otheguy and Zentella (2012), in bilingual speakers, the absence of relevant constraints in the bilinguals' other languages can reinforce or weaken certain constraints, depending on which language the bilingual uses the most. In the case of Equatorial Guinea, Spanish is used in formal and educative settings whereas the native languages are used in familial settings. Thus speakers who spend time in formal settings get the opportunity to practice the rules that they have been taught regarding the selection of Spanish overt pronouns, whilst speakers who spend the majority of their time in informal settings do not get to practice these rules. In effect, the outcome of the overt SPP rates suggests the latter.

Since overt SPP rates are only one aspect of the mechanisms involved in SPE and they are largely due to a combination of different factors (Travis, 2007), the discussion of the significant factor groups and the constraints within them follow in the next section.

Significant factors in Equatoguinean Spanish SPE

Overall, the findings from this study indicate that for the most part, L2 Equatoguinean bilingual speakers are sensitive to the same constraints that influence SPE in the Spanish of native speakers and other bilingual speakers. The significant factors attested in this study largely are consistent with those of previous studies. Furthermore, these results corroborate that linguistic factors are stronger predictors of overt SPPs than non-linguistic factors in L2 and bilingual speech.

The significant factor groups reported in this study (grammatical person and number, ambiguity, verb type, and reference) are similar to those found in the majority of previous studies. The factor group grammatical person and number has been a consistent significant factor in all SPE studies. Specifically, singular pronouns have been found to condition SPE much more strongly than plural pronouns, and the Equatoguinean variety is no exception. In the present study, the first person singular form *yo* (2328 tokens) was expressed more than double the number of times any other token was expressed. This is similar to previous studies such as Lastra and Butragueño (2015), Orozco (2015), Michnowicz (2015) and several others. The high frequency of verbs in the first person singular could be attributed to the egocentric nature of verbal communication (Silva-Corvalán, 1994; Posio, 2011). Moreover, the majority of the questions asked during the interview were based on the interviewee's life experiences,

wishes, and desires. For this reason, most of the narration was from the speaker's point of view.

Also, in line with previous studies, plural verbs favored the use of null subjects (e.g., Cameron, 1993; Michnowicz, 2015; Orozco & Guy, 2008) in this study. Plural verbs were used mainly when speakers were referring to actions they had done collectively. For instance, punishments received in school, pranks played on teachers, classmates, and family. According to Cameron (1993, p. 306) plural subjects tend to produce null subjects because, "if we conceive of plural subjects as sets, we find that discourse is typically structured so that the great majority of plural subjects occur in contexts where their set members are either explicitly or inferable present within the immediate preceding discourse." Furthermore, because of the unique verbal ending of plural verbs, it is easy for the speaker and interlocutor to identify the referent and this would obviate the need for the expression of the overt subject pronoun.

Regarding ambiguity, overall, ambiguous verb forms favored the occurrence of the overt pronoun while non-ambiguous verb forms did not, in accordance with formal instruction guidelines. In this study, many of the ambiguous verbs were verbs that occurred with the imperfect and conditional tenses. This study thus concurs with studies that have found a similar trend (e.g., Cameron, 1993; Cerrón-Palomino, 2014; De Prada Pérez, 2015; Lastra & Martín-Butragueño, 2015; Shin, 2014).

A possible explanation for the expression of overt pronouns with ambiguous verb forms could be related to the controversial Functional Hypothesis that has found support in some studies (e.g. Hochberg, 1986; Otheguy & Zentella, 2012) but not in others (Cameron, 1996; Ranson, 1991; Samper, 1990). According to this hypothesis, there is a

relationship between the use of overt pronouns and the weakening of inflectional morphemes (Hochberg, 1986). Thus, varieties with a weakened or deleted /s/ consonant tend to have higher overt pronoun rates because the speaker tends to use the overt pronoun to disambiguate between the second person singular originally marked with the suffix *-s* and the third person singular. Evidence in support of this hypothesis was for instance found in Otheguy and Zentella (2012) study. In this study, the researchers found 37 percent of second person singular verbs (*tú*) where the /s/ ending is pronounced and 80 percent of second person singular verbs when the coda /s/ was absent. I argue that similar to the case of overt pronouns appearing in a compensatory function to mitigate the ambiguating effects of final /s/ deletion in verb forms, the expression of overt subject pronouns with ambiguous verb forms in the current study could possibly be related to the intention of the speaker to provide clarifying information regarding the identity of the referent of the verbal subject. This is information that would have been lost due to the identical morphological endings of the first and third person singular verb forms in the conditional, present subjunctive and the imperfect tense.

Concerning verb class, psychological verbs and stative verbs favored the overt subject pronoun whereas motion verbs and speech acts did not. These results go along the lines of what previous studies had found, albeit this factor group in Equatoguinean Spanish was not as strong.

In addition, like previous studies, a switch in reference favored the occurrence of the overt pronoun whereas the maintenance of the same reference disfavored its use despite being the lowest ranked variable in the category of linguistic factor groups. Nonetheless, this finding is not peculiar to the present study. For instance, Michnowicz'

(2015) study of monolinguals and bilinguals in contact with Maya in Yucatan Spanish argued that in comparison to monolinguals, the bilingual speakers demonstrated a decreased sensitivity to reference.

Concerning education, the above results support studies that have found a significant effect for education. The results of the present study depict that the high school educated speakers favor the use of the overt variant whereas the elementary school and college educated disfavor it. This result is partially in line with Ávila-Jimenez (1996) for instance, who found that unskilled workers favored the use of the overt pronoun more than the college-educated. However, why would speakers with high school education use more overt SPPs than speakers with elementary school education? One possibility is that the former ended their formal instruction in Spanish at a time when they were adolescents, and therefore crucially self-centered, uttering the overt first person singular SPP *yo* at a considerably high rate. This lack of continuity in their instruction of Spanish would have caused them to maintain this adolescent-like pattern and thus fail to adjust to more standard uses of SPE

For this explanation to hold, the high school educated speakers should exhibit considerably higher overt *yo* rates than their elementary school and college-educated counterparts, whereas the former's rate of overt *tú, él/ella, nosotros, ellos/ellas* SPPs should not be higher than those of the latter. A closer look at the distribution of the overt first person singular pronoun *yo* by speakers' education in table 9 reveals that this is precisely the case: the highest production of overt *yo* was among the high school educated (63.1%) in contrast to the college educated (32.9%) and those with elementary education (3.9%).

Table 9. Cross-tabulation of Education and person/number overt SPP forms

		Elementary	High school	College	Total
Yo	tokens	23	370	193	586
	%	3.9%	63.1%	32.9%	100
Tu	tokens	2	23	47	72
	%	2.7%	31.9%	65.2%	100
Él/ella	tokens	4	12	31	47
	%	8.5%	25.5%	65.9%	100
Nosotros	tokens	1	8	18	27
	%	3.7%	29.6%	66.6%	100
Ellos/ellas	tokens	1	8	19	28
	%	3.5%	28.5%	67.8%	100

Table 9 also supports my hypothesis that speakers who go beyond a high school education and attain a college education learn to moderate the use of overt SPPs and reduce their rate of *yo* expression up to about a half (63.1% to 32.9%). .

In addition, table 9 shows that high school educated speakers do not exhibit higher rates of overt pronouns *tú, él/ella, nosotros or ellos/ellas* when compared to their elementary school and college educated counterparts, as hypothesized above. All of these facts support my account that high school educated Equatoguinean Spanish speakers only apparently favor overt SPPs as a whole: they really promote the use of overt *yo* only, although at such a high rate that it obscures the effect of the education factor group.

Speaker as a random factor was also analyzed in this study and was found to be significant. Even though the participants represent a small number of the entire Equatoguinean population, the statistical significance of the speaker variable provides a statistical validation for the results obtained from the present study. It shows that the

linguistic factors derived from this study are significant in spite of the variation that exists within speakers.

The findings of this study may be understood as further validation of the universality of linguistic factors that condition SPE among L2 speakers, bilingual speakers and monolingual speakers. The interlanguage of these L2 Equatoguinean speakers is not only systematic but is also influenced by the same constraints that influence Spanish spoken by speakers of other varieties of Spanish, at least with regard to SPE.

The next section offers a discussion of Spanish L2 speakers' SPE and the discourse/pragmatic interface.

L2 Speakers and the Discourse/Pragmatics Interface

The results of this study support the prediction of the IH concerning switch reference. These results show that reference is one area where Equatoguinean L2 speakers differ from monolingual, bilingual and other L2 speakers analyzed in previous studies. As shown by the statistical analysis, Equatoguineans have a decreased sensitivity to a change of reference across verbs.

According to Sorace and Filaci (2006) “narrow syntactic properties are completely acquirable in a second language, even though they may exhibit significant developmental delays, whereas interface properties involving syntax and another cognitive domain may not be fully acquirable” (p. 340). As the selection of reference pronouns is a property that requires the integration of syntactic and discourse/ pragmatic interfaces, in line with the tenets of the IH, it poses a challenge to the Equatoguinean L2

speakers. However, this extra difficulty does not necessarily entail a boost in the overt SPP rate, as this study shows: the rate could also be lowered.

For instance, Pérez-Leroux and Glass (1997) found advanced L2 learners having problems with the syntax-discourse interface in a task where they were supposed to translate a sentence with several referents into Spanish. Interestingly, in these tasks, learners expressed more null subjects than overt subjects, even in cases where a topic shift required the expression of an overt SPP.

Along the same lines, Geeslin and Gudmestad (2008) also found an underproduction of overt pronouns in a topic shift, where an overt pronoun would have been pragmatically needed. In this study, they examined the subject position of finite verbs among native speakers and non-native speakers of Spanish (English speaking learners of Spanish) to find out which forms these speakers produced the most. The researchers found a subtle difference in both groups; the tendency of the non-native speakers to use more null forms than overt pronouns, in comparison to the native speakers. Specifically, the non-native speakers used null pronouns 70% of the time whereas the native speakers, 67.6% of the time. This non-native like acquisition of the discourse/pragmatic constraint has been attested in other languages apart from Spanish. For instance, among English Italian bilingual children (Serratrice, Sorace & Paoli, 2004), and English native speakers acquiring Italian (Sorace & Filiaci, 2006).

The results of the present study prove that although Equatoguinean speakers are sensitive to other factors that constrain SPE in monolingual, bilingual and L2 speech, largely, they have a decreased sensitivity to discourse pragmatic properties that constrain Subject Pronoun Expression in Spanish.

Summary

The goal of this chapter was to present the results obtained by running multiple regression analyses with the statistical program Rbrul. To answer the first research question, the overt SPP rate in Equatoguinean Spanish is 19.1%, an indication of a low pronoun variety. With respect to the second research question, the significant linguistic factors influencing SPE in Equatoguinean Spanish, in order of magnitude are *person and number, ambiguity, verb class, and reference*. Two extra-linguistic factor groups were significant: education and the random variable of individual speaker. With respect to research question three, on the one hand, the tenets of the IH hypothesis regarding problems with switch reference hold in this study. The switch reference constraint in these speakers is weakened when compared to studies in other monolingual and bilingual speech communities. As depicted by the results, reference is the least favoring significant linguistic factor group. On the other hand, the low pronoun rate attested in this data goes against the prediction of the IH, which states that the weakening of the switch reference constraint would boost the rate of overt SPPs among bilinguals. Finally, with respect to the fourth research question, contact with the local languages does not seem to influence the underlying grammars of the speakers, with respect to SPE. This is despite the fact that Fang and Bube to a large extent do not use subject pronouns and Fá d'Ambu uses a fixed subject verb order. The possible explanation provided by the study is the role of Spanish as a *lingua franca* in Equatorial Guinea. The next chapter concludes the dissertation.

CHAPTER 7

CONCLUSION

The present study has examined SPE in the Equatoguinean variety of Spanish. To the best of my knowledge, it is the only large scale variationist study to examine SPE in this variety of Spanish in the last 25 years. The examination of SPE in this variety not only brings to light an understudied bilingual/multilingual community but also, contributes to several subfields in linguistics, specifically, to the field of language variation, language contact studies, and second language acquisition studies.

This chapter is divided into four sections. In the first section, I elaborate on the contributions of the study. Secondly, I discuss the limitations of the study. Thirdly, I provide directions for future research. The last section summarizes and concludes the dissertation.

Contribution to the Field of Language Variation Research

The contribution of this study to the field of language variation cannot be underestimated. To date, very few studies have examined SPE in Spanish among second-language speakers using the variationist approach (Abreu, 2012; Geeslin & Gudmestad, 2008, 2010, 2011; Gudmestad & Geeslin, 2010; Gudmestad, House & Geeslin, 2013). Most studies involving SPE in L2 Spanish have looked at formal factors such as the Null Subject Parameter (Montrul, 2004; Sorace, 2004) rather than data-based patterns that condition the expression or omission of overt SPPs. The use of sociolinguistics methods not only shows the variability in L2 speaker's speech but also leads to an understanding of the constraints that govern SPE.

In addition, these results provide additional evidence of the importance of linguistic factors in influencing SPE, in comparison with extra-linguistic factors. The literature review revealed that pronoun rates differ among varieties and that the influencing factors have been *person/number*, *reference*, *ambiguity*, and *verb type*. These results have been quite uniform across studies, and the results of this study are no exception.

Moreover, the few SPE studies that have used sociolinguistic methods to study SPE in L2 varieties have been conducted by comparing the overt SPP rates and constraint rankings of L2 and native speakers (e.g., Abreu, 2012; Geeslin & Gudmestad, 2008). The present study bridges this gap by investigating a context where there are hardly native Spanish speakers to use as a comparative model, but instead the focus is centered on Spanish L2 speakers who are bilingual/multilingual, and native speakers of different African languages.

Additionally, variationist studies investigating SPE have tended to use ANOVA, correlations, cross-tabulations, and fixed models regression analyses such as GoldVarb in analyzing which linguistic and extra-linguistic factors account for the expression of the overt pronoun. However, in this study, I used the mixed-effects model software Rbrul because it offers several advantages over other software. For instance, Rbrul can account for both fixed effects and random effects in the same analysis. Individual speaker was thus accounted for in this study as a random effect. By accounting for the speaker, this study was able to avoid the overestimation of social variables usually caused by the characteristics of a few speakers skewing the data in a particular direction (Johnson,

2009). In addition, Rbrul allows for the analysis of continuous and not only discrete variables, which allows one to analyze variables like age in a natural fashion.

Finally, by conducting this study, more light has been shed on the Spanish variety of Equatorial Guinea- the only African country with Spanish as an official language.

Contribution to the Field of Language Contact Research

The contribution of this study can also be seen in the fields of language contact, as Equatoguinean Spanish is a variety in contact with various African languages, Indo-European languages, and creoles. Thus, this study makes an original contribution to the field of language contact research, as well as provides support for predictions, hypothesis, and conclusions derived from previous sociolinguistic studies regarding Spanish contact and subject pronoun expression. For instance, a possible explanation provided in this study for the non-significance of L1 is the *lingua franca* status of Spanish in Equatorial Guinea.

On a larger scale, the material gathered on SPE in Equatorial Guinea adds to the knowledge of linguistic variation in Subject Pronoun Expression and the use of Spanish in a stable bilingual African Spanish-speaking society. This study could not have been conducted in any other society except Equatorial Guinea, a country where the African L1's predominantly do not use subject personal pronouns.

Contribution to Second Language Acquisition Studies

Equatorial Guinea was an ideal scenario for testing SLA hypotheses such as the IH. Since the IH's predictions are expected to hold true regardless of the SPE characteristics of the L1, Equatorial Guinea was suitable for two reasons: a) Spanish is

spoken there as an L2-only variety and b) the African L1s predominantly lack the use of subject personal pronouns.

The results of this study show that the Equatoguinean variety is a variety with a low subject pronoun expression rate in spite of the prediction of the IH of high pronoun rates. Although the rates of expression do not tell much about the variability of their expression, by examining the constraints and constraint rankings, one finds that these L2 speakers, largely exhibit native-like characteristics regarding the choice of the overt subject pronoun.

Also, the investigation of SPE among these Equatoguinean speakers brings to light an area where these L2 speakers differ from monolinguals and other second language and bilingual speakers: reference. The fact that Equatoguinean speakers show a decreased sensitivity to the reference constraint supports the IH prediction that discourse/pragmatic constraints are difficult for second language and bilingual speakers to fully acquire.

Limitations of the study

This study is not free from limitations. The main drawback is the unbalanced sample size for the social factor groups; sex, L1, and education. For instance, with respect to sex, there were 18 males and 12 females. Concerning the native languages, there were 12 native speakers of Fang, 14 of Bube, three of Kombe, and only one of Fá d'Ambu. With respect to education, more than half of the speakers had a college education; two speakers with elementary education, eight with high school education and 20 with a college education. It is possible that with a much more balanced sample, some of these factor groups would have shown different trends and patterns.

Furthermore, the interviews were conducted in the capital (Malabo), where there was access to speakers with different native languages who had relocated to Malabo. This frequent contact with speakers of different native languages may have led to dialectal leveling and thus may be a factor as to why the native language was not a predictor of SPE.

A major challenge, however, in this study has been the lack of information regarding SPE rates in the local languages that served as L1s to compare it to Spanish. At the moment, only descriptive accounts are provided because of scarce material focusing on SPE in these local languages.

Another possible limitation is the *Hawthorne Effect* or *Observers paradox*. Participants are likely to modify their speech because they know they are being recorded (Brown, 1988). I recognize that this "...should not be regarded as a defect of observation, but rather as a characteristic intrinsic to it" (Gobo, 2008, p. 10). It is because of this weakness that the interviews were conducted for at least 40 minutes (Tagliamonte, 2009). Nonetheless, it was still obvious that some participants were more relaxed than others were, and this may have affected their production.

Future Research

Equatorial Guinea is a country that exhibits linguistic diversity that may be of interest to researchers in the field of bilingualism and second language acquisition studies. In this study, participants' spoke Fang, Bube, Kombe and one speaker spoke Fá d'Ambu. Future research should expand the sample size and include comparable sets of speakers with different native languages. Other languages spoken such as Balengue,

Ndowe, Pichi English, and French may provide more insight into the contact situation in Equatorial Guinea.

Another possibility for a future study could be to focus on speakers of Fang as an L1 since it is the most dominant local language. Moreover, considering that Equatoguinean Spanish is an L2, a future study should look at learners in Equatorial Guinea at different stages of development (as was done in Geeslin and Gudmestad, 2008).

Moreover, future research should analyze priming as a linguistic factor group, as this has been found to be significant in previous studies in which it was included. Finally, the present study mentioned the possibility of Spanish being a *lingua franca* as an explanation for the insignificance of native languages as a predictor of SPE. This assumption might be addressed in future studies. These studies could investigate SPE in other (more rural) parts of Equatorial where regional varieties have not undergone dialect leveling as they had done in Malabo. Specifically, the island of Annobon and the mainland region where Fa d’Ambu and Fang respectively are spoken would provide this type of data.

Summary

This dissertation focused on answering four research questions:

- (a) What is the null and overt pronoun rate in Equatoguinean Spanish? How does this rate compare with pronoun rates in other varieties of Spanish?
- (b) What linguistic and social constraints favor the use of SPE in Equatoguinean Spanish?

(c) Does SPE in Equatoguinean Spanish support or contradict universal trends of SPE in bilinguals, as proposed by the Interface Hypothesis?

(d) What does the use of SPE reveal about language contact between Spanish and the local languages that are spoken in Equatorial Guinea?

To address these questions, a sociolinguistic study was conducted with 30 Equatoguinean Spanish speakers living in Malabo. These speakers were comprised of 18 males and 12 females. The youngest participant who took part in the study was 18 years of age and the oldest was 64 years of age. These speakers took part in sociolinguistic interviews that lasted between 40 minutes and one hour during which they narrated important life events, stories and anything else they felt comfortable sharing. The recorded interviews were later transcribed and analyzed.

The envelope of variation followed the Principle of Accountability (Labov, 1972). Thus, all tensed forms with human reference were included. However, all cases in which the null pronoun failed to alternate with the overt pronoun were excluded. Thus, verbs with subjects that tend to be null phrases, emphatic *mismo*, discourse markers, and imperatives were excluded. Following the line of previous studies, the linguistic factor groups were *person and number*, *reflexivity*, *verb type*, and *ambiguity*. The extra-linguistic factors were *age*, *sex*, *L1* and *education*. There were 3980 tokens submitted to Rbrul for the analysis.

To answer the first research question, the pronoun rate of Equatorial Guinea was 19.1%. This places Equatorial Guinea as a variety with one of the lowest SPE. To answer the second research question, the results of the quantitative analysis revealed that in order of magnitude, person and number, ambiguity, verb class and reference were significant

linguistic factor groups that condition SPE in Equatorial Guinea. Education and individual speaker were the only significant social factor group. Concerning the third research question, the results of this study support the tenets of the IH in one respect, but refute them in another. Reference is the lowest ranked significant predictor of SPE indicating that these L2 speakers do have a decreased sensitivity to the discourse/pragmatic constraint, as the IH postulated (Sorace, 2011). Nonetheless, the low overt SPP rate attested in this study contrasts with the overproduction of overt SPPs the IH expected. Concerning the fourth research question, the results suggest that the native languages do not influence SPE expression in Spanish even though Fang and Bube use verbal prefixes instead of subject pronouns and Fá d'Ambu uses a fixed subject verb order. The study proposed the position of Spanish as *lingua franca* as a way of explaining the non-significance of the local languages.

Although it still leaves unanswered questions relating to bilingualism and SPE in Equatorial Guinea, the present study nonetheless reveals that SPE cannot be explained or defined by one line of evidence alone. As long as language continues to evolve, language contact continues to exist, and society continues to change, SPE will continue to be an interesting field of study in Spanish sociolinguistics.

REFERENCES

- Abreu, L. (2009). Spanish subject personal pronoun use by monolinguals, bilinguals and second language learners. (Unpublished doctoral dissertation). University of Florida, Gainesville FL.
- Abreu, L. (2012). Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish. In K. Geeslin, & M. Díaz-Campos (Eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium* (pp. 1-8). Somerville, MA: Cascadilla Proceedings Project.
- Alba, O. (2004). *Cómo hablamos los dominicanos. Un Enfoque Sociolingüístico*. Santo Domingo: Grupo León Jiménez.
- Alfaraz, G. (2015). Variation of overt and null subject pronouns in the Spanish of Santo Domingo. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 3-16). Washington DC: Georgetown University Press.
- Argyri, E., & Sorace, A. (2007). Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*, 10(1), 79-99.
- Ávila-Jiménez, B. (1995). A sociolinguistic analysis of a change in progress: Pronominal overtness in Puerto Rican Spanish. *Cornell Working Papers in Linguistics*, 13, 25-48.
- Avila-Jimenez, B. (1996). *Subject pronoun expression in Puerto Rican Spanish: A sociolinguistic, morphological, and discourse analysis* (Order No. 9608242). Available from ProQuest Dissertations & Theses Global. (304287797). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/304287797?accountid=4485>
- Barbosa, P., Duarte, M., & Kato, M. (2005). Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4, 11-52.
- Barnes, H. (2010). Subject pronoun expression in bilinguals of two null subject languages. In K. Arregi, Z. Fagyal, S. Montrul & A. Tremblay (Eds.), *Selected Papers from the 38th Linguistic Symposium on Romance Languages (LSRL)* (pp. 9-22). Amsterdam: John Benjamins.
- Barrenechea, A. M., & Alonso, A. (1977). Los pronombres personales sujetos en el español hablado en Buenos Aires. In Blanch (Ed.), *Estudios sobre el español hablado en las principales ciudades de América* (pp. 333-349). Mexico City: Universidad Nacional Autónoma de México.
- Bayley, R. & Pease-Álvarez, L. (1997). Null pronoun variation in Mexican-descent children's narrative discourse. *Language variation and change*, 9(3), 349-371.

- Bayley, R., Cárdenas, N. L., Schouten, B. T., & Salas, C. M. V. (2012). Spanish dialect contact in San Antonio, Texas: An exploratory study. In K. Geeslin, & M. Díaz-Campos (Eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium* (pp. 48–60). Somerville, Massachusetts: Cascadilla Proceedings Project.
- Bayley, R., & Tarone, E. (2012). Variationist approaches to second language acquisition. In P. Robinson (Ed.) *Routledge Encyclopedia of Second Language Acquisition*. (pp. 679-682). New York: Routledge.
- Belletti, A., Bennati, E., & Sorace, A. (2007). Theoretical and developmental issues in the syntax of subjects: evidence from near-native Italian. *Natural Language and Linguistic Theory*, 2, 657-68.
- Bentivoglio, P. (1980). *Why canto and not yo canto? The problem of first-person subject pronoun in spoken Venezuela Spanish*. (Unpublished master's thesis). University of California, Los Angeles.
- Bentivoglio, P. (1987). *Los sujetos pronominales de primera persona en el habla de Caracas*. Caracas: Universidad Central de Venezuela.
- Besari, T. M. (2005). Breve apunte sobre el español en Guinea Ecuatorial. *El Español en el Mundo: anuario del Instituto Cervantes 2005*, 255-262. Instituto Cervantes. Retrieved from http://cvc.cervantes.es/lengua/anuario/anuario_05/morgades/p01.htm
- Bessett, R. M. (2018). Testing English influence on first person singular “yo” subject pronoun expression in Sonoran Spanish. *Contemporary trends in Hispanic and Lusophone linguistics*, 15, 355-372.
- Bibang Oyee, J. (1990). *Curso de lengua Fang*. Malabo: Centro Cultural Hispano-Guineano.
- Bibang Oyee, J. B. (2002). *El español guineano: interferencias, guineanismo*. Malabo, Centro Cultural Hispanoguineano.
- Bolekia Boleká, J. (1991). *Curso de lengua bubu*. Malabo: Centro Cultural Hispanoguineano.
- Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14(2), 136-193.
- Bullock, B. E., & Toribio, A. J. (2004). Introduction: Convergence as an emergent property in bilingual speech. *Bilingualism: Language and Cognition*, 7(2), 91-93.
- Cameron, R. (1992). *Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation* (Order No. 9227631). Available from ProQuest Dissertations & Theses Global. (303996702). Retrieved from

- <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/303996702?accountid=4485>
- 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. (1994). Switch reference, verb class and priming in a variable syntax. *Papers from the Regional Meeting of the Chicago Linguistics Society: Parasession on variation in linguistics theory*, 30, 27-45.
- Cameron, R. (1995). The scope and limits of switch reference as a constraint on pronominal subject. *Hispanic Linguistics*, 6(7), 1-27.
- Cameron, R., & Flores-Ferrán, N. (2004). Perseveration of subject expression across regional dialects of Spanish. *Spanish in Context*, 1(1), 41-65.
- Campbell, L. (2004). *Historical linguistics: An introduction*. Cambridge, MA: MIT Press.
- Campos, A. (2003). The decolonization of Equatorial Guinea: the relevance of the international factor. *The Journal of African History*, 44(1), 95-116.
- Carminati, M. N. (2002). *The processing of Italian subject pronouns* (Order No. 3039345). Available from ProQuest Dissertations & Theses Global. (288028052). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/288028052?accountid=4485>
- Carvalho, A. M. (2006). Políticas lingüísticas de séculos passados nos dias de hoje: O dilema sobre a educação bilíngüe no norte do Uruguai. *Language Problems and Language Planning*, 30(2), 149-171.
- Carvalho, A. M., & Child, M. (2011). Subject pronoun expression in a variety of Spanish in contact with Portuguese. In J. Michnowicz & R. Dodsworth (Eds.), *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* (pp.14-25). Somerville, MA: Cascadilla Press.
- Carvalho, A. M., & Bessett, R. M. (2015). Subject pronoun expression in Spanish in contact with Portuguese. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 143-163). Washington DC: Georgetown University Press.
- Carvalho, A. M., Orozco, R., & Shin, N. L. (2015). *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown University Press.
- Castillo Barril, M. (1964). El español en la Guinea Ecuatorial. *Español actual*, 3, 8-9.

- Cedergren, H. J., & Sankoff, D. (1974). Variable rules: Performance as a statistical reflection of competence. *Language*, 333-355.
- Cerrón-Palomino, Á. (2014). Ser más pro o menos pro: variación en la expresión de sujeto pronominal en el castellano limeño. *Lingüística*, 30(1), 61-83.
- Cerrón-Palomino, Á. (2016). (No) English interference on US southwest Spanish?: A look at variable subject expression in Phoenix Spanish-English bilinguals. *Sociolinguistic Studies*, 10(3), 383-408.
- Cerrón-Palomino, Á. (2018). Null-subject encounter: Variable subject pronoun expression in the Spanish of Quechua-Spanish bilinguals in the Central Peruvian Andes. *International Journal of Bilingualism*, 1-19, doi: org.ezproxy1.lib.asu.edu/10.1177%2F1367006918763175
- Cifuentes, H. (1980). Presencia y ausencia del pronombre personal sujeto en el habla culta de Santiago de Chile. *Boletín de Filología*, 31(2), 743-752.
- Claes, J. (2011). ¿Constituyen las Antillas y el Caribe continental una sola zona dialectal?: Datos de la variable expresión del sujeto pronominal en San Juan de Puerto Rico y Barranquilla, Colombia. *Spanish in context*, 8(2), 191-212.
- Chambers, J. K. (2003). *Sociolinguistic theory: linguistic variation and its social significance*. Malden and Oxford: Blackwell Publishers.
- Chirilă, E. M. (2015). *Identidad lingüística en Guinea Ecuatorial: diglosia y actitudes lingüísticas ante el español* (Unpublished Master's thesis). University of Bergen, Bergen, Norway.
- Datos del Censo de Población de Guinea Ecuatorial. (2015, September). Guinea Infomarket. Retrieved from <https://www.guineainfomarket.com/datos-del-censo-de-poblacion-de-guinea-ecuatorial/>
- De Granda, G. (1985). Un caso de transferencia léxica intercolonial: Cuba-Fernando Poo (Bioko). *Anuario de Letras. Lingüística y Filología*, 23, 131-159.
- De Granda, G. (1990). El español de Guinea Ecuatorial. Sobre un fenómeno sintáctico: la marcación en superficie de los pronombres personales sujeto. *Thesaurus-Instituto Caro y Cuervo*, 45(2), 332-354.
- Duarte, M. (1995). *A Perda do Princípio "Evite pronome" no Português Brasileiro*. (Unpublished Doctoral dissertation). University of Campinas, São Paulo, Brazil.
- Echegaray, C. G. (1951). Notas sobre el español en África Ecuatorial. *Revista de Filología Española*, 35, 106-118.

- Echegaray, C. G. (1959). *La lengua baseque (Guinea Española) (Vol. 1)*. Madrid: Consejo superior de investigaciones científicas.
- Ellis, R. (1985). Sources of variability in interlanguage. *Applied linguistics*, 6(2), 118-131.
- Enríquez, E. V. (1984). *El pronombre personal sujeto en la lengua española hablada en Madrid* (Vol. 4). Editorial CSIC-CSIC Press.
- Erker, D., & Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*, 526-557.
- Essah, L. (accepted). Subject Pronoun Expression in Equatoguinean Spanish. *Divergencias: Revista de estudios lingüísticos y literarios*.
- Flege, J. E., MacKay, I. R., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, 23(4), 567-598.
- Flores-Ferrán, N. (2004). Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? *Language variation and change*, 16(1), 49-73.
- Flores-Ferrán, N. (2007). A bend in the road: Subject personal pronoun expression in Spanish after 30 years of sociolinguistic research. *Language and Linguistics Compass*, 1(6), 624-652.
- Frynas, J. G. (2004). The oil boom in Equatorial Guinea. *African Affairs*, 103(413), 527-546.
- Geeslin, K., & Gudmestad, A. (2008). Variable subject expression in second-language Spanish: A comparison of native and non-native speakers. In M. Bowles, R. Foote, S. Perpiñán, & R. Bhatt (Eds.), *Selected Proceedings of the 2007 Second Language Research Forum* (pp. 69-85). MA: Cascadilla Proceedings Project.
- Geeslin, K., & Gudmestad, A. (2010). An exploration of the range and frequency of occurrence of forms in potentially variable structures in second-language Spanish. *Studies in Second Language Acquisition*, 32(3), 433-463.
- Geeslin, K., & Gudmestad, A. (2011). Using sociolinguistic analyses of discourse-level features to expand research on L2 variation in forms of Spanish subject expression. In *Selected proceedings of the 2009 Second Language Research Forum* (pp. 16-30). Somerville, MA: Cascadilla Proceedings Project.
- Geeslin, K., & Gudmestad, A. (2016). Subject expression in Spanish: Contrasts between native and non-native speakers for first and second-person singular referents. *Spanish in Context*, 13(1), 53-79.

- Geeslin, K., Linford, B., & Fafulas, S. (2015). Variable subject expression in second language Spanish. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 191-209). Washington DC: Georgetown University Press.
- Gobo, G. (2008). *Doing ethnography*. Los Angeles, CA: Sage Publications.
- Gomashie, G. A. (2019). Language Vitality of Spanish in Equatorial Guinea: Language Use and Attitudes. *Humanities*, 8(1), 33.
- Granda Gutiérrez, G. (1984). Perfil lingüístico de Guinea Ecuatorial. In L. Flórez (Ed.), *Homenaje a Luis Flórez: estudios de historia cultural, dialectología, geografía lingüística, sociolingüística, fonética, gramática y lexicografía* (pp. 119-195). Bogotá: Instituto Caro y Cuervo.
- Granda Gutiérrez, G. (1988). El Español en el África Subsahariana. *África 2000. Revista de Cultura*, 7 (4), 15.
- Granda Gutiérrez, G. D. (1990). El español de Guinea Ecuatorial. Sobre un fenómeno sintáctico: la marcación en superficie de los pronombres personales sujeto. *Thesaurus: boletín del Instituto Caro y Cuervo*, 45(2), 332-354.
- Gudmestad, A., & Geeslin, K. L. (2010). Exploring the roles of redundancy and ambiguity in variable subject expression: A comparison of native and non-native speakers. In C. Borgonovo, M. Español-Echevarría, and P. Prévost (Eds.), *Selected proceedings of the 12th Hispanic linguistics symposium* (pp. 270-283). Somerville, MA: Cascadilla Proceedings Project.
- Gudmestad, A., House, L., & Geeslin, K. L. (2013). What a Bayesian analysis can do for SLA: New tools for the sociolinguistic study of subject expression in L2 Spanish. *Language Learning*, 63(3), 371-399.
- Gumperz, J. J. (1982). *Discourse strategies* (Vol. 1). Cambridge University Press.
- Gutiérrez, M. J. (1994). Simplification, transfer, and convergence in Chicano Spanish. *Bilingual Review/La Revista Bilingüe*, 19(2), 111-121.
- Harvie, D. (1998). Null subject in English: Wonder if it exists. *Cahiers Linguistiques D Ottawa*, 15-26.
- Hochberg, J. G. (1986). Functional compensation for/s/deletion in Puerto Rican Spanish. *Language*, 62(3), 609-621.
- Hoffman, M. F., and Walker, J. A. (2010) Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, 22(1), 3767. Doi: <https://doi.org/10.1017/S0954394509990238>. 220.

- Holmquist, J. (2012). Frequency rates and constraints on subject personal pronoun expression: Findings from the Puerto Rican highlands. *Language Variation and Change*, 24(2), 203.
- House, J. (1999). Misunderstanding in intercultural communication: Interactions in English as a lingua franca and the myth of mutual intelligibility. In Gnutzmann Jenkins, J. (Ed.) *Which pronunciation norms and models for English as an International Language?* (pp.73-89). Tübingen: Stauffenburg.
- Ivanov, I. P. (2009). *Second language acquisition of bulgarian object clitics: A test case for the interface hypothesis* (Order No. 3368912). Available from ProQuest Dissertations & Theses Global. (304900096). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/304900096?accountid=4485>
- Klein, W. (1986). *Second language acquisition*. Cambridge, England: Cambridge University Press.
- Jia, L., & Bayley, R. (2002). Null pronoun variation in Mandarin Chinese. *University of Pennsylvania Working Papers in Linguistics*, 8(3), 9.
- Johnson, D. E. (2008). Rbrul. Retrieved March 6, 2020 from <http://www.ling.upenn.edu/~johnson4/Rbrul.R>
- Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and linguistics compass*, 3(1), 359-383.
- Klintonberg, R. (1979). *Equatorial Guinea-Macias Country: The Forgotten Refugees; An International University Exchange Fund (IUEF) Field Study on the Equatorial Guinea Refugee Situation*. International Univ. Exchange Fund.
- Koban, D. (2010). *Overt versus null subject pronoun variation in the turkish spoken in turkey and in new york city* (Order No. 3481504). Available from ProQuest Dissertations & Theses Global; Sociological Abstracts. (906496603). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/906496603?accountid=4485>
- Kraus, J. R. (2010). *The "business" of state-building: The impact of corporate social responsibility on state development in Equatorial Guinea* (Order No. 3446952). Available from ProQuest Dissertations & Theses Global. (856604806). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/856604806?accountid=4485>
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19(3), 273-309.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

- Labov, W. (1969). Contraction, deletion and inherent variability of the English copula. *Language* 45, 715-762.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (1973). The social setting of linguistic change. *Diachronic, Areal, and Typological Linguistics*, 195-249.
- Labov, W. (1982). Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor. *Language in society*, 11(2), 165-201.
- Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh and J. Sherzer (Eds.), *Language in use: Readings in Sociolinguistics* (pp. 28–53). Englewood Cliffs, NJ: Prentice Hall.
- Labov, W. (2001). *Principles of linguistic change Volume 2: Social factors*. Malden and Oxford: Blackwell Publishers.
- Labov, W. (2006). *The social stratification of English in New York City*. New York: Cambridge University Press.
- Lafond, L., Hayes, R., & Bhatt, R. M. (2001). Constraint Demotion and Null-Subjects in Spanish L2 Acquisition. Romance Syntax, Semantics and L2 Acquisition. In J. Camps and C. Wiltshire (Eds.), *Selected Papers from the 30th Linguistic Symposium on Romance Languages* (pp. 121-135). Amsterdam: John Benjamins.
- Lapidus, N. & Otheguy, R. (2005). Contact induced change? Overt non-specific ellos in Spanish in New York. In L. Sayahi & M. Westmoreland (Eds.), *Selected Proceedings of the 2nd Workshop on Spanish Sociolinguistics* (pp. 67-75). Somerville, MA: Cascadilla Press.
- Lastra, Y., & Butragueño, P. M. (2015). Subject pronoun expression in oral Mexican Spanish. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 39-57). Washington DC: Georgetown University Press.
- Lipski, J. M. (1984). Observations on the Spanish of Malabo, Equatorial Guinea: Implications for Latin American Spanish. *Hispanic Linguistics*, 1(1), 69-96.
- Lipski, J. M. (1985). *The Spanish of Equatorial Guinea*. Tübingen: Max Niemeyer.
- Lipski, J. M. (1985). Literary ‘Africanized’ Spanish as a research tool: Dating consonant reduction. *Romance Philology* 49, (2), 130-167.

- Lipski, J. M. (1999). Creole-to-creole contacts in the Spanish Caribbean: the genesis of Afro Hispanic language. *Publications of the Afro-Latin American Research Association (PALARA)*, 3, 5-46.
- Lipski, J. M. (2000). The Spanish of Equatorial Guinea: research on la hispanidad's best-kept secret. *Afro-Hispanic Review*, 11-38.
- Lipski, J. M. (2002). The Spanish of Equatorial Guinea 1: research on la hispanidad's best-kept secret. *Afro-Hispanic Review*, 70-97.
- Lipski, J. M. (2004). The Spanish language of Equatorial Guinea. *Arizona Journal of Hispanic Cultural Studies*, 8(1), 115-130.
- Lipski, J. M. (2008). El español de Guinea Ecuatorial en el contexto del español mundial. In G. Rosique & G. Jahn (Eds.), *La situación actual del español en África* (pp.79-117). Madrid: Casa de África/SIAL.
- Lipski, J. M. (2014). ¿Existe un dialecto “ecuatoguineano” del español? *Revista iberoamericana*, (248-9), 865-882.
- Long, M. (2003). Stabilization and fossilization in interlanguage development. In C. Doughty, & M. Long (Eds.), *The handbook of second language acquisition* (pp. 487-535). Oxford: Blackwell.
- Lozano, C. (2002). Knowledge of expletive and pronominal subjects by learners of Spanish. *ITL-International Journal of Applied Linguistics*, 135(1), 37-60.
- Lozano, C. (2009). Selective deficits at the syntax-discourse interface. In Y. Leung, N. Snape, M. Sharwood-Smith (Eds.), *Representational deficits in SLA: Studies in honor of Roger Hawkins*, (pp. 127-166). Amsterdam: John Benjamins.
- Lott, D. (1983). Analyzing and counteracting interference errors. *ELT Journal*, 37(3), 256-261.
- Manjón-Cabeza Cruz, A., Pose Furest, F., & Sánchez García, F. J. (2016). Factores determinantes en la expresión del sujeto pronominal en el corpus PRESEEA de Granada. *Boletín de filología*, 51(2), 181-207.
- Manso Luengo, A. J., and Bibang Oyee J. (2014). El español en Guinea Ecuatorial. In Javier Serrano Avilés (Ed.), *La enseñanza del español en África Subsahariana* (pp.310-322) Madrid: Catarata. Retrieved from <https://cvc.cervantes.es/lengua/eeas/capitulo13.htm>
- Martinez-Sanz, C. (2011). *Null and overt subjects in a variable system: The case of dominican spanish* (Order No. NR98031). Available from Linguistics and Language Behavior Abstracts (LLBA); ProQuest Dissertations & Theses Global. (1356845002). Retrieved from

<http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/1356845002?accountid=4485>

- Mayol, L. (2009). *Pronouns in Catalan: information, discourse and strategy*. (Unpublished doctoral dissertation). University of Pennsylvania, USA.
- Michnowicz, J. (2015). Subject pronoun expression in contact with Maya in Yucatan Spanish. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 101-119). Washington DC: Georgetown University Press.
- Migeod, F. H. (1971). *The Languages of West Africa*. London: Kegan Paul, Trench, Trübner & Co., Ltd.
- Miró Vera, R., & De Pineda, A. (1982). Determinación sociolingüística de la presencia/ausencia del pronombre personal sujeto. *Sociolingüística andaluza*, 5, 37-44.
- Mohamadou, A. (2008). Acercamiento al «espaguifranlés», el español funcional de Guinea Ecuatorial. *Cauce*, 31, 213-229.
- Morales, A. (1980). La expresión de sujeto pronominal, primera persona, en el español de Puerto Rico. *Boletín de la Academia Puertorriqueña de la Lengua Española*, 8(3), 91-102.
- Morales, A. (1986). *Gramáticas en contacto: Análisis sintáctico sobre el español de Puerto Rico*. San Juan: Editorial Playor.
- Morales, A. (1989). Hacia un universal sintáctico del español del caribe. El orden SVO. *Anuario de Lingüística Hispánica* 5, 139-152.
- Morales, A. (1997). La hipótesis funcional y la aparición de sujeto no nominal: el español de Puerto Rico. *Hispania*, 80 (1), 153-165.
- Montes Miró, R. (1986). Factores discursivos en el análisis de los pronombres personales sujeto en español. *Morphé* (1).45-71.
- Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and cognition*, 7(2), 125-142.
- Nguen, D. (2009). *La creación semántica y léxica en el español de Guinea Ecuatorial* (Unpublished doctoral dissertation). Universidad Complutense de Madrid, Spain.

- Nistal, G. (2007). El caso del español en Guinea Ecuatorial. *Instituto Cervantes de Cervantes*, 73–76. Retrieved from https://cvc.cervantes.es/lengua/anuario/anuario_06-07/pdf/paises_08.pdf
- Nistal, G. (2009). Peculiaridades del español de Guinea Ecuatorial y los préstamos del español en las lenguas indígenas. In Stolz, Thomas, Dik Bakker & Rosa Salas Palomo (Eds.), *Romanisierung in Afrika. Der Einfluss des Französischen, Portugiesischen und Spanischen auf die indigenen Sprachen Afrikas* (pp.1-15). Universitätsverlag Dr. N. Brockmeyer.
- Nsue Otong, C. (1986). Guineanismos o español de Guinea Ecuatorial. *Muntu. Revue Scientifique et Culturelle du CICIBA*, (4-5), 265-268.
- Ocampo, F. (1990). El subjuntivo en tres generaciones de hablantes bilingües. In JJ. Bergen (Ed.), *Spanish in the United States: Sociolinguistic Issues* (pp. 39-48). Washington D.C.: Georgetown University Press.
- Ortiz López, L. (2011). Spanish in contact with Haitian Creole. In Díaz-Campos, Manuel (Ed.), *The handbook of Spanish sociolinguistics* (pp. 418–445). Wiley: Blackwell.
- Orozco, R. (2007). The impact of linguistic constraints on the expression of futurity in the Spanish of New York City Colombians. In K. Potowski & R. Cameron (Eds.), *Spanish in contact: Educational, social, and linguistic Inquiries* (pp.311-327). Amsterdam: John Benjamins.
- Orozco, R. (2009). La influencia de factores sociales en la expresión del posesivo. *Lingüística*, 22, 35-60.
- Orozco, R. (2010). Variation in the expression of nominal possession in Costeño Spanish. *Spanish in context*, 7(2), 194-220.
- Orozco, R. (2015). Pronominal variation in Colombian costeño Spanish. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 17-37). Washington DC: Georgetown University Press.
- Orozco, R. (2016). Subject pronoun expression in Mexican Spanish: ¿Qué pasa en Xalapa? *Proceedings of the Linguistic Society of America*, 1, (7), 1-15.
- Orozco, R. & Guy, G. (2008). El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana? In M. Westmoreland & J. Thomas (Eds.), *Selected Proceedings of the Fourth Workshop on Spanish Sociolinguistics* (pp. 70-80). Somerville, MA: Cascadilla.
- Ostler, N. (2010). *The last lingua franca: English until the return of Babel*. Bloomsbury Publishing USA.

- 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* 83 (4), 770-802.
- Otheguy, R., & Zentella, A. C. (2012). *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. New York: Oxford University Press.
- Paolillo, J. C. (2002). *Analyzing linguistic variation: Statistical models and methods*. Stanford: CSLI Publications.
- Park, S. (2013). Testing the interface hypothesis: Acquisition of English articles by Korean L2 learners. In J. Amaro (Ed.), *Proceedings of the 12th Generative Approaches to Second Language Acquisition Conference* (pp. 155-161). Cascadilla Proceedings Project.
- Péllissier, R. (1964). *Los territorios españoles de África*. Madrid: Consejo Superior de Investigaciones Científicas, Instituto de Estudios Africanos.
- Pérez-Leroux, A. T., & Glass, W. (1997). *Contemporary perspectives on the acquisition of Spanish*. Somerville, MA: Cascadilla Press.
- Pinheiro, J. C. & Bates D. (2001). *NLME: Software for Mixed Effects Models*. Retrieved from <http://stat.bell-labs.com/NLME/index.html>.
- Poplack, S. (1980). Sometimes I'll start a sentence in spanish y termino en espanol: toward a typology of code-switching1. *Linguistics*, 18(7-8), 581-618.
- Poplack, S. (1993). Variation theory and language contact. In D. Preston (Ed.), *American dialect research* (pp. 251-286). Amsterdam: Benjamins.
- Poplack, S. & Tagliamonte, S.A. (2001). *African American English in the Diaspora: Tense and Aspect*. Malden, Blackwell. In N. Smelser, & P. Baltes (Eds.), *International Encyclopedia of the Social and Behavioral Sciences* (pp. 2016-2065). Amsterdam: Elsevier Science.
- Posio, P. (2011). Spanish subject pronoun usage and verb semantics revisited: First and second person singular subject pronouns and focusing of attention in spoken Peninsular Spanish. *Journal of Pragmatics*, 43(3), 777-798.
- Prada Pérez, A. (2009). Subject expression in Minorcan Spanish: Consequences of contact with Catalan. The Pennsylvania State University. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 143-163). Washington DC: Georgetown University Press.
- Prada Pérez, A. (2015). First person singular subject pronoun expression in Spanish in contact with Catalan. In A. Carvalho, R. Orozco & N. Lapidus Shin (Eds.),

- Subject pronoun expression in Spanish: A cross-dialectal perspective* (pp. 121-142). Washington DC: Georgetown University Press.
- Quilis, A., & Casado Fresnillo, C. (1992). Phonology and phonetics of the Spanish-language spoken in Equatorial Guinea. *Revue de linguistique Romane*, 56(221-22), 71-89.
- Quilis, A., & Casado Fresnillo, C. (1995). *La lengua española en Guinea Ecuatorial*. Madrid: Universidad nacional de educación a distancia.
- Ranson, D. L. (1991). Person marking in the wake of /s/-deletion in Andalusian Spanish. *Language Variation and Change*, 3(2), 133-152.
- Rosengren, P. (1974). *Presencia y ausencia de los pronombres personales sujetos en el español moderno*. Stockholm: Almqvist & Wiksell.
- Rothman, Jason. (2008). Aspectual Morphology use in Adult L2 Spanish and the Competing Systems Hypothesis: When Pedagogical and Linguistics Rules Conflict. *Languages in Contact*, 8(1), 74-106.
- Ruiz Martínez, A. (2002). La enseñanza del español como lengua extranjera en Guinea Ecuatorial y la interferencia de las lenguas indígenas. In M. Pérez, & J. Coloma (Eds.), *El español, lengua de mestizaje y la interculturalidad. Actas del XIII Congreso de Asele* (pp. 762-770). Murcia, España: Asele.
- Sankoff, D. (1988). Sociolinguistics and syntactic variation. *Linguistics: the Cambridge survey*, 4, 140-161.
- Sankoff, G. (1980). *The social life of language*. University of Pennsylvania Press.
- Sankoff, G. (2002). 25 Linguistic Outcomes of Language Contact. In J. K. Chambers, P. Trudgill, and N Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 638-668). Oxford: Blackwell.
- Samper, J. (1990), *Estudio sociolingüístico del español de Las Palmas de Gran Canaria*. Las Palmas: La Caja de Canarias.
- Selinker, L. (1972). Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching*, 10(1-4), 209-232.
- Serratrice, L., Sorace, A., & Paoli, S. (2004). Crosslinguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and cognition*, 7(3), 183-205.
- Shin, N. L. (2013). Women as leaders of language change: A qualification from the bilingual perspective. In A. Carvalho & S. Beaudrie (Eds.), *Proceedings of the 6th*

- International Workshop on Spanish Sociolinguistics* (pp. 135-147). Somerville, MA: Cascadilla Proceedings Project.
- Shin, N. L. (2014). Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change*, 26(3), 303-330.
- Shin, N. L., & Otheguy, R. (2009). Shifting sensitivity to continuity of reference: Subject pronoun use in Spanish in New York City. In M. Lacorte & J. Leeman (Eds.), *Español en Estados Unidos y en otros contextos: Cuestiones sociolingüísticas, políticas y pedagógicas* (pp. 111-136). Madrid: Iberoamericana.
- Shin, N. L., & Otheguy, R. (2013). Women as leaders of language change: A qualification from the bilingual perspective. In A. Carvalho & S. Beaudrie (Eds.), *Selected Proceedings of the 6th Workshop on Spanish Sociolinguistics* (pp.135-147). Somerville, MA: Cascadilla Proceedings Project.
- Shin, N. L., & Montes-Alcalá, C. (2014). El uso contextual del pronombre sujeto como factor predictivo de la influencia del inglés en el español en Nueva York. *Sociolinguistic Studies*, 8(1), 85-110.
- Shin, N. L., & Van Buren, J. (2016). Maintenance of Spanish subject pronoun expression patterns among bilingual children of farmworkers in Washington/Montana. *Spanish in Context*, 13(2), 173-194.
- Silva-Corvalán, C. (1982). *Subject expression and placement in Mexican-American Spanish*. Cambridge: Cambridge University Press.
- Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles*. Oxford: Oxford University Press.
- Silva-Corvalán, C. (2001). *Sociolinguística y Pragmatia del Español*. Washington D.C.: Georgetown University Press.
- Silva-Corvalán, C. (2014). *Bilingual language acquisition: Spanish and English in the first six years*. Cambridge University Press.
- Silveira Neto, A., De Souza. (2012). *Subject expression in brazilian portuguese: Construction and frequency effects* (Order No. 3517592). Available from Linguistics and Language Behavior Abstracts (LLBA); ProQuest Dissertations & Theses Global. (1032784459). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/1032784459?accountid=4485>
- Simons, G. F., and Fennig C. D. (2018). *Ethnologue: Languages of the World*, 21st ed. Dallas: SIL International. Retrieved from <https://www.ethnologue.com/ethnoblog/gary-simons/welcome-21st-edition>

- Soares da Silva, H. (2006). *O parâmetro do sujeito nulo: Confronto entre o português e o espanhol*. (Unpublished Doctoral dissertation). Universidade Federal do Rio de Janeiro, Brazil.
- Solomon, J. (1998). External constraints and variable subject expression in Yucatec Spanish. Paper presented at the annual meeting of New Ways of Analyzing Variation (NWAYE) twenty-seven, Athens, Georgia.
- Solomon, J. (1999). *Phonological and syntactic variation in the spanish of valladolid, yucatan* (Order No. 9958201). Available from ProQuest Dissertations & Theses Global. (304547157). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/304547157?accountid=4485>
- 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. (2006). Gradedness and optionality in mature and developing grammars. *Gradience in grammar: Generative perspectives*, 106-123.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic approaches to bilingualism*, 1(1), 1-33.
- Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22(3), 339-368.
- Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, 13(2), 195-210.
- Sundiata, I. K. (1990). *Equatorial Guinea: colonialism, state terror, and the search for stability*. Boulder: Westview Press.
- Tagliamonte, S. (2011). *Variationist sociolinguistics: Change, observation, interpretation* (Vol. 40). John Wiley & Sons.
- Tagliamonte, S. A. (2006). *Analyzing Sociolinguistic Variation*. Cambridge and New York: Cambridge University Press.
- Thomason, S. G. (2001). *Language contact: an introduction*. Edinburgh & Washington, DC: Edinburgh University Press - Georgetown University Press.
- Thomason, S. G. (2008). Social and linguistic factors as predictors of contact-induced change. *Journal of language contact*, 2(1), 42-56.

- Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Toribio, J. (1994). Dialectal variation in the licensing of null referential and expletive subjects. Aspects of Romance linguistics. In C. Parodi, C. Quicoli, M. Saltarelli and M. Zubizarreta, (Eds.), *Selected papers from the 24th Linguistic Symposium on Romance Languages* (pp. 409-432). Washington: Georgetown University Press.
- Toribio, J. (2004). Convergence as an optimization strategy in bilingual speech: Evidence from code-switching. *Bilingualism: Language and Cognition*, 7(2), 165-173.
- Torres Cacoullous, R., & Travis, C. E (2010). Variable yo expression in New Mexico: English influence. *Journal of the Rivera-Mills & Villa*, 203, 185-206.
- Torres Cacoullous, R., & Travis, C. E (2011). Testing convergence via codeswitching: Priming and the structure of variable subject expression. *International Journal of Bilingualism*, 15 (3), 241-267.
- Torres Cacoullous, R., & Travis, C. E. (2015). Foundations for the study of subject pronoun expression in Spanish in contact with English: Assessing interlinguistic (dis)similarity via intralinguistic variability. In A. Carvalho, R. Orozco, & Shin, N. (Eds.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*. (pp. 83–102). Georgetown: Georgetown University Press.
- Travis, C. E. (2005). The yo-yo effect: Priming in subject expression in Colombian Spanish. In R. Gess & E. Rubin (Eds.), *Selected Papers from the 34th Linguistic Symposium on Romance Languages* (pp. 329-349). Amsterdam: John Benjamins.
- Travis, C. E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language variation and change*, 19(2), 101-135.
- Travis, C. E., & Cacoullous, R. T. (2012). What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics*, 23, 711-748.
- Travis, C. E & Lindstrom A. (2016). Different registers, different grammars? Subject expression in English conversation and narrative. *Language Variation and Change*, 28 (1), 103-128.
- Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, 23(2), 215-242.

- Tsimpli, I. M. & Mastropavlou, M. (2008). Feature interpretability in L2 acquisition and SLI: Greek clitics and determiners. In J. Liceras, H. Zobl & H. Goodluck (Eds.), *The role of formal features in second language acquisition*, (pp. 142-183). Mahwah, NJ: Lawrence Erlbaum.
- Tsimpli, I., & Sorace, A. (2006). Differentiating interfaces: L2 performance in syntax-semantics and syntax-discourse phenomena. In D. Bamman, T. Magnitskaia, and C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development* (pp. 653-664). Cascadilla Press: Somerville, MA.
- Vuskovich, M. (2013). *A contrastive analysis of spanish and fang: An L2 approach to equatorial guinean spanish as spoken in bata city* (Order No. 3606374). Available from Linguistics and Language Behavior Abstracts (LLBA); ProQuest Dissertations & Theses Global. (1492334601). Retrieved from <http://login.ezproxy1.lib.asu.edu/login?url=https://search-proquest-com.ezproxy1.lib.asu.edu/docview/1492334601?accountid=4485>
- Williams, A., & Kerswill, P. (1999). Dialect levelling: change and continuity in Milton Keynes, Reading and Hull. *Urban voices: Accent studies in the British Isles*, 141-162.
- Winford, D. (2005). Contact-induced changes: Classification and processes. *Diachronica*, 22(2), 373-427.
- White, L. (2011). The interface hypothesis: How far does it extend?. *Linguistic approaches to bilingualism*, 1(1), 108-110.
- White, L. (2007). Linguistic theory, universal grammar, and second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction*, (pp. 37-55). Mahwah, NJ: Erlbaum.
- Yakpo, K. (2009). Complexity revisited: Pichi (Equatorial Guinea) and Spanish in contact. *Simplicity and complexity in creoles and pidgins*. In G. Nicholas, Faraclas & T. Klein (Eds.), *Simplicity and complexity in creoles and pidgins* (pp. 183-215). London: Battlebridge.
- Zamponi, R. (2009). *Focus in Fang and other Bantu A70 languages*. Unpublished manuscript, University of Siena, Siena.

APPENDIX A
LANGUAGE BACKGROUND QUESTIONNAIRE

- a. (Pseudonym):
- b. Sex:
- c. Age:
- d. Education:
- e. Birth place:
- f. How long have you been living in Equatorial Guinea?
- g. Do you travel outside Equatorial Guinea? If yes, how often?
- i. How old were you when you first learnt Spanish?
- h. How often do you speak Spanish?
- i. Do you speak any other languages apart from Spanish? If yes, name them.
- j. Which language do you often use at home?
- k. How many years of schooling (elementary school through university) did you complete?
- l. Where do you currently live in Equatorial Guinea?

APPENDIX B
INTERVIEW QUESTIONS

My name is _____ and I will be conducting this interview. We will first start with demographic questions. If at any point you do not feel comfortable answering the questions, just let me know. We can also get rid of your responses if you no longer want to be a part of the interview.

1. How old are you?
2. What is your occupation?
3. Were you born in Equatorial Guinea? If yes, which part of the country were you born? If your answer is no, how long have you been in Equatorial?
4. Where are your parents from? How long have they been where they reside?
5. Do you have brothers and sisters? How many?

Education

6. Did you attend any of the schools in your neighborhood? Which one? Was it far or close to your house?
7. Did you have any difficult professors? For what reason will a professor reprimand a child? What was the worst punishment you ever received? What was the worse joke a student ever played on a professor?
8. Have you ever been punished for something you did not do? Have you ever behaved badly in class? What happened?
9. Have you ever had a professor who you felt was unfair to you? Have you ever had a favorite professor? What made this professor your favorite?
10. Have you ever copied notes during an exam in class? If yes, were you caught?
11. Did you have a best friend in school? If yes, what made this person your best friend? Did you ever have a boyfriend or girlfriend in school? What made you like this person?
12. Was there anyone in your school who you did not get along with?
13. Have you ever engaged in a heated conversation? What cause it?
14. Did you attend the university? What did you study? Have you graduated? If no, how much long do you have left?

Music

15. What type of music do you listen to? What is the name of your favorite band? What type of music do they sing? Have you ever been to a concert? If no, which type of concert would you have liked to go to? If yes, which concert did you attend?
16. Which type of music do you detest?
17. Do you prefer music in Spanish or in the local dialects? Why?
18. Do you play a musical instrument?
19. Do you buy CDs or you download music from the internet?
20. Do you like to dance?

Food

21. What is your favorite food?
22. Do you know how to cook?
23. Can you describe how you will prepare a meal of your choice?
24. What is your opinion regarding fast food?

Hobbies

25. What is your favorite hobby?

26. Are you an indoor or outdoor person?
27. Which sports do you like? Do you prefer to watch sports on the television or in a stadium?
28. What is your favorite team?
29. Do you spend a lot of time on Facebook or twitter? What pages on the internet do you usually visit
30. Do you like to read? If yes, what type of books do you like to read? If no, what do you prefer to do in your free time?
31. What is your favorite television program?
32. Do you watch the news on TV or you get your news from the newspaper?

Transportation

33. What form of transportation do you usually use?
34. Do you have a car? What model is this car?
35. Do you prefer cars from the United States, Germany, Asia or elsewhere?
36. Do you prefer an automatic transmission or a manual one? Why?
37. Do you like motorcycles?

Travel

38. Do you like to travel?
39. Which areas have you visited in the past?
40. What was your favorite trip?
41. If you had the opportunity to travel, where will you go?138

APPENDIX C
IRB APPROVAL



EXEMPTION GRANTED

Alvaro Cerron-Palomino
International Letters and Cultures, School of (SILC)
489/727-0238
Alvaro.Cerron-Palomino@asu.edu

Dear Alvaro Cerron-Palomino:

On 2/5/2019 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Subject pronoun expression in Equatoguinean Spanish
Investigator:	Alvaro Cerron-Palomino
IRB ID:	STUDY00009343
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	<ul style="list-style-type: none"> • questionnaire, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • consent form, Category: Consent Form; • Subject Pronoun Expression in a Spanish L2-Only environment: The case of Equatorial Guinea.docx, Category: IRB Protocol; • recruitment form, Category: Recruitment Materials; • interview questions, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions);

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (2) Tests, surveys, interviews, or observation on 2/5/2019.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,



IRB Administrator

cc: Lillie Padilla
Alvaro Cerron-Palomino