Linguistic Outcomes of the Wayuunaiki-Spanish Language Contact Situation

Nelson José Méndez-Rivera

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
PhD in Spanish

Department of Modern Languages and Literatures
Faculty of Arts
University of Ottawa

ABSTRACT

The study of Spanish in contact with Wayuunaiki has received limited attention in generative and variationist analyses. In particular, the possible influence of this indigenous language on some parts of the Spanish language has not been investigated or has been only briefly addressed. This dissertation aims to fill this existing gap by studying two morpho-syntactic variables: (i) the distribution of null and overt subjects (NOS) as portrayed by the Null Subject Parameter and overt subject personal pronoun (SPP) expression as traditionally researched in variationist studies, and (ii) the issue of subject-verb agreement within the theory of features. To carry out these studies, we collected spontaneous data from 27 Wayuunaiki-Spanish bilinguals and five Spanish monolingual speakers. This Spanish monolingual group served as the vernacular benchmark.

The NOS and SPP expression are among the most studied topics in Hispanic linguistics, but they have never been systematically researched in Guajiro Spanish. By analyzing these issues in our dissertation, we want to contribute new data to their study and to the properties and factors affecting them, in order to widen the knowledge of how they function in this Spanish language contact situation. In the generative analysis of the null/overt subjects we investigate whether the distinction that occurs in Wayuunaiki between stative and active verbs and the participants' proficiency in Spanish have an impact on the distribution of NOS in Guajiro Spanish. In the variationist study, a number of independent variables widely believed to constrain variable SPP expression are factored into the investigation to find out how

they behave in this particular language contact situation. These variables include person/number, TAM, switch reference, priming effects, etc.

The second research topic is subject-verb agreement. This is only investigated from the generative grammar perspective, specifically by employing the concept of features to explain the nature of the subject-verb mismatches produced by the Wayuunaiki dominant Wayuunaiki-Spanish bilinguals. We specifically investigate whether the Spanish conjugation system poses a problem to the bilingual speakers' ability to produce the native Spanish verb forms vis-à-vis the monolingual cohort and whether this ability is shaped by the participants' proficiency in Spanish.

The results of the three studies contribute to the field of Hispanic linguistics from three different perspectives. The study of NOS adds a new dimension to the pro-drop parameter: the possible role that Wayuunaiki's double conjugation may play in the distribution of null and overt subject pronouns in Guajiro Spanish. The variationist study provides new data on the topic of SPP expression in a variety of Spanish which, in this specific case, has as contact language an understudied indigenous language. The analysis of the subject-verb mismatches that occur in Guajiro Spanish allows us to differentiate between the status of null and overt subjects with respect to subject-verb agreement and to differentiate between this contact variety and the Colombian vernacular benchmark.

ACKNOWLEDGEMENTS

I want to extend my gratitude to the following people for helping me in the process of completing this dissertation.

To the 27 Guajiro people and the five monolingual Spanish speakers who accepted to be interviewed so that I could collect the necessary data to complete this dissertation.

To my supervisor, Professor Juana Liceras and my co-supervirsor Professor Stephen Levey. Each one of them in their area has provided me with the guidance I needed to write this thesis.

To professor Jose Alvarez for giving me invaluable feedback, examples, and commentaries about the Wayuunaiki language.

To my oral defence committee, professor Joyce Bruhn de Garavito from The University of Western Ontario, Professor Liliana Sánchez from the University of Illinois at Chicago, Professor Elena Valenzuela from the University of Ottawa and Professor Anahí Alba de la Fuente from the Université de Montréal.

To my wife, Ana Cabrera whose support has provided me with the motivation I needed to finish this dissertation.

To all of my relatives and friends who have been systematically asking me when I would finish this project. As I did not want to disappoint any of them, I made the resolution to finish what I had started.

Finally, to the University of Ottawa for its support with the admission and excellence scholarships. To APTPUO for its funding to attend conferences, to the

Faculty of Graduate and Postdoctoral Studies for the funding to visit La Guajira to collect the data, and to my colleagues at the Language Acquisition Research Laboratory for their friendship and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The study of the Spanish spoken by the Guajiro people of Venezuela and Colombia is exiguous and more research is required to understand this variety of Spanish and to broaden our understanding of the linguistic phenomena that have never been reported and/or analysed. The distribution of null and overt pronominal subjects (NOS) and the issue of overt subject personal pronoun (SPP) expression, for instance, have never been examined in Spanish in contact with Wayuunaiki and the problem of subject-verb agreement has received limited attention. A case in point concerns Pimienta (2008), who defends the idea that Wayuunaiki's conjugation system affects how the bilingual speakers produce the Spanish subject-verb agreement requirements, and Ramírez Cruz (2009) who claims that the Wayuunaiki speakers have not completely learned the Spanish conjugation system.

Having scrutinized the research carried out on Guajiro Spanish, the time is ideal to offer new perspectives regarding the two aforementioned topics. This dissertation is consequently built around the study of the distribution of NOS, overt SPP expression and subject-verb agreement in Guajiro Spanish. Concerning the realization of the pronominal subject, this topic receives a double treatment, since it is addressed from a generative and a variationist approach. On the other hand, subject-verb agreement is examined from the generative perspective only.[1]

The rationale behind the idea of using two distinct frameworks for one issue comes from the argument that they are complementary rather than exclusive (Adger & Smith, 2005; Cornips

[1] The study of subject-verb agreement is also feasible from the variationist perspective. In our case, however, it was not our main aim to carry out that study because our goal was to contribute to the variationist research with the study of the SPP expression. Furthermore, having just 431 mismatched tokens would have made a variationist analysis with Varbrul very difficult. The limited number of tokens and the irregular distribution of the data would not have generated sufficient statistical power to be helpful or revealing as to what was happening in Guajiro Spanish regarding this issue. Another issue was the comparability with the Spanish benchmark group, which would have been hampered by the limited number of mismatched tokens produced by this group.

& Corrigan, 2005). A linguistic topic can receive an analysis from both sides and our understanding of it can be wider. For example, the study of the overt expression of the pronominal subject has been widely investigated across different varieties of Spanish (Silva-Corvalán, 1982; Bentivoglio, 1987; Cameron, 1993, 1995; Lowther, 2004; Flores-Ferrán, 2004, 2005, 2007; Hurtado, 2005; Travis, 2007; Otheguy, Zentella & Livert, 2007; Orozco & Guy, 2008; Torres-Cacoullos & Travis, 2010; Carvalho & Child, 2011; Sánchez Arroba, 2011; De Prada, 2009; Otheguy & Zentella, 2012; Holmquist, 2012; Abreu, 2012; Shin, 2012; and Michnowicz, 2015, *inter alia*), but it has received less attention from a combined perspective (Martínez-Sanz, 2011, Moreno, 2019). This type of combined investigation provides, therefore, patterns of variations from the variationist sociolinguistics perspective and theoretical explanations from the generative grammar analysis.

This dissertation is organized in the following manner:

Chapter 2 provides an overview of the Guajiro people. Their socio-historical background is presented, as well as some facts about how this indigenous group lives together with the other surrounding population. This chapter also gives a brief account of the Wayuunaiki-Spanish contact situation and some general information related to the Wayuunaiki language.

Chapter 3 offers the theoretical frameworks that are used to analyse the two issues under examination. First, the arguments to marry theoretical linguistics and sociolinguistics to study language variation and change are presented. Subsequently, each paradigm is discussed separately, and the main theoretical tenets and views of each theory about language are described.

Chapter 4 is about the participants and the nature of the data sources. First of all, how we address and how we entered the community and interviewed the indigenous population. Second, we address the sociolinguistic data for the monolingual Spanish and bilingual speakers, as well as

how the bilingual informants were assessed in terms of their Spanish proficiency. We also describe how the data extracted from the interviews were analysed.

Chapter 5 reports the results for the study of the distribution of null and overt pronominal subjects (NOS). In this chapter, this issue is investigated from the perspective of the generative grammar. We analyse whether Wayuunaiki verb classes (stative vs. active transitive/intransitive verbs) and Spanish proficiency influence the rate and distribution of null and overt subject pronouns in Guajiro Spanish, the way in which the verb types favours NOS realization, and how the results of the bilingual group relate to the Spanish native benchmark. For the bilingual group, the overt subject pronoun realization was preferred with stative verbs, while the monolingual cohort preferred active intransitive verbs. The results were identical for the intermediate and advanced speakers, which rules out a possible effect of Spanish proficiency. However, the results for the bilingual and the monolingual group differed, supporting the idea that the Wayuunaiki-Spanish bilinguals and the monolingual Spanish speakers' grammars are not entirely the same regarding this specific linguistic sub-component.

Chapter 6 presents the findings of the analysis of overt SPP expression as carried out within the framework of variationist sociolinguistics. First, we obtain the rate of overt SPP usage and then we examine if the factor groups of 'grammatical person/number', 'TAM', 'priming effects', 'switch reference', 'reflexive use of the verb', etc. favour or disfavour the likelihood of a pronominal subject being overtly realized.[2] The Wayuunaiki-Spanish speakers and the monolingual Spanish speakers differ in the rate of SPP expression, but rate discrepancies between comparison groups do not constitute adequate proof of divergent linguistic conditioning in this

[2] Priming effects in this dissertation refer to the production of an overt SPP triggered by the presence of a previous overt SPP as produced in a conversational context elicited through interviews. In this thesis the term priming effects is not meant to refer to an experimental technique.

sector of the grammar. Moreover, the slight differences found in the constraint rankings associated with the comparison groups do not support the argument of a radical dissimilarity between the systems under analysis.

Chapter 7 presents the study of subject-verb agreement from the perspective of generative grammar. By using the concept of features, the production of mismatched clauses found in the bilingual data is analysed. To account for the possible effects of Spanish vernacular phenomena such as the /s/ and /n/ elision, the results of the bilingual group are compared to those of the Spanish native group. The findings strongly support the conclusion that 3rd person singular is used as the default conjugation by the bilingual speakers when the Spanish conjugation is not retrieved. This claim is supported by McCarthy's (2004) Morphological Underspecification Hypothesis (MUH). Concerning the effect of language proficiency, the results give support to the argument that the mismatched sentences of the bilingual cohort are temporary. The higher the Spanish proficiency, the smaller the number of this type of sentences is found in the data, as hypothesized by the Missing Surface Inflection Hypothesis (MSIH) of Prévost and White (2000b), and Prévost (2008).

Chapter 8 summarizes the findings of this dissertation and offers the general conclusions of this thesis, the limitations, as well as suggestions for further research on Guajiro Spanish.

2. THE GUAJIROS: SOCIO-HISTORICAL BACKGROUND, THE WAYUUNAIKI-SPANISH CONTACT SITUATION AND THE WAYUUNAIKI LANGUAGE

2.1 INTRODUCTION

In this chapter, the specific terminology used in this dissertation is explained as well as what we refer to when using the terms Guajiro and Wayuunaiki. In order to contextualize the research detailed in this thesis, a portrait of the Guajiro people will be provided along with a description of their current political and economic situation.[3] Some historical facts concerning the relationship between the Guajiros and the Spanish colonizers/settlers will be presented as well. Moreover, we will discuss the Wayuunaiki-Spanish language contact situation as well as providing a brief description of the Wayuunaiki language.

2.2 TERMINOLOGICAL PRECISION

In this dissertation, the term *Guajiro Spanish* refers to the Spanish spoken by the Guajiro people. We base this decision on the fact that this variety is specifically spoken by members of the Guajiro community and their Spanish has some features that differentiate it from the local Spanish variety. For example, issues of gender assignment as in (1a), in which the expected native Spanish gender is masculine asi in (1b), but the NP is realized with a feminine article and noun (Álvarez, 2000; Atencio, 2014; Oquendo, 2003), and issues with subject-verb agreement as in (2a), where the pronoun and the verb do not match (Ramírez, 2009; Pimienta, 2008) as they do in the native Spanish variety or the Spanish vernacular to which the Wayuunaiki speakers are routinely exposed as in (2b).

[3] Sociolinguistic details of the Guajiro participants of this dissertation are presented in Chapter 4.

(1) a. Un**a** velori**a** que dura tres días
a.ART.F.S wake.F.S that lasts three days
b. Un velori**o** que dura tres días
a.ART.M.S wake.M.S that lasts three days
'A wake that lasts three days'.

(Atencio, 2014, p. 68)

(2) a. Nosotros no conoce nada
We.**1P** not know.**3S** anything
b. Nosotros no conocemos nada
We.**1P** not know.**1P** anything
'We don't know anything'.

(Ramírez, 2009, p. 175)

The decision to name Guajiro Spanish a variety follows Hudson's (1996) and Ferguson's (1972) idea that a language variety can be identified when there are linguistic traits such as sounds, words, and/or grammatical features that are unequivocally correlated to an ethnic group (the Guajiro people) that is circumscribed to a geographical area (La Guajira Peninsula) (Wardhaugh, 2006).

The terms *Guajiro* or *Goajiro* are used interchangeably to refer to the language and people of the La Guajira Peninsula. However, in this dissertation the term *Guajiro* is assigned to the people and *Wayuunaiki* is used to name the indigenous language spoken by the Guajiro people. Regarding the peninsula where the Guajiros live, the phrase *La Guajira Peninsula* is used. However, we sometimes simply use the phrase *La Guajira* alone to mean the region where the Guajiros primarily live. We will also use the phrase *indigenous people(s)* to refer to the Guajiros and to other native peoples living in La Guajira and in the Americas.

Finally, we borrow the word *alijuna* from Wayuunaiki. The Guajiros call foreigners 'alijunas', which includes any stranger or person who is not one of them. In La Guajira, the word is heard and used in reference to Spanish-speaking people. Therefore, when we use it in this

dissertation, we are referring to a non-Guajiro person, mainly a westernized person, since Guajiros also have names for members of other indigenous tribes.

2.3 A Contemporary Perspective on the Guajiros

The Guajiros are an indigenous group of people of Arawak origin who migrated from the Amazon River and Río Negro River region (the area where the city of Manaus in Brazil is presently located) to a place now called La Guajira (Figure 2.1). In this territory, extending to parts of present-day Colombia and Venezuela (Pérez, 2006), the Guajiros settled and have flourished. According to the latest census information, there are 415,498 Guajiros in Venezuela (INE, 2012) and 380,460 in Colombia (DANE, 2018). The Guajiros and their language are the most vigorous indigenous group in these two South American countries (Pineda, 2005).

Fig. 2.1. *Localization of La Guajira. Google Maps. Accessed 23 Sep. 2019.*

The Guajiros reside mainly in the Department of La Guajira in Colombia and the State of Zulia in Venezuela. Although both countries have established a border, for the Guajiros, La Guajira is just one ancestral land despite the modern political division imposed on their territory. Ethnically and politically, the Guajiros consider themselves as one nation (Arellano, 1987; "Declaración de

la Nación Wayúu,[4] 2009") and they seek autonomy, equality, and respect for their identity. Many Guajiros remain in rural areas and live in settlements called *rancherías* (Figures 2.2, 2.3, and 2.4). Rancherías are compounds consisting of several concrete houses or thatched sheds of mud walls, usually belonging to members of the same family (González, 1973). There are also certain pockets of Guajiros living in large urban areas, such as in Maracaibo and Barranquilla, as well as in small towns close to these cities. There are Guajiros working as hacienda labourers and as domestic staff far away from their tribal land.[5] Generally, these Guajiros have migrated in search of better living conditions and job opportunities. The migration of the Guajiros to the city depends also on the season. In the dry season, as resources (water and pastures) in their territory diminish or disappear, they move to the city or haciendas to find work. In contrast, they return to their homes during the rainy season as there is an abundance of resources and they can work with their livestock (Vergara, 1987). Around 40% of Guajiros are herdsmen (Gobernación de la Guajira, 2012).

Fig. 2.2. *View of a ranchería. Photo by Geiner Amaya, 2018.*

[4] Declaration of the Guajiro Nation.
[5] Haciendas are large estates found in Latin America.

Fig. 2.3. *A school. Photo by Geiner Amaya, 2018.*

Fig. 2.4. A soccer field. Photo by Geiner Amaya, 2018.

Due to seasonal constraints, the economy of the Guajiros in their rural setting is mixed and limited. It is mainly based on raising goats and sheep, producing handicrafts such as bags and hammocks, farming corn, beans, melon, watermelon, and pumpkins, and extracting salt (Ramírez, 2007). Livestock is extremely important for the Guajiros as it gives them prestige and may function as currency or as a way to measure a family's wealth. For example, if there is a dispute between two families, one way of solving the problem is by paying the other family with a certain amount of sheep, goats, and/or mules. Livestock can also be used as a dowry (García, López & Agudo, 1996).

Guajiros living in rural areas travel to markets in cities or small towns to sell their handicrafts and meat surpluses in order to earn money to purchase manufactured goods or products they cannot produce themselves. Guajiros living close to the sea base their economy and subsistence mainly on fishing. Poverty is pervasive among the Guajiros, leading many of them to seek other means of survival. Following a practice that has been well documented for a number of years, many Guajiros are also in the business of smuggling gasoline, rice, flour, cooking oil, and other staples into Colombia (Ramírez, 2007). They do this to take advantage of the price gap existing between products in Colombia and Venezuela. Because the Venezuelan government subsidizes gasoline and many products of daily use, they are thus cheaper there than in Colombia. Smuggling has been a common activity among some Guajiros since colonial times, yet this is not typical of the majority of Guajiros and is only common to residents of Colombia and Venezuela living along the border. Nevertheless, this situation has led Guajiros into conflicts with several groups of people as they compete with alijunas to purchase products that are sometimes scarce, and they get into trouble with Venezuelan and Colombian authorities by engaging in these illegal activities. Additionally, they are subject to being pursued and assassinated by criminal groups that want to control their territory to carry out unlawful activities such as drug and weapons smuggling, kidnappings, and murders (Ramírez, 2007).

2.4 A Brief Account of the Colonization and Settlement of La Guajira

The relationship between the Guajiros and Spanish speakers has persisted for nearly 500 years. In the final week of August 1499, Alonso de Ojeda was the first Spaniard to reach the La Guajira Península (Paredes, 2006). The first allusion to La Guajira region can be found in Fernando Colon's 1527 map of America and Diego Ribero's 1529 map (Arellano, 1987). The very first mention of the word Guajiros (Goaxiros in the original text) comes from Fray Pedro Simon in

Noticias Historiales de las Conquistas de Tierra Firme en las Indias Occidentales[6] (1627), where Fray Pedro Simón describes an indigenous group of people who is very violent, difficult to master, very skilled in the art of war, and who had even harmed Spanish settlers (Arellano, 1987).

Between the 16th and 18th centuries, the Guajiros burned down several Spanish settlements and each one of the Spanish military expeditions to conquer La Guajira failed (Barrera, 1990). The relationship between the Spaniards and this indigenous people was never easy as was demonstrated by the uprising of the Guajiros in 1769. This event would lead to the independence of the Guajiro territory from Spanish dominion (Polo, 2011). By 1820, Colombia was an independent state, and by extension the Guajiro territory.

Throughout the rest of the 19th century, hostilities continued but were now directed at the government of the new Republic of Colombia. To avoid, or at least to mitigate this, the government appointed as governors people known and respected by the Guajiros. In 1826, Vice President Santander also proclaimed that indigenous peoples in general would be protected, treated as Colombian citizens, and "civilized" (De La Pedraja, 1981). In this effort to "civilize" the Guajiros, several steps were taken including creating and recognizing an official Guajiro territory in 1846 and reintroducing Catholic missionaries to help in the assimilation of the Guajiro people to the new Colombian society which was strongly based on Spanish traditions.

During the 19th and the 20th century, Guajiros started a slow process of integration. They traded cattle, sheep, goats, alcohol, leather, and textiles with other inhabitants of the region. They extracted salt, smuggled products, and even started dressing like alijunas. By the mid-1850s, droughts, diseases, and famine forced some Guajiros to migrate to urban centers (Pérez, 2006).

[6] *News on the Historical Conquests of the Mainland in the West Indies.*

Currently, Guajiros are "*indigenous people with trucks"* (Perrin, 1995). They are struggling to cope with external factors such as lack of fresh water, unemployment, and limited health services, while at the same time keeping some of their traditions. Despite the prolonged contact with Spanish speakers and other indigenous peoples, they have kept their language, religion, and social organization (Etxeberria, 2012).

The Guajiros are fully recognized at present and have access to some ethno-bilingual schooling in their territory, their traditions are studied and shared, and there is a revalorization of the Guajiro culture. In addition, Wayuunaiki is a co-official language along with Spanish in Colombia (Constitución de Colombia, 1991) and Venezuela (Constitución de la República Bolivariana de Venezuela, 1999) in the territories where it is spoken.

2.5 The Onset of the Wayuunaiki-Spanish Language Contact Situation

The Wayuunaiki-Spanish contact situation is well known and is reported, for instance, by Lipski (2010) and Palacios and Calvo-Pérez (2008) who consider it as one of the most salient in the indigenous-Spanish context as the two languages have been in contact for very long time. Undoubtedly, Guajiros were exposed to the Spanish language with the arrival of the first colonizers in La Guajira in the 16th century. The conquering process started with the Spanish founding settlements that were sometimes destroyed by the Guajiros and other indigenous peoples. Once the Spanish colonizers discovered the potential business opportunity of the pearl industry, they enslaved the indigenous peoples to make them find pearls in the sea (Polo, 2005; Ángel, 2014).

With this colonization came the proselytization and education of the indigenous peoples. However, this enterprise failed several times. In the beginning, the Guajiros were difficult to subdue and the most significant challenge was the use of Spanish to teach them the Catholic religion and to make them work for the Spanish settlers. The first Capuchins arrived in the New

World in 1647 and in La Guajira in 1696. They were rapidly expelled in 1701 after an insurrection by the indigenous people (Polo, 2010). Subsequent missions were initiated but aborted as well. First, due to the failure to convert indigenous peoples to Christianity, and second, because of the indigenous peoples' animosity. Probably understanding this language barrier, Priest Rafael Celedón together with Ezequiel Uricoechea wrote one of the first Wayuunaiki grammars: *Gramática, Catecismo y Vocabulario de la Lengua Goajira*[7] (1878), so that subsequent missionaries could learn the Wayuunaiki language and could help in the task of evangelizing the Guajiros (De La Pedraja, 1981; Pérez, 2006). The Capuchins not only learned Wayuunaiki, but they also taught Guajiros the Spanish language.[8]

In other regions of the New World, there were, nevertheless, documents in Aboriginal languages. Particularly in Mexico, booklets in Latin, Spanish, and Nahuatl were used to spread the Gospel among the indigenous peoples. On one hand, the use of Spanish by the indigenous peoples was encouraged. By knowing Spanish, it would be easier to indoctrinate them since Aboriginal languages were believed to be poor in vocabulary to convey important ideas. In addition, by stripping the indigenous peoples of their languages, the Spanish thought the indigenous peoples would abandon their beliefs and what the Spanish considered to be idolatries and vices. On the other hand, evangelizers made efforts to master the diverse indigenous languages as it became mandatory for them to know the general indigenous language of the place where they would be ordained to work (Torre, 1962).

[7] *Grammar, Catechism and Vocabulary of the Wayuunaiki Language.*

[8] In other parts of South America there were already by the second half of 1500s and the first half of the 1600s grammars, vocabularies, and Christian doctrines in indigenous languages (Tercer concilio limense (1583-1591), 2017). In the case of the Wayuunaki, a grammar was late to appear because La Guajira was not easy to colonize because the Spaniards and the Guajiros had a very hostile relationship.

The Guajiro's relationship with the Spanish settlers was very disproportional. This situation was apparent, for instance, in the participation of the indigenous people in the colonial government and their interactions with the representatives of the Spanish Crown. Those Aboriginal people who knew Spanish were privileged. The Crown expressly prohibited the participation of indigenous peoples in any government position, unless the candidate or applicant was able to communicate in the Romance language, as it was established in the Royal Charter of June 25, 1690 (Torre, 1962). In commercial relationships Spanish was favoured as Guajiros were forced to use Spanish to trade in areas in which they were a minority or while dealing with alijunas.

The situations described above are still persistent today. However, there is a broader recognition and understanding of the Guajiros' reality. They still have limited use of their language in the Spanish-speaking world since press, education, and government negotiations and assistance are mainly produced and carried out in Spanish. We will address in more detail the current Wayuunaiki-Spanish contact situation in the ensuing section.

2.6 The Wayuunaiki-Spanish Language Contact Situation in Modern Times

Pérez Van-Leenden (1998) affirms that in the 19th century and in the beginning of the 20th century, the links between the Guajiros and the other inhabitants, mainly Spanish speakers of La Guajira, increased. This context, in which the Guajiros are in a minority position, has led to the emergence of a diglossic situation. This situation is characterized by a state of affairs in which Spanish has more prestige than Wayuunaiki and the indigenous language is circumscribed to limited spheres of interaction.

La Guajira can be sectioned into three regions (Candelier, 1994): Upper, Middle, and Lower Guajira (Figure 2.5).

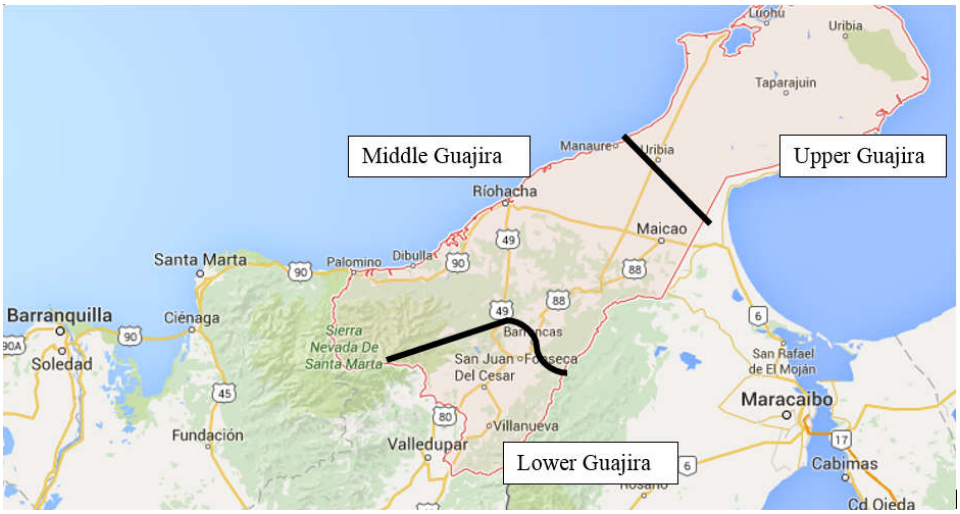


Fig. 2.5. *Division of La Guajira. Google Maps. Accessed 23/09/2019*

Upper Guajira is the less developed zone, where the use of Wayuunaiki is more common and the highest number of monolingual Wayuunaiki speakers can be found (Mejía, 2011). In this area, in some cases, they live very isolated from urban areas, and suffer from lack of fresh water and have limited access to formal education and health services (Ramírez, 2007). The data for this dissertation was collected in the jurisdiction of Uribia, a rural municipality in the Upper Guajira, but in proximity to the Middle Guajira.

In *Middle Guajira*, there is a higher proportion of bilinguals, but there are still monolingual Guajiros (Mejía, 2011). In this area, Guajiros have more contact with alijunas and the Spanish language. Although the Guajiros' homes remain mainly monolingual and their language is favoured in many contexts, the Spanish language is present in their homes via the media (e.g., radio and television) (Etxeberria, 2012). The presence of the State by means of hospitals and schools is more visible, and larger urban centers like Maicao attract Guajiros for trading or employment.

In *Lower Guajira*, the presence of the Spanish-speaking world is dominant. Wayuunaiki-Spanish bilingualism is very widespread. In some cases, there are even Guajiros who are

monolingual in Spanish, and in many contexts, Spanish is preferred among the bilinguals (Mejía, 2011; Etxeberria, 2012).

Wayuunaiki is also present in cities outside of La Guajira such as Barranquilla and Maracaibo (González, 2006). However, in these cases, its presence is limited to the Guajiros' homes or to interactions among the Guajiros (Atencio, 2014).

In general, there is a conscious commitment to preserve the language. Much of their ancestral knowledge, traditions, and beliefs are transmitted through their language, which is mainly oral (González, 2006). When Guajiros gather together for social interactions like a wedding, a burial, a visit to the *piache* (the Guajiro shaman), or when they visit their relatives, Wayuunaiki is the predominant language (Mejía, 2011). How then does Spanish infiltrate into the Guajiro world?

If the Guajiros do not go to the Spanish-speaking world, Spanish comes to them through the media. As more Guajiros have access to radio and television, Spanish comes to them through entertainment on TV and radio shows and broadcasts. For instance, in their rancherías, they carry out their daily and household activities in Wayuunaiki, while at the same time they receive Spanish input via the media.

Schooling plays also an important role. Once Guajiros get a formal education, the presence of Spanish also increases. Although they can get some education in Wayuunaiki, this is limited to the first years of primary education. When the transition to Spanish is completed, this will be the dominant language of instruction. In schools where bilingual ethno-education is provided, this is constrained to the availability of educators able to speak Wayuunaiki. Therefore, once children start going to school and learning Spanish, they become active agents that take Spanish home with them (Mejía, 2011).

Another way in which Spanish comes to the Guajiro world is through inter-ethnic marriages with a Spanish-speaking person. Since Spanish speakers generally do not speak Wayuunaiki, Spanish starts displacing Wayuunaiki at home. In some cases, there is a deliberate decision to raise their children only in Spanish so that they will have better opportunities in the future (Mejía, 2011).

In daily interactions with alijunas, Guajiros are obliged to speak Spanish most of the time. Guajiros use Spanish when they attend church services, when trading with alijunas, and when seeking social and health assistance from the government. They are forced to speak Spanish even if they have never received any education on how to speak the language and their mastery of Spanish is rudimentary. This situation creates various degrees of bilingualism so that some Guajiros speak native-like Spanish and others communicate with great difficulty. The outcome, with the less proficient and even sometimes with advanced speakers, is a Spanish that differs somehow from the Spanish variety spoken in the region.

2.7 The Wayuunaiki Language: A Concise Description

Wayuunaiki, a language spoken by about 500,000 people in Colombia and Venezuela, belongs to the Arawak family of languages (Álvarez, 2017). This family includes a group of languages (Bare, Garifuna, Lokono, etc.) spoken in many parts of the Americas in countries like Brazil, Peru, Venezuela, Colombia, Surinam, Honduras, and Belize. Although the exact number of languages of this family is uncertain, there are about forty Arawakan languages currently spoken (Aikhenvald, 1999, p. 65).

The Arawakan languages are polysynthetic, agglutinating, and mostly head-marking. Regarding verbal morphology, the verb is an obligatory element in the clause, and it can be divided into three types: active transitive, active intransitive, and stative verbs (Aikhenvald, 1999). In some

Indigenous languages, there are two types of verbal conjugations, as in Wayuunaiki, an analytical conjugation and a synthetic conjugation (Olza & Jusayú, 1978).

Being an agglutinating language, Wayuunaiki depends extensively upon prefixes and suffixes to mark case tense, aspect, mood, number, and gender (Álvarez, 2007). Adjectives as an open class do not exist nor does the copula, although there is a verb meaning "to be" or "to exist" ('*eewaa*'), it behaves like a stative verb and does not have a special function like auxiliary verbs do in other languages (Álvarez, 2005, and 2007).

Wayuunaiki, like Spanish, has two gender values: masculine is for beings that are biologically male (men and male animals) and feminine embraces all remaining entities (objects and humans and animals that are biologically female) (Álvarez, 2007). Masculine is the marked gender in Wayuunaiki (Regúnaga, 2005), whereas it is the default gender in Spanish (Harris, 1991). Nouns in Wayuunaiki can also have another classification as they can be absolute and relative. The absolute class comprises nouns that do not always need to reflect a possessor. The second category refers to nouns that always depend on another entity (Álvarez, 2017). Their difference is noticeable because the absolute nouns don't need a personal prefix: *atpanaa* 'rabbit', whereas the relative nouns are always stated with one of the personal prefixes, for example: ***t****asiipü* 'my nephew', ***j****üpana* 'her liver'.

Regarding pluralization, if just one element marks for plural, that is enough to understand that pluralization is in effect, as in (3), in which the plural suffix *'-irua'* that is used to mark the plural number can be left out (Álvarez, 2007), because the number twenty *piama shiki* is in the phrase and renders redundant the pluralization of the noun *paa'a* 'cow'.

(3) Aya'lajüshi taya piama shikii paa'a(irua).
a-ya'laja-shi taya piama shikii paa'a-irua
0-buy-M I two ten cow-PLU
'I bought twenty heads of cattle'.

(Álvarez, 2017, p. 256)

For temporality, Wayuunaiki has 11 tenses or tempora (Table 2.1), 18 aspects (3 of them have double meaning), as well as 12 aspectual-temporary forms (Knabenschuh de Porta, 1998). Regarding the tenses, the main opposition is present-future. The present tense can never convey the Spanish meaning of future, however it can cover the Spanish simple past, present perfect, and imperfect tenses (Alvarez, 2017, p.72). Wayuunaiki divides the past tenses further to consider more temporal nuances. There is a remote past, close past, and active past.

Table 2.1. Tenses of the Wayuunaiki language

Absolute tenses	Relative tenses	Absolute-relative tenses
General tense	Simultaneous gerund	Previous future
Future	Anteriority gerund	Previous past
Remote past	Completed participle	
Close past	General conditional	
Active past		

Two aspects of the Wayuunaiki grammar that were not mentioned in this chapter, the conjugation systems and pronominal subject realization, will be treated in chapter 5, 6, and 7, respectively.

2.8 SUMMARY

In this chapter, we have provided a description of the relationship between the Guajiros and the Spanish speaking people, as well as of their languages. Although Guajiros and their language have experienced hardship, their future is not uncertain as their population has increased in recent years and there is at least some access to schooling in their language. However, it is also a fact that the Wayuunaiki language is restricted in many cases to home, rituals, Guajiro social

events, and Spanish is the *de facto* language that Guajiros must learn to function in the outer Spanish-speaking world. This situation requires Guajiros to learn Spanish, most of the time without any formal instruction.

Summarizing, this dissertation operates, therefore, with the working hypothesis that Guajiro Spanish differs in non-trivial ways from the local variety of Spanish spoken by native speakers. In order to elucidate the nature of these differences, we will conduct a systematic analysis of the distribution of null and overt subject pronouns, the overt SPP expression and subject-verb agreement in Guajiro Spanish, and we will compare the resultant findings with counterpart phenomena in the local vernacular Spanish benchmark.

3. THEORETICAL FRAMEWORK

3.1 INTRODUCTION

Research done on language variation and change has been mainly carried out within the field of sociolinguistics. However, there has been increasing effort invested in drawing on the combined insights of variationist and theoretical linguistics to elucidate a range of syntactic phenomena. In this sense, there has been a growing initiative to combine observations from the generative and variationist frameworks to analyse and present a cohesive view of syntactic variation that is not necessarily contradictory but rather complementary (Adger & Smith, 2005; Cornips & Corrigan, 2005; Adger & Trousdale, 2007; Martínez-Sanz, 2011).

The idea is that instead of being two incompatible paradigms, they can provide explanations that are meaningful to each other. For example, Meechan and Fooley (1994, p. 82) say that variationist analyses allow us to obtain a genuine observation of grammatical structures as they are used by the speakers. Quantitative analyses can reinforce theoretical explanations of structural analyses by providing information on how language is organized in patterns of usage and across local varieties and dialects of a language. They can provide information about the variable or categorical behavior of an element in a dialect or cross-linguistically, providing linguistic theory with information about the limits of variation (Cornips & Corrigan, 2005).

Theoretical linguistics focuses mainly on language acquisition and can inform variationist sociolinguistics of the acquisition process and how language change is a by-product of the learners using innovative settings for the parameters associated with Universal Grammar (Cornips & Corrigan, 2005, p. 3). From this paradigm, quantitative analysis can also theoretically enhance the exploration of syntactic variation and help us understand how variation is constrained by the nature of possible grammars. Sociolinguistics can also study microvariation in closely related languages

drawing on the apparatus of generative linguistics and using formal explanations of the generative tradition to clarify and interpret patterns of usage found in quantitative analysis (Cornips & Corrigan, 2005).

Building upon the enterprise of presenting a combined analysis for external and internal linguistic constraints, this dissertation integrates these two frameworks to study novel and unanalysed data from Guajiro Spanish. These two framewroks will be used specifically to study the issue of the distribution of NOS and overt SPP expression. Blending these two frameworks will allow us to produce an analysis which, on the one hand, will allow us to understand theoretically whether Guajiro Spanish behaves as a typical null-subject language regarding the expression of null and overt pronominal subjects. On the other hand, the variationist study will allow us to uncover patterns of variation related to the overt expression of the pronominal subject with a view to understanding what motivates these patterns.

In the case of the subject-verb agreement issue, this problem will be analysed from the generative grammar perspective. This type of analysis will inform us how we can explain theoretically the mismatches produced by the bilingual speakers.

In the rest of the chapter, we present a description of the theoretical constructs of Generative Grammar and how variation is approached within this paradigm. The last part of the chapter deals with the key tenets, working principles, and how data is analysed in variationist and comparative sociolinguistics.

3.2 Language Variation and the Generative Grammar Tradition

Up until recently, the study of inter-speaker variation was not of interest for generative grammarians. This was mainly the case because variation was outside the scope of this paradigm. Theoretical linguistics was concerned with the psychological dimension of language (how

language is represented in the mind) and with language acquisition. In other words, with the idea that language is internal to the mind of the speaker. However, there has been a substantial change in that perspective due to the growing interest of theoretical linguistics in studying dialectal variation and how parameters operate in different varieties of a specific language, including developing systems such as those of native and non-native learners, and for the necessity of formal linguistics to account for parametric variation, as seen in Martínez-Sanz (2011) when studying Dominican Spanish.

3.2.1 GENERATIVE GRAMMAR (GG)

Generative Grammar (Chomsky, 1965 and subsequent work; henceforth GG) is the study of language assuming that grammar is an innate human faculty generated by means of subconscious mental processes. In this regard, the aim of this theory is to model those processes in order to come up with a language syntax theory (Carnie, 2012).

One important theoretical aspect of GG is that the subconscious set of procedures by which grammar is generated can be formalized through grammatical rules, principles, or hypotheses. For example, a rule determines the fact that in Spanish the position of the subject can be either preverbal or postverbal.

While generative linguists derive their analyses from native speakers' grammatical intuitions, including their own if they are native speakers of the language being analysed, the nature or source of the data in generative analysis also comes from so-called grammaticality judgment tests/tasks (Henry, 2003) in which informants are required to judge the well-formedness of a sentence, that is to say, whether a sentence is grammatical or well-formed or ungrammatical or ill-formed. By doing this, the generative researcher is trying to tap into the subconscious

knowledge that we as human beings have of our language in order to describe this knowledge and theorize about it.

In this regard, for GG the focus of study is the speakers' language competence or I-language (internal language), or language as a mental property (Adger & Trousdale, 2007). E-language (external language), the language that is manifested through produced or performative speech is less important (Lasnik & Lohndal, 2010) because this language comes out with imperfections, and no matter how big the collected or analyzed data may be, it will not be able to contain all the grammatical sentences that a language can generate (Carnie, 2012).

However, what is the origin of the knowledge we have of language —not of a particular language like English or Spanish, but of the biological faculty— that allows us to speak following certain principles? For generative grammarians, the answer is short: language is innate, and we are born with a capacity that makes our knowledge of language subconscious and universal —what is known as Universal Grammar. Nonetheless, the parameters of the grammar must be still acquired or set. To accomplish this, we benefit from a built-in device: The Language Acquisition Device (LAD), that, when we were children helped us to process and adjust linguistic data to our own specific language, be it Portuguese, Spanish, etc. (Carnie, 2012; Sessarego, 2014).

Because one of the main aims of GG is to establish how grammar is represented in the mind/brain of the speaker/hearer (Henry, 2003), the speakers' own intuitions or grammaticality judgment tests are relevant for determining grammatical competence. Although the model is based on native intuitions and those intuitions are used by the linguist to describe how language is represented in the mind, the reality is that the speakers have intuitions on the language variety they speak or according to their level of bilingualism because the persons who complete the

gramaticality judgement tests come from different backgrounds, which implies that all varieties of languages are taken into consideration.

Even though at the early stages the problem of inter- and intra-speaker variation was not a priority within this framework, Chomsky (1981) acknowledges that languages have principles (e.g., all language have a subject) which are governed by UG and are parametrically variable and sets out a new perspective and expands the research agenda of the theory of syntax. In the following sections, we explore the different generative attempts to explain language variability.

3.2.2 PRINCIPLES & PARAMETERS (P&P)

Language variation started gaining more attention within the generative tradition with the advent of the P&P framework (Chomsky, 1981 and subsequent work). As our linguistic innate knowledge is amenable to formalization, it was proposed the existence of a finite set of general linguistic principles or abstract grammatical rules, and a finite group of parameters that must be set following language specific particularities. In the first case, an example of a principle would be that all languages must exhibit a subject (the so-called Extended Projection Principle), be this covert (implicit) or overt (explicit). In the second case, one parameter that has received ample attention in the literature is the Pro-drop Parameter, also known as the Null Subject Parameter, which accounts for the fact that a language can be [± *pro-drop*] or [± null subject]. A principle is then universal, and a parameter is language-specific and allows for limited syntactic variation (Lasnik & Lohndal, 2010).

Within the P&P approach, one commonly proposed idea is that language variation is found at the parametric level because of the different settings a parameter can have across different languages/dialects. From this perspective, cross-linguistic variation is the result of the setting of parametric values during the acquisition process of a given language or a given language variety.

Henry (1995) uses the parametric approach to account for dialect variation in Belfast English and Standard English. In Belfast English it is possible to have sentences like (4), where the overt subject of the imperative is postposed.

(4) Go you away. Henry (1995, p. 45)

This is possible, according to Henry, because there is more than one grammar for imperatives. In one of these two grammars, the verb has the option of rising or not and, on the other, the subject can optionally rise or stay in situ before Spell-out because variation is tied up to optional movement. From Henry's perspective, variation implies the presence of multiple grammars that result from parameter-setting divergences or choices. This has implications for language acquisition as well. Facing the presence of grammars with different parameter settings, children possibly choose the parameter setting that best reflects or covers most of the data to which they are exposed.

3.2.3 THE MINIMALIST PROGRAM (MP)

With the development of the MP (Chomsky, 1995), the locus of parameters changes. Rizzi (2014, p. 20), following Borer's (1984, p. 29) ideas, argues that "parameters are specified in the functional lexicon of particular grammars". Variation is then a phenomenon linked to the features of the functional categories (Picallo, 2014). Features are at play by determining some syntactic operations and rules (Merge, Agree, Move, Spell-out), when the head of an item enters the syntax (Rizzi, 2014).

In the MP, there is a division on how to approach cross-linguistic variability: on the one hand, microparametric variation and, on the other hand, macroparametric variation as in the traditional parameter view. The microparametric approach is best suited to deal with dialects of the same language (Sessarego, 2014) or closely related languages such as Romance languages

where microparametric choices are very similar. The second approach is more suited to non-genetically related languages, or to compare groups of languages in which differences are better captured by more ample assumptions than by small surface properties (Snyder, 2001). A macro-parameter is "a single characteristic affecting a large number of syntactic categories" (Picallo, 2014, p. 6). However, Baker (2008) acknowledges that both approaches are necessary to explain cross-linguistic differences.

The macro and the micro-parametric approaches to language variation are internalist, since they allocate fundamental importance to syntactic factors as determinants of language variation, particularly to the assumption of parameter-setting and the idea that syntax is feature-driven. Another approach within the MP considers that language variation is affected by external factors such as frequency, probability, and conventionalization, minimizing the impact of syntax and leaving out the prominence of parameters to explain cross-linguistic variation (Picallo, 2014).

One particular approach to intra-speaker language variation is put forward by Adger & Smith (2005). They use the machinery of features within the Minimalist Program to account for variation in Buckie English (Scotland). They argue that variation is produced because of how the lexical items are specified at the end of the derivation. The different phonological outputs or morphosyntactic forms are caused because when the entire feature checking process is finished, the derivation produces items sharing the same interpretable feature specification. However, these items have different uninterpretable specification, which permits different phonological or syntactic realizations but with one possible semantic interpretation. An example is the alternation between *was* and *were* in Buckie English, in which *was* can appear in contexts where *were* is expected (*you was, we was*). They are different phonological productions or allomorphs, but they still convey the same meaning.

Adger & Smith (2005) assume that there is one invariant system. Variation is motivated by the selection of a given lexical item from those which the speaker has available in their lexicon. The choice of a lexical item is motivated by the speakers' interactions, processing requirements, social identity, frequency, and lexical access, and these constraints can be found at the individual and/or community level. The mechanisms allowing variation, however, are located at the individual level.

In subsequent work, Adger (2006) advances another proposal to analyse variation within GG. He introduced the Combinatorial Variability (CV) approach which looks at predicting the frequencies of variants present in a corpus:

> Once one controls for the effect of context, not only do we predict the possibility of stable variability within the grammar, but, on the minimal assumption of random choice of equivalent lexical items, we predict the correct frequencies of variants found in corpora, even though we will not incorporate probabilities into the grammar itself (Adger, 2006, p. 50).

Other proposals to analyse variation within the GG, such as Parrot's (2007), are based on distributed morphology (Halle & Marantz, 1993, 1994). From this perspective, variation comes from competition in the morphology "since there cannot be two outputs for the same input" (Embick, 2008, p. 66).

Barbiers (2014, p. 198) proposes that doubling or the phonological expression of a morphosyntactic morpheme, word, or phrase, twice or more within a clause is "an important source of cross-linguistic and intra-linguistic variation. Syntactic doubling involves local redundancy of features". The sentence in (5a) contains three negations whereas (5b) contains just one. Both

sentences, however, produce the same meaning. The different use of the negation elements in the two sentences is an example of syntactic doubling.

(5) a. At the end of the month, nobody ain't got no money.
b. At the end of the month, nobody has money.

(Barbiers, 2008, p. 2)

More recently, Smith & Cormack (2015, pp. 236-238) propose to contribute to the concept of features in Chomsky's Minimalist Program by using a concise version of Combinatorial Categorical Grammar (CCG). They keep the use of underdetermined features and the encoding of parameters as features of functional heads but argue that microparameters can be encoded in lexical heads. They specifically argue for the use of Combinators as functional heads. Combinators are lexical items that play a role in the syntax and that allow their features to "range over category as well as items from the lexicon." They also propose to make extensive use of diploid (twofold) features. According to these authors, with CCG Combinators, "LF can be produced by categorial Merge, and feature-based displacement can be confined to PF effects", two outcomes that contribute to simplify the task of linguistic analysis and the task of acquisition.

3.3 The diachronic approach to Language Variation in the Generative Tradition

Kroch (1989, 1994) proposes the Competing Grammars Hypothesis (CGH) to explain diachronic morphosyntactic variation. From this perspective, the existence of various options implies the presence of different grammars at the disposal of the speaker and not the realization of competing forms within the same grammar. These options are diachronically unstable, a situation which eventually produces the disappearance of one of them.

Kroch (1994) proposes that competing variants can be assumed as morphological doublets, which arise thanks to dialect/language contact and other sociolinguistic factors or processes. Mentally, "there is more than one system of grammatical knowledge in the head of the native

speaker, and variation then boils down to decisions that the speaker makes about which grammatical output to choose" (Adger & Smith, 2005, p. 15).

The decline of a competing option is due to processing factors and economy. First, the more complex the form, the less it is used, which in turns disfavours its frequency. Second, the lexicon rejects the storage of already existing equivalent forms to maximize capacity. Then, the rising of a unique form can be explained in psycholinguistic terms.

For instance, in English, *do*-support became categorical after five centuries of its first appearance, the period during which it existed in competition with simple forms (verb-subject-not) to produce negatives and questions as in (6).

(6) a. How like you this sonnet?
b. How do you like this sonnet?

(Kroch, 1989, p. 143)

In (6a), the question is asked without the help of the auxiliary, and in (6b), *do*-support is used to ask the question, the norm in Modern English.

In L2 acquisition the Competing Grammar Hypothesis is introduced to account for syntactic optionality during parametric change (Zobl & Liceras, 2004). According to this hypothesis, learners of an L2 internalize two or more conflicting grammatical representations when they are confronted with ambiguous data. These representations compete with each other to gain usage, which creates a situation of apparent syntactic optionality.

Using CGH as a framework helps to explain important aspects of L2 development and especially the optionality that arises during the acquisition of new parametric options as well as how L1 and L2 acquisition differ. In this sense, parametric change in diachrony and L2A share some fundamental features that support the idea that L2 acquisition is not neccesarily faulty because of the impairment produced by the effects of a critical period, as explained by Beck, 1998;

Clahsen & Muysken, 1989, Hawkins & Chan, 1997; Tsimpli, 2003; Tsimpli & Roussou, 1991 to account for the difficulty of L2 learners to acquire formal features or values which are non-existent or different from their L1.

Regarding our dissertation, for the study of the distribution of NOS and verb class we will take into consideration the property that defines the Null Subject Parameter, specifically the absence or presence of referential subjects, and for the study on subject-verb agreement we will use the concept and role of features within the Minimalist Program.

3.4 VARIATION THEORY

Variation theory is the study of language as it is manifested or used in its natural social context. In order to carry out variationist analyses, this theory has combined methodological elements of several disciplines, such as anthropology, sociology, linguistics, and statistics (Labov, 1972; Sankoff, 1982; Poplack, 1993). Variation theory is founded on three core constructs about language: orderly heterogeneity, language change, and language's social dimension (Tagliamonte, 2006, p. 6-7).

The first core idea suggests that the speaker has "alternative ways of saying the same thing" or expressing a similar grammatical function within her speech and this alternation is statistically regular, the notion of *inherent variability* (Labov, 1969). However, language variation is not random, but patterned and constrained by internal and external determinants (Labov, 1971). Speakers may have many possibilities to convey a given meaning, nonetheless, those speakers' options may be determined by analysing the distribution and linguistic conditioning of competing variants in discourse (Kiesling, 2011).

The second idea advocates the notion that language is dynamic and not a static system, in the sense that language changes and is not invariant (Poplack & Tagliamonte, 2001, p. 92). The

language spoken today differs from the language spoken 300 years ago. Let's illustrate this with the use of *haya/haiga* (1st and 3rd person singular subjunctive of *haber*) in Spanish: the *haiga* form has been displaced by *haya* and its use is nowadays common only among rural and/or uneducated people (Johnson & Barnes, 2013), although this form was completely accepted some centuries ago.

The third core concept refers to the existence of sociolinguistic patterns, which are the correlations between features of the linguistic environment with extra-linguistic variables (e.g., age, sex, and ethnicity).

Tagliamonte (2006, 2013) points out that the methodology, analysis, and explanation of the linguistic phenomena from the variationist perspective must take into account certain paramount concepts including the notion of the vernacular, the speech community, form/function asymmetry, the linguistic variable, the quantitative approach, the principle of accountability, and the delimitation of the variable context. We discuss these concepts in greater detail in the following paragraphs.

For variationist sociolinguistics, the vernacular, or everyday speech, is the crucial object of study. This is the kind of language that is supposed to be free of hypercorrection, which the speaker acquires early in preadolescence, and offers the most systematic data for linguistic analysis (Labov, 1972; 1984). In order to gain access to the vernacular, the variationist linguist must enter the speech community and participate as an observer or as a participant. This allows them to assess the way the speakers use language in their particular sociocultural situations (Weinreich, Labov & Herzog, 1968).

Variation theory also claims an asymmetrical relationship between form and function. This relationship refers to expressing the same referential meaning or a similar grammatical function

via different forms or a form having different functions. (Labov, 1972; Tagliamonte, 2006). In this sense, there are cases where there is "a difference in linguistic form but no change in the linguistic context and no apparent change in meaning" (Walker, 2010, p. 8). The idea of the linguistic variable or the use of different linguistic forms to say approximately the same thing is central to variationist linguistics (Labov, 1969; Tagliamonte, 2006; Walker, 2010). The set of variants available to the speaker are located at the different levels of the language: phonological, morphological, lexical, syntactic, pragmatics/discourse, and suprasegmental levels (Kiesling, 2011). For example, in Spanish the pronominal personal subject —a syntactic variable— can be expressed by means of an overt or null subject, comprising two variants.

Once a linguistic variable is recognized, its variable contexts must be circumscribed or delimited. This procedure necessitates finding and defining all possible contexts in which there is variation or not; the so-called variable context or the envelope of variation. For instance, the subject personal pronoun is categorically null in Spanish in chronological or time period sentences as in (7), since this kind of phrase does not need a denotational subject, but it is variable in many other contexts such as in contrastive sentences like (8), where the subject can be present or absent.

(7) a. **Ø** Hace treinta minutos que espero el autobús
Make.3S thirty minutes that wait.1S the bus
'I have been waiting for the bus for thirty minutes.'

(8) a. Tengo 5 años en Ottawa, pero (yo) no hablo inglés
Have.1S 5 years in Ottawa, but (I) not speak.1S English
'I have been living in Ottawa for 5 years, but I don´t speak English.'

The context of a variable can be defined following a form-based or a function-based approach. Walker (2010, 2015) says that a form-based approach is most suitable for phonological and lexical variables because it is easier to comply with the requirements of semantic equivalence. For grammatical or discourse variables, the approach is to take a linguistic function and observe

all the possible forms that embody this function. When semantic equivalence is invoked, Sankoff (1988) proposes that differences among competing forms may be neutralized in discourse. That is to say, a form can convey different meanings, but in practice, all possible meanings are not appropriate for the context, the speaker, or the interlocutor (Walker, 2010b). Neutralization in discourse is said to pave the way for linguistic change (Sankoff, 1988).

Furthermore, when the variable context is delimited, the produced or unproduced variants must be considered, in other words *the principle of accountability* should be respected (Labov, 1972b). In this sense, one must (i) "identify the total population of utterances in which the feature varies", (ii) "decide on the number of variants which can be reliably identified and set aside those environments in which the distinctions are neutralized", and (iii) "identify all the sub-categories which would reasonably be relevant in determining the frequency with which the rule in question applies." (Labov, 1969, pp. 729-730).

Language variation can also be systematically, and quantitatively handled and diverse linguistic and social contexts can be taken into account through multivariate analysis (Hazen, 2011). The quantitative method is then a prominent aspect of variation analysis (Tagliamonte, 2006), and whenever a choice of a variant has been motivated by language internal or external factors, statistical analysis can be used to find and explain this linguistic variability (Sankoff, 1988). The tool most commonly used to "uncover the latent grammatical structure in produced language" is variable rule analysis (Roy, 2008, p. 1), which we explain more in detail in the next section.

3.4.1 VARIABLE RULE ANALYSIS (VARBRUL)

Variation theory assumes that language varies in an orderly manner and this variation is constrained or influenced by internal and/or extra-linguistic factors including a speaker's demography, surrounding syntactic, morphological, and phonological contexts, and discourse-related functions (Tagliamonte, 2006). Naturalistic data is also unevenly distributed as "the number of occurrences of each context depends on its relative frequency in discourse. Hence the number of cases per context is highly variable and many combinations of factors may not occur at all" (Sankoff, 1988, p. 986). Facing this situation, variable rule analysis is the method commonly used in variationist sociolinguistics to quantitatively analyse this type of data and understand the linguistic phenomena (Young & Bayley, 1996; Tagliamonte, 2006).

According to Sankoff (1988, p. 984), the use of variable rule analysis can be invoked when there is an evident choice among the variants of a sound, structure, or word. Although it may look heterogeneous, random, or free, language variation is governed by and implicit set of rules and the speakers' selection of a form is systematic.

Sankoff (1988, p. 987) says that the main objective of a variable rule analysis is "to allocate contextual effects on the choice process among the different cross-cutting factors which make up the contexts, in a way which is able to correct both for dependent factors and for interacting factors". In this sense, variable rule analysis provides three crucial lines of evidence to interpret data: (i) it gives us the statistical significance of the explanatory variables or factor groups, the relative contribution of independent social and linguistic variables hypothesized to condition variant choice, (ii) the magnitude effect of the significant independent variables or groups as assessed by the range value (e.g., the difference between the highest and lowest factor weights in a significant factor group, expressed as a whole integer), which helps to establish the hierarchy of

significant factor groups, and (iii) the constraint hierarchy of the levels of a group, from the highest to the lowest, that is a window on a portion of the underlying grammar (Poplack & Tagliamonte, 2001, p. 92).

In summary, variable rule analysis allows us to extract data regularities and tendencies and it provides information about what contextual factors under analysis (dis)favor the choice of one variant over another and what their relative strengths are (Tagliamonte, 2013).

3.4.2 COMPARATIVE SOCIOLINGUISTICS

In their exploration of what they refer to as Early African American English, Poplack and Tagliamonte (2001) draw on a comparative sociolinguistics methodology which integrates the comparative approach of historical linguistics, dialectology, and variationist sociolinguistics. According to Tagliamonte (2013b, p. 128), "comparative sociolinguistics in particular is concerned with the relationship of linguistic variation in one body of materials to another".

Comparative sociolinguistics is then aimed at comparing elements/features of one language variety to another. First, by finding out and examining systematic patterns of variability through variationist analysis. Second, by situating those patterns of variations in relation to other close or related language varieties. For example, the conditioning factors obtained for one dialect are compared or contrasted to the conditioning factors of another variety or dialect.

The comparison is done by "using the lines of evidence provided by statistical modelling" (Tagliamonte, 2013b, p. 130) in a chosen adequate conflict site or structure that is amenable to comparative analysis, on the basis that this structure or element varies between languages or dialects as different environmental constraints make grammatical systems different (Poplack & Meechan, 1998, p. 132). The idea of conflict site is very important for this thesis as this is the key heuristic for the work that we undertake here in which we will compare two conflict sites or

structures: the realization of the pronominal subject and subject-verb mismatches to determine if there is contact-induced change.

If there are differences, the comparative analysis will reveal them in the observed patterning of the features under analysis (Poplack, 2018, p. 26). The comparison of constraint hierarchies or ranking of factor weights within a factor group will help us to determine if the systems under comparison share the same linguistic conditioning on variable usage and the same direction of effects. We will be able to find as well if there are non-trivial parallels in the 'grammar' underlying patterns of variation.

Divergences in the constraint hierarchies are important distinctions that inform us of the presence of different linguistic conditions in the comparison varieties. In this respect, dissimilarities in the ranking of constraints provide possible evidence that the language varieties under analysis are not part of the same system. If the constraint ranking of variables is parallel but the strength of a significant effect, as gauged from the range value differs, this could mean the feature under investigation is at a different stage of development across the varieties (Tagliamonte, 2013b).

3.5 SUMMARY

In this chapter we have presented the arguments for a cohesive approach to the study of syntactic variation and how the generative paradigm and the variationist framework can be exploited to elucidate variation in a language contact situation. We have described how Generative Grammar and variationist linguistics see language - the former as an internal aspect of the human mind and the latter as an external element amenable to observation. Although each theory approaches language differently and differs in its data collection processes as well as in their analytical procedures, both disciplines offer perspectives that can enrich our understanding of the

processes and outcomes of language contact. In this regard, this dissertation makes use of both; the principles and proposals of the theory of Generative Grammar to analyzse (NOS), variationist sociolinguistics to study overt (SPP) expression, and it makes use of the theory of features within the Generative Grammar to investigate subject-verb agreement morphology.

4. THE STUDY: DATA AND PARTICIPANTS

4.1 INTRODUCTION

In this chapter, we will describe the monolingual and bilingual participants and the data elicitation techniques used in this dissertation. We will describe how we approached the Guajiro community and how we collected spoken language data using a sociolinguistic interview protocol (Labov, 1984).

4.2 SOURCE AND NATURE OF THE DATA

The research described here is based on spontaneous speech data that was elicited with the help of standard sociolinguistic interviews (Labov, 1984) and digitally recorded.[9] For this, an interview protocol based on different topics was elaborated (see Appendix A). The interview protocol consisted of a questionnaire organized in several conversational modules, but open enough to allow the participants to converse about any topic. The questionnaire was designed with the aim of promoting informal conversations and eliciting narratives dealing with the speakers' personal experiences. This type of discourse is desired as it triggers vernacular language use which is necessary to study patterns of usage **(**Lavob, 1981, p. 8)**.** For example, when a participant narrates a surgical intervention in one of our interviews, he is more focused on telling the chain of events than in monitoring his language. In this kind of situation, the participants shift style and use the vernacular. The interviews also allow us to obtain "the full range of demographic data necessary for the analysis of sociolinguistic patterns" such as language history, age, use of languages, education, etc. (Labov, 1984, p. 8).

[9] We used an Olympus digital voice recorder, model VN-702PC.

To elicit the data, we traveled to the town of Uribia, Colombia, to interview bilingual speakers and to the city of Maicao to interview monolingual Spanish speakers.[10] In Uribia, a local person helped us get in contact with Wayuunaiki speakers. We interviewed the participants entirely in Spanish and recorded them, on average, for around 17 minutes each.

The monolingual Spanish speakers were interviewed for 22 minutes each on average. The entire sample used to extract tokens from the oral interviews is composed of one hour and fifty-two minutes (1:52) for the monolingual group interviewed in Maicao and seven hours and forty minutes (7:40) for the bilingual group interviewed in Uribia. The number of tokens varied (Table 4.1) according to the requirements of each study, as we explain in chapters 5, 6 and 7.

Table 4.1. N of tokens analysed per study

Study	Bilingual tokens	Monolingual tokens	Totals
SPP analysis by verb type	2448[11]	481	2929
SPP variationist analysis	3493	705	4198
Subject-verb mismatches	2336	391	2727
Totals	8278	1577	9854

As mentioned, the protocol that we used for the interviews was composed of several sections. These sections were made up following Martínez-Sanz's conversational modules (2011, pp. 270-271). This questionnaire worked as a guide to structure the interviews and to follow a certain order. However, the interviews were conducted to elicit conversational and spontaneous

[10] Uribia is the indigenous capital of Colombia. The major part of its population is composed of Guajiros (Gobernación de la Guajira, 2012). More than 90% of the population in this municipality live in rural settlements (Martínez, 2019).

[11] The discrepancy between the token number for SPP variationist analysis, the NOS analysis by verb type and the subject-verb agreement study reflect the fact that the three analyses took into consideration different verbal forms. In the case of subject-verb agreement only verbs in the present tense were analyzed to avoid ambiguous forms like *yo comía*, *usted comía tú comía(s)* in the imperfect. In the case of the SPP analysis, we included tokens in questions and relative clauses, whereas for the NOS analysis by verb type only declarative clauses were included. Phrases with modals, and different types of periphrases were not included because it would have added complications to the clear-cut differentiation between the different types of verbs (i.e. stative vs. active, etc.).

.

style responses. We asked participants about their personal experiences and their opinions about the Wayuunaiki and Spanish languages, their community and social practices, their family, their work, their travels, and uncommon events in their town, all the while giving them the flexibility to expand on other topics of their interest.

The data collected through the interviews came in the form of narratives, comments, and simple answers to questions. The interviews were later transcribed following Spanish standard orthographic conventions and the criteria established for the transcription of the *Corpus sociolingüístico del habla de Mérida-Venezuela*[12] (Table 4.2) (Domínguez & Mora, 1998).

Table 4.2. Conventions for the transcriptions

Full stop	period	.
Questions	question marks	¿?
Exclamations	exclamation marks	¡!
Pauses	comma	,
Hesitations and false starts	ellipsis	…
Elision of sounds	parenthesis	( )
Incomprehensible words or passages		XXX
Convention for the transcription of interjections	ajá, mjm, mm, uf, uy, ay, ey, eh.	
Repetitions were kept. For example, *hola, hola.*		
Contractions were transcribed as they were produced: *pal* (para el).		
When the participants used words in Wayuunaiki, we transcribed them as we heard them.		
Double brackets were used for non-linguistic sounds: laughing, coughing, third person intervention, interruptions in the recording.		

Since our analysis is mainly morpho-syntactic, a detailed transcription of phonetic phenomena was not implemented except for verb endings. In this case, the variable elision or weakening of the /s/ and /n/ sounds was indicated as we needed that information for the analysis

[12] *Sociolinguistic corpus of spoken Spanish of Merida-Venezuela.*

of the verb-subject mismatches in chapter 7. In (9) and (10), we present an interview excerpt of two bilingual speakers. In these two extracts, the interviewee is identified with the first three letters of their name.

(9) **Entrevistador**: ¿y qué hacen tus otros hermanos? ¿A qué se dedican?	**Interviewer**: and what do your other siblings do? What do they do for a living?
CAT: mis hermanas; hay unas que están estudiando, hay unas que ya viven con los maridos, cada quien (es)tán regados, dos hermanos míos viajan pa(ra)… hay unos que viajan pa(ra)… mis hermanos que viajan pa(ra) Maracaibo y unos que (es)tán trabajando.	**CAT**: my sisters, there are some who are studying. The others are already living with their husbands, each one lives on her own. Two of my brothers travel to… some of them travel to… my brothers travel to Maracaibo and some others are working.
Entrevistador: ¿qué hacen en Maracaibo?	**Interviewer**: what do they do in Maracaibo?
CAT: los hermanos míos (es)tán, son de chofer, manejan carro y los otros son vigilantes o trabajan en parcela por allá.	**CAT**: my brothers are… they are drivers, they drive cars. The other ones are security guards, or they work on farms over there.
Entrevistador: ok.	**Interviewer**: ok.
(10) **Entrevistador**: ok, ok, eh… háblame de las costumbres, de las tradiciones que hacen en las rancherías ¿cómo son los bailes y las fiestas?	**Interviewer**: ok, ok, eh.. tell me about your customs and traditions you have in the rancherías. How are the dances and parties?
ALB: o sea, el baile más común ahí es la yonna, la yonna del wayuu que e(s), que e(s) la chichamaya.[13]	**ALB**: I mean, the most common dance there it is the yonna. Guajiro's yonna, which is the chichamaya.
Entrevistador: ¿cómo, cómo, cómo es eso?	**Interviewer**: How, how, how is that?
ALB: o sea, un… o sea… o sea una chichamaya, o sea, eso lo ve en sueño primero pa(ra) esa fiesta, un sueño, o un encierro.	**ALB**: I mean, a..., I mean... I mean… a chichamaya, I mean…, that is what they first see in their dreams, for the party: a dream, a confinement.
Entrevistador: ¿un sueño, un encierro, cómo así? ¿cómo un encierro?	**Interviewer**: a dream, a confinement, how come? what is a confinement?

[13] The *yonna* or *chichamaya* is a Guajiro traditional dance. This dance is carried out to celebrate an important event in the community. For example, to interpret a dream, to celebrate the passage to adulthood of a girl or to celebrate peace between two families or two clans.

ALB: o sea, por ejemplo, en la ley… en la ley del… en la ley del wayuu, o sea, por ejemplo, un pela(d)o o un viejo se, se accidenta lo, lo encierra(n), lo encierra(n) ahí un mes, dos meses y después cuando va a salir ahí, ahí hacen la chichamaya, la yonna pal que sale.	**ALB**: I mean, for example, in the law… in the law… in the Guajiro's law, I mean, if a young man or a an old man has an accident, they confine him, they enclose him during a month or two and later, when he is about to quit the confinement, they dance the chichamaya, the yonna in honour of the person who quits the confinement.

4.3 The Participants

For this study, we use data collected from 27 native Wayuunaiki speakers and five monolingual Spanish speakers. The bilingual and monolingual participants were all recruited in La Guajira Department, Colombia, and they were interviewed in their communities, place of work, or their homes. During the recruitment process, the participants were told they would participate in a confidential study and that they could have access to the data and results once the research was finished. For the illiterate participants, we orally explained the objective of our study to them and when it was needed, an intermediary acting as an interpreter explained in Wayuunaiki the aim of the study and the relevance and importance of their participation. To comply with the Research Ethics Board of the University of Ottawa, all the participants completed and signed a consent form (Appendix B).

To determine the participant's language experience and usage, they were asked to complete a language background questionnaire. With this instrument, we asked them about the contexts and situations where they used the languages they spoke. In this same questionnaire (Appendix C), we collected data related to age and gender. From the interview, we obtained information for two more variables: the onset of learning of Spanish in the case of the bilingual speakers, and literacy for both groups.

4.3.1 THE BILINGUAL PARTICIPANTS

From our initial contact with the Guajiro people, we realized that it would be difficult for us to access the community since we were identified as non-community or outgroup individuals and did not speak the language. To overcome these barriers, we earned the trust of a local contact who helped us negotiate interviews with other Guajiros in the local community. This intermediary was our first interviewee, and we capitalized on this first interview to recruit additional participants using the tried-and-tested 'snowball' protocol (i.e., the friend-of-a-friend approach). This approach is important to help a stranger/outsider gain access to the community and participate in the community's exchange and obligations relationships. It also helps researchers to increase their chances of success to interact and observe the community (Milroy, 1980, p. 53).

This intermediary took us through the town and introduced us to his family, some of whom accepted to be interviewed. As we needed a substantial number of participants, we walked around the town, introducing ourselves. The intermediary talked to other Guajiros on our behalf to explain the reasons for our presence there as well as our objectives. As some Guajiros feel threatened by alijunas, they sometimes answer in Wayuunaiki even though they could speak Spanish. Our intermediary was a Guajiro himself and therefore helped us overcome the potential problems involved in the negotiations. His presence during the interviews was also convenient to generate confidence and mitigate hypercorrection from the participants and gain access to the vernacular. Milroy (1980, p. 25) considers this situation to be a "matter of good luck" since the presence of another local during the interview helps to reduce anxiety and the peculiarity of the situation.

The individuals who accepted to participate in our study were native Wayuunaiki speakers. They all learned Wayuunaiki first and Spanish later. Some of them learned Spanish when they

started attending primary school and others learned it informally by being in contact with Spanish speakers or because a more proficient friend or relative had taught Spanish to them.

The bilingual participants had different levels of Spanish proficiency. To assess their knowledge of Spanish, three monolingual Spanish speakers from Colombia acted as judges and listened to audio clips of these speakers, classifying them according to their perception of the participants' Spanish proficiency on a 9-point scale, which was subsequently used for classifying speakers as basic, intermediate or advanced (Table 4.3).

Table 4.3. Instrument to assess bilingual speakers' proficiency in Spanish[14]

	1	2	3	4	5	6	7	8	9
Fluency									
Vocabulary									
Pronunciation									
Grammar									

The judges had to rate each speaker with respect to four different components: fluency, vocabulary knowledge, pronunciation, and grammar. The judges had to mark each one of these sections between 1 (the weakest performance) and 9 (the strongest performance). A score between 1-9 was assigned for incipient proficiency, 10-18 for intermediate speakers, and 19-27 for advanced speakers. This assessment produced 13 speakers for the intermediate group and 14 speakers for the advanced group. None of the speakers were rated as having a low proficiency in Spanish according to this assessment tool.

[14] The instrument used to assess the Wayuunaiki spekaers' proficiency in Spanish was designed by us. The idea was to create an instrument that would permit us to obtain their proficiency by assessing four different components in a scale from 1 to 9.

Table 4.4. Sociolinguistic information of the bilingual participants

Participant	Age	Sex	Literacy	Spanish Proficiency	Onset of learning of Spanish
ORL	18	M	Literate	Advanced	Early
GEI	18	M	Literate	Advanced	Early
CRI	18	M	Illiterate	Intermediate	Late
CAT	28	F	Literate	Advanced	Late
FAB	19	M	Literate	Advanced	Late
ELI	20	M	Literate	Advanced	Late
MIS	20	M	Illiterate	Intermediate	Late
MIG	21	M	Literate	Advanced	Early
NER	21	M	Literate	Intermediate	Early
ADE	22	M	Literate	Advanced	Late
ZUR	23	F	Literate	Advanced	Early
ALB	23	M	Literate	Advanced	Late
RAM	23	M	Illiterate	Advanced	Late
ABR	28	M	Literate	Advanced	Late
CON	28	F	Literate	Advanced	Early
MAR	30	F	Illiterate	Intermediate	Late
EST	31	F	Illiterate	Intermediate	Late
ANE	32	F	Illiterate	Intermediate	Late
ISR	32	M	Literate	Intermediate	Late
ROS	33	F	Illiterate	Intermediate	Late
BEN	37	F	Illiterate	Intermediate	Late
SAR	40	F	Literate	Advanced	Early
LUZ	42	F	Illiterate	Advanced	Early
ANI	43	F	Illiterate	Intermediate	Late
NAR	52	M	Literate	Intermediate	Late
MAI	55	F	Illiterate	Intermediate	Late
JOS	60	M	Illiterate	Intermediate	Late

Regarding social data, the age of the youngest participant is 18 years (three individuals), and the oldest is 60 years. The mean age is 30. There are 15 males and 12 females. Concerning literacy in Spanish, 15 participants are literate and 12 are illiterate. We did not ask the bilingual participants about literacy in their native language. In any case, Wayuunaiki is still mainly oral.

To classify the speakers by onset of learning of Spanish, we use the term “early bilingual” to designate individuals who started learning their second language before the age of six and a “late bilingual” to those who started learning their L2 after the age of six, an age at which L2

learners begin losing their child-like linguistic abilities (Hytelstam, 1992 in Montrul, 2008). The information used to classify the speakers as early/late bilingual was self-reported by the participants, who were asked to provide the age at which they started learning Spanish.

The domain of usage of Wayuunaiki was also assessed and the participants were asked with whom they used Wayuunaiki and Spanish. The participants reported that they spoke their native language in most of their daily interactions: with their parents, children, partners, and friends (see Figure 4.1). In some cases, they reported speaking only Spanish when they have an alijuna monolingual partner or when they interact with an alijuna monolingual friend. In other situations, they use both languages, such as at home when the two languages are spoken indiscriminately with a bilingual partner.

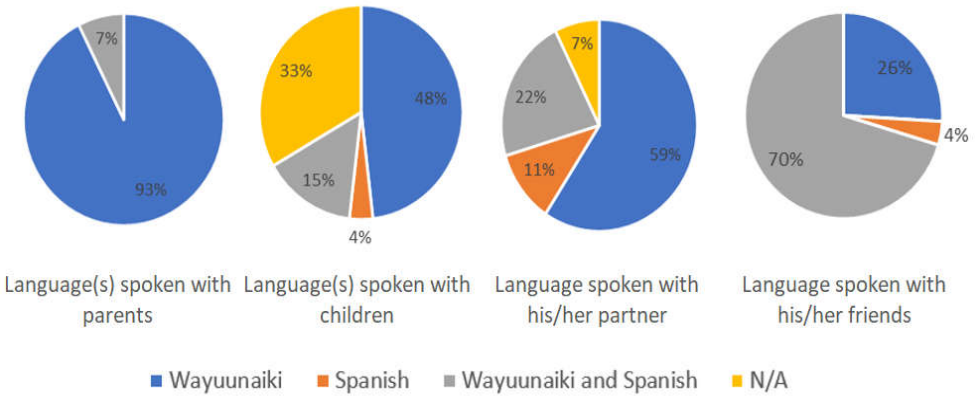


Fig. 4.1. Participants' usage of Wayuunaiki and Spanish[15]

[15] N/A means the participant did not have children of their own or did not have a partner, hence they did not provide any information about these types of interactions.

4.3.2 THE MONOLINGUAL PARTICIPANTS

In order to have a control group representing the variety of Spanish in which the Guajiros are immersed, we interviewed monolingual Spanish speakers in the city of Maicao, a Spanish-dominant city in La Guajira.[16] The participants reported speaking and carrying out their daily activities entirely in Spanish. Nonetheless, they encounter Guajiros on a regular basis. One of the participants even indicated to know a few words, but nothing meaningful to start or hold a conversation.

Two of the participants are male and three are females (Table 4.5). The youngest participant is 21 years old and the oldest is 69 years old. The mean age is 38. Four participants are literate, and one is illiterate. They were born in La Guajira and have lived in Maicao for most of their lives.

Table 4.5. Monolingual participants' social information

Participant	Age	Sex	Literacy	Education	Language
YAN	21	F	literate	Some university education	Spanish
YOE	22	F	literate	Some university education	Spanish
YAQ	40	F	literate	High school completed	Spanish
END	34	M	literate	High school completed	Spanish
FAV	69	M	illiterate	None	Spanish

For the data collection of the monolingual speakers, we used the same questionnaire that guided us during the interviews with the bilingual speakers. We also explained the nature of our research and our reason for being in La Guajira, and they accepted to help us out. In (11), we present a transcription excerpt of a monolingual Spanish speaker.

[16] The city is mainly Spanish speaking and is in the Middle Guajira. However, the majority of the population is alijuna. Around 50,000 people are of native origin in an area of more than 150,000 inhabitants (Gobernación de la Guajira, 2012).

(11) **Entrevistador**: ¿cuánto pagaste de servicio?	**Interviewer**: how long did you stay in the Army?
END: un año y medio, 18 meses.	**END**: a year and half, 18 months.
Entrevistador: ¿y el hecho de que usted estuviera estudiando no le perdonaron para que no fuera al servicio?	**Interviewer**: and the fact you were at school, did not help you avoid conscription?
END: no, nada, porque estaban agarrando así en caliente.	**END**: not at all, they were recruiting on the spur of the moment.
Entrevistador: ok, ¿y alguna experiencia espectacular en el ejército? ¿algo que le pasó en el servicio o tuvo algún problema, un enfrentamiento, tuvo alguna…?	**Interviewer**: ok, did you have any special experience in the Army? Did something happen to you? a problem, a confrontation, did you have any…?
END: sí, hay un dato muy curioso que nosotros los muchachos que estábamos acá, en este lado de la Guajira, yo estuve preste en el batallón Cartagena, batallón que está en Riohacha, nosotros pensábamos que aquí en la Guajira no había guerrilla y nosotros andábamos normal, porque pensábamos: no, aquí no hay guerrilla, pero hubo una vez que sí, nos dimos cuenta que nos dispararon; ey, hay guerilla, andábamos todos pilas, la guardia; ey pilas. Eso fue algo que nos marcó la vida así porque cambiamos de una.	**END**: yes, there is a curious event that happened to us guys here, on this side of La Guajira, I served in the Cartagena battalion, this battalion is located in Riohacha, we thought that here in La Guajira there were not guerillas and we were behaving as we normally used to do because we thought there were not presence of guerillas, but there was one time we realized they were here when they shot at us, hey, there are guerrillas and since then we have all been careful, we patrolmen, hey, attention. That was something that affected our lives and that experience made us change immediately.

4.4 Data Analysis

The monolingual and bilingual data of this study are subject to a comparative analysis. As explained earlier, the distribution of NOS will be studied from the Generative Grammar, the overt SPP expression from the comparative sociolinguistics perspective, and the issue of subject-verb agreement from the Generative Grammar framework only.

For the generative analysis, we will obtain the production rates of subject-verb agreement mismatches of the bilingual group in order to determine whether these mismatched sequences are

typical of Guajiro Spanish or a replication of the Spanish vernacular benchmark. For the NOS study, we will obtain the rates and distribution of overt and null subject pronouns for the bilingual and monolingual speakers in order to compare them. To get these rates, we will tabulate the tokens according to whether the pronominal subject is null or overt and whether the pronominal subject is realized with a stative or an active (transitive/intransitive) verb.

For the sociolinguistic analysis, we will conduct a Varbrul analysis to obtain the ranking of constraints associated with the effect of independent variables on variable SPP realization in the monolingual and bilingual groups. The rates of usage of overt SPPs for each group will also be obtained and compared.

4.5 SUMMARY

In this chapter, we have provided information about the participants' profiles and the situations and contexts in which they use one or the two languages. We have also described how we were able to access the Guajiro community to carry out the interviews. The bilingual group is composed of advanced and intermediate speakers of Spanish and represents a rather adequate sample for studying the linguistic phenomena targeted in this dissertation. Even though access to the bilingual speakers was difficult at the beginning, the attempt proved to be extremely rewarding as we now have data that allows us to uncover how the Wayuunaiki-Spanish contact situation has contributed to the formation of a contact variety of Spanish. For the monolingual group, we have five native speakers. The data from the native speakers serves as the benchmark for the comparative analysis, in which we intend to determine whether the grammar of the bilingual and monolingual Spanish speakers under study differs or coincides.

5. THE USE AND DISTRIBUTION OF GUAJIRO SPANISH NULL AND OVERT PRONOMINAL SUBJECTS (NOS)

5.1 INTRODUCTION

The aim of this chapter is two-fold: (i) to analyse whether Wayuunaiki's two sets of conjugations, the *synthetic* and the *analytical*, influence the rate and distribution of null and overt pronominal subjects (NOS) and how verb class affects Guajiro Spanish, and (ii) to compare the results of the bilingual group with the local monolingual baseline to find out whether there is a relationship between the type of conjugation and the distribution of null and overt pronominal subjects (NOS) and whether and how that relationship is manifested in the case of the bilingual and the control groups.

With these aims in mind, this chapter is organized as follows. First, an overview of the Null Subject Parameter (NSP) regarding Spanish is presented. Second, the functioning of specific properties of this parameter is compared in Wayuunaiki and Spanish. Third, the research questions and hypotheses of this chapter are presented. Fourth, the results of the data analysis are discussed. The chapter closes with a summary and the conclusions of the investigation.

5.2 THE NULL SUBJECT PARAMETER

In order to investigate the distribution of null and overt subject pronouns in Guajiro Spanish, it is important to provide a description of the syntactic properties attributed to the languages that, in the classical inception of the Null Subject Parameter (NSP) (Perlmutter, 1971; Taraldsen, 1978; Chomsky, 1981; Rizzi, 1982), behave like Spanish or Italian. These properties have been described as follows:

A. Omission of referential subject pronouns. That is to say, the subject can be present or absent and this absence is grammatically correct:

(12) Ø estudiamos todos los días *Ø study every day
study$_{[1p-p]}$ all the days
'We study every day'.

B. Expletives must be null:

(13) Ø Nieva mucho en Ottawa en invierno. *Ø snows a lot in Ottawa in winter
snows$_{[3p-s]}$ a lot in Ottawa in winter
'It snows a lot in Ottawa in the winter'.

C. Presence of postverbal subjects in declarative sentences:

(14) Nelson ha cocinado hoy. ~ Ha cocinado Nelson hoy.
Nelson has$_{[3p-s]}$ cooked today *Has cooked Nelson today
'Nelson has cooked today'.

In this original definition of the parameter, properties such as the *that-t filter* violation and long *wh*-movement were also included. Apart from the previous properties, other interpretative constraints associated with null subject languages have been proposed. For example, Montalbetti (1984) proposed the Overt Pronoun Constraint (OPC) and Carminati (2002) put forward the Position of Antecedent Hypothesis (PAH). The Overt Pronoun Constraint states that null and overt subjects are interpreted differently in Spanish as bound variables. Overt pronouns and *pro* have the same referential capabilities although overt pronouns must be interpreted with a specific reference. On the other hand, if the antecedent varies, as in the case with interrogative pronouns or if the subject of the main clause is indefinite or quantified, the null subject pronouns can be bound to this reference as in (15). Overt subject pronouns of the subordinate clause cannot be bound to this subject, though. In this sense, L2 learners of Spanish must learn the properties of the OPC to interpret and use correctly the co-reference of overt/null subjects in subordinate clauses.

(15) a. Todo estudiante$_i$ cree que *pro$_i$ es* inteligente.
Every student$_i$ thinks that (he $_i$) is intelligent.

b. Todo estudiante$_i$ cree que él$_{*i/j}$ es inteligente.
Every student$_i$ thinks that he $_j$ is intelligent.

(Camacho, 2013, p. 30)

The Position of Antecedent Hypothesis states that null-subject languages have a processing preference to resolve intrasentential anaphoras. Null subjects as in (16) prefer an antecedent in the specifier of the inflected phrase (SPEC IP), the subject position. Overt pronouns as in (17), on the other hand, prefer the antecedent that is lower in the clause, the complement position. The implication of antecedent assignment is that L2 learners of Spanish do not necessarily follow the same strategies that L1 speakers of Spanish apply due to cross-linguistic variation and different processing strategies.

(16) Cuando Mario$_i$ llamó a Juan, ∅$_i$ había apenas terminado de comer.
When Mario$_i$ telephoned to Juan, (he)$_i$ had just finished eating.

(17) Cuando Mario llamó a Juan$_i$, él$_i$ había apenas terminado de comer.
When Mario telephoned to Juan$_i$, he$_i$ had just finished eating.[17]

Rich inflection has also been associated with Null-Subject Languages (NSL) ever since Perlmutter (1971) divided languages in relation to surface structure constraints, one of them being the possibility of depicting null subjects with inflected verbs. Camacho (2016, p. 3), referring to rich inflection, states that "it is this property that allows a language to either identify *pro*, or to have pronominal agreement with referential properties".

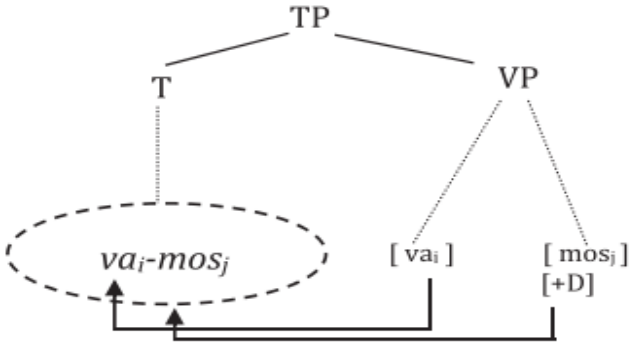
Moreover, the NSP has been reformulated and new ways of defining the parameter have been advanced. Alexiadou and Anagnostopoulou (1998, pp. 494, 522) state that Spanish can have null subjects because Spanish's strong agreement morphology has a categorial feature [+D]. Spanish agreement affixes behave like subject pronouns, have semantic features [+interpretable]

[17] Examples (16) and 1(7) were translated from Italian into Spanish from Carminatti (2002, p. 58)

and have independent lexical entries. In this regard, the verbal root and the [+D] affix are separate items in the numeration.

Nevertheless, the main point is how EPP checking operates in Spanish; EPP is parametrized and EPP is checked by merging the verb with the agreement affixes, which have nominal features (move/merge X^o), as is represented in the tree structure in (18).[18]

(18) Vamos.
Go-PRS.1PL
We go.

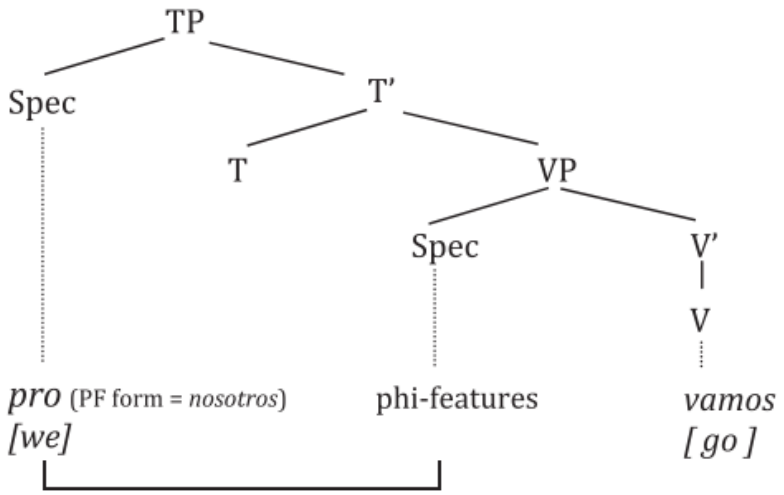


Taken from Liceras and Fernández (2016)

Taking a different approach, Holmberg (2005) and Sheehan (2006) suggest that languages like Spanish can have zero pronouns because this is a PF matter. Pronouns are not spelled-out due to economy principles, which prevent the system from producing duplicates. In relation to EPP checking, first, the null subject rises to Spec, TP. Even if the null subject pronoun has phi-features, it does not have [D] or deictic/referential properties. Second, the verb and the agreement morphology move to T, and third, the uninterpretable features in T and the interpretable features of V are valued (Liceras and Fernández, 2016, p. 7), as illustrated in (19).

[18] The extended projection principle (EPP) was proposed by Chomsky (1981) to state that all sentences need a subject, although in some cases the subject can be implicit, as is the case in Spanish. Liceras (2010, p.266) says that this principle has also been used "to refer to a feature which is bundled with other features to trigger movement. In other words, it has somehow taken over the role of the [+/-strong] value of features".

(19)



Taken from Liceras and Fernández (2016)

The NSP has been contested many times and the cluster of properties put to the test. This has led to the proposal that there are languages that do not comply with all the properties attributed to the [+Null Subject] option or to the [-Null Subject] option. Camacho (2013) and D'Alessandro (2015) provide a summary of the different types of languages, depending on how they realize or omit the subject. There are: (i) canonical NSLs such as Italian or Spanish; (ii) radical NSLs such as Chinese, Japanese, or Korean, which lack rich inflection but still allow null subjects thanks to discourse-specific properties; (iii) partial NSLs, in which the null subjects are restricted to some particular structures or the featural configuration of the subject pronouns, among them Finnish, Russian, and Marathi; and (iv) expletive NSLs where only the expletive pronouns can be null, as is the case in Dutch and Haitian Creole.

One relevant issue for the purpose of this dissertation is: where does Wayuunaiki lie in the spectrum of languages of the world in relation to the expression of the subject personal pronoun? In the next section, the realization of pronominal subjects in this language is explored in order to try to achieve a better understanding of their status.

5.3 WAYUUNAIKI

To the best of our knowledge, Wayuunaiki has not been classified in relation to the Null Subject parameter. This indigenous language displays two sets of conjugations: the analytical conjugation and the synthetic conjugation. What we need to determine is how this double conjugation relates to the properties of a canonical NSP language like Spanish.

5.3.1 REFERENTIAL NULL SUBJECTS

As mentioned before, one property of Spanish is its ability to have null subjects which are recovered from the morphological characteristics of its verbal system including the [+person] and [+number] features.

In the case of Wayuunaiki, the subject is realized differently, and its expression depends on the type of verb. In the **synthetic conjugation**, which is only used by active transitive verbs, the subject personal pronoun is expressed via a pronominal prefix (Table 5.1) incorporated to the verb as in (20).[19]

(20) Te'reechi Juan.
ta-'ra-ee-chi Juan
1S-see-FUT-M Juan
'I will see Juan'.

Olza & Jusayú (1978, p. 169)

[19] Active transitive verbs can use the synthetic or the analytical conjugations. The use of one or the other conjugation depends on definiteness and focus. Álvarez (2002) states that the use of the synthetic conjugation requires a definite direct object, whereas, in the analytical conjugation the direct object can be definite or not. Our investigation does not take into account this difference because it was beyond the scope of this dissertation to include these fine-grained lexical differences. It is our intention to provide a global analysis for the time being and encourage further research on this topic.

Table 5.1. Wayuunaiki's personal pronouns and pronominal prefixes

	Analytical conjugation **Personal Pronouns**	***Synthetic conjugation*** **Pronominal Prefixes**	**Spanish Personal pronouns**	**English Personal pronouns**
1S	taya	ta-	Yo	I
2S	pia	pü-	tú/usted	You
3SM	nia	nü-	Él	He
3SF	shia/jia	sü-/jü-	Ella	She
1P	waya	wa-	nosotros/as	We
2P	jia	sa-/jü-	Ustedes vosotros/as	You
3P	naya	na-	ellos/as	They

The prefixes displayed by the synthetic conjugation have the features [+number], [+person], and the feature [+gender] available only for the 3rd person singular. In the analysis available for this conjugation, the pronominal prefixes are not considered full subjects, which means that the position for the subject slot is empty or zero (Matera, 2001), so the order of the constituents in this conjugation is ØVO. Thus, with respect to this conjugation, Wayuunaiki can be considered to be a null subject language.

In fact, we could assume that the prenominal prefixes which bear the features [+number] and [+person] of the Wayuunaiki's synthetic conjugation are the equivalent of the post-nominal suffixes which occur in Spanish and which also bear the [+number] and [+person] features, as shown in Table 5.2.

Table 5.2. Spanish present tense verb endings

		-AR conjugation	**-ER conjugation**	**-IR conjugation**
1S	yo	-o	-o	-o
2S	tú/usted	-as	-as	-as
3SM	él	-a	-a	-a
3SF	ella	-a	-a	-a
1P	nosotros/as	- amos	- emos	- imos
2P	ustedes	-an	-en	-en
3P	ellos/as	-an	-en	-en

The **analytical conjugation** is displayed by stative verbs and active verbs (transitive and intransitive). This conjugation requires overt subject pronouns like the ones available in Spanish (Table 5.1). However, the dominant position of the subject for this conjugation is postverbal, namely the order is VSO (Álvarez, 2017). The Wayuunaiki verb morphology used with this conjugation presents the features [+number] and [+gender] only for the singular persons. In this conjugation Wayuunaiki does not mark person, as shown in (21). In this point Wayuunaiki differs from standard English, in which person is marked with a /s/ for the 3rd person singular in the present tense, and also from Spanish that marks person across the entire conjugation.

(21) Ekerotshi taya tepialu'um*üin.*
a-kerola-shi taya ta-pia-lu'u-m*üin*
0-enter-M I 1S-house-inside-towards
'I enter the house.'

Álvarez (2002, p. 3)

Regarding the omission of the subject in this conjugation, Álvarez (2004) indicates that the subject can be omitted in some contexts. We have noticed that the subject can be omitted in the analytical conjugation right after it has been presented in a previous sentence, as shown in the dialogue *The visit* below in turns (I) and (J) (Table 5.3). Ana asks about Josefa's daughter and Josefa answers with *O'unüsü (shia) apalaalajüin Uribiamüin* [she went to Uribia], omitting the pronoun *shia*. In the following sentence, Josefa reprises the pronoun when she adds: *Aikküsü* ***shia*** *kaa'ulo'uta* [she is selling goat leather there]. In that sentence, we believe that the subject could have be omitted because there is topic continuity, as it would be the case in Spanish. However, *shia* is repeated without any switch reference.

Table 5.3. The visit (Mansen & Mansen 1984: 35-36)

A	**Ana:**	Antüsü	**pia**,	Josefa.	
		anta-sü	pia	Josefa	
		arrive-FEM	you	Josefa	
		'You have arrived, Josefa'.			
B	**Josefa:**	Aa.	Antüsü	**taya**.	
		aa	anta-sü	taya	
		yes	arrive-FEM	I	
		'Yes. **I** have arrived'.			
C	**Ana:**	¿Kasachiki?			
		kasa-chiki			
		what-about			
		'What's new?'			
D	**Josefa:**	Nnojotsü	kasachiki.	Waraitüsü	ne'e.
		nnojulu-sü	kasa-chiki	waraita-sü	ne'e
		not.be-FEM	what-about	walk-FEM	just
		'Nothing new. **I** am just walking'.			
E	**Ana:**	¿Anayaasüje'e	**pia**?		
		ana-yaa-sü-je'e	pia		
		good-IRR-FEM-POND	you		
		'Are **you** fine?'			
F	**Josefa:**	Anasü	**taya**.	¿Jama	**pia**kat?
		ana-sü	taya	jama	pia-kat
		good-FEM	I	how	you-ART:FEM
		'I am fine. How about **you**?'			
G	**Ana:**	Anasü	**taya**.	¿Jalashi	pu'wayuuse?
		ana-sü	taya	jala-shi	pü-'wayuuse
		good-FEM	I	where-MAS	2s-spouse
		'I am fine. Where is your husband?'			

H	**Josefa:**	¿**Nia**kai?	Isa'aya.	Makatüshi	miichipa'a.
		nia-kai	isa'aya	makata-shi	miichi-pa'a
		he-ART:MAS	there	stay-MAS	house-area
		'Him? He is there. **He** stayed at home'.			
I	**Ana:**	¿Jama	puchon?	¿Jalasü	**shia**?
		jama	pü-chon	jala-sü	shia
		how	2s-daughter	where-FEM	she
		'How is your daughter? Where is **she**?'			
J	**Josefa:**	O'unusu	apalaalajüin	Uribiamüin.	Aiküsü
		o'una-sü	apalaalaja-in	Uribia-müin	aika-sü
		go-FEM	walk-GER	Uribia-to	sell-FEM
		'She went walking to Uribia. **She** is selling goatskin'.			
K	**Ana:**	¿O'unusü	wayuu	wainma?	
		o'una-sü	wayuu	wainma	
		go-FEM	people	many	
		'Did many people go?'			
L	**Josefa:**	Aa.	Jee	cha'aya,	aya'lajüshii
		aa	jee	cha'aya	aya'laja-shii
		yes	and	there	buy-PLU
		'Yes. And there **they** are buying food'.			
M	**Josefa:**	Otta **taya**kat, ajulijaasü **taya** suukula. Opoolojüsü **taya** kepein			
		Otta taya-kat ajulijaa-sü taya suukala opoolojo-sü taya kepein			
		But I-ART:FEM look.for-FEM I sugar brew-FEM I coffee			
		'As for **me**, **I** am looking for sugar. **I** am boiling/brewing coffee'.			
N	**Ana:**	Eesü	**ta**ma'ana	suukala	palitchon.
		ee-sü	ta-ma'ana	suukala	paliru-chon
		be-FEM	1s-possession	sugar	little-DIMIN
		'I have very little sugar'.			

shia	kaa'ulo'uta.		
shia	kaa'ula-'uta		
she	goat-skin		
naya	eküülü.		
naya	Eküülü		
they	Food		

O **Josefa:** Anasü tia. Keraasü.
ana-sü tia keraa-sü
good-FEM that finish-FEM
'It is fine. That is all'.

P **Ana:** ¡Nna! ¡Atak! Aja'lajaasü. **Shia**iria ne'e.
nna atak aja'lajaa-sü shia-iria ne'e
behold wow finish-FEM she-PAST just
'There you are! Wow! **It** is all'.

Q **Josefa:** O'unusü **taya**.
o'una-sü taya
go-FEM I
'I am leaving'.

R **Ana:** Anakaja. ¿Jouja **pia**?
ana-kaja jouja pia
good-EMPH when you
'Good. When **you** (are coming back)?'

S **Josefa:** Watta'a.
watta'a
tomorrow
'Tomorrow.'

Whether the pronoun can be omitted in sentences with additional intervening clauses remains to be uncovered. The Wayuunaiki's analytical conjugation lacks the feature [+person], and the feature [+gender] (available in Wayuuaniki only for the singular persons) is missing in Spanish. However, if null subjects can be recovered from the previous sentences in limited contexts in this conjugation, then Wayuunaiki is more like English (a non-null subject language) but somehow more flexible in the sense that Wayuunaiki can license null subjects via context. Some null pronominal subjects are allowed, granted they are recoverable from the previous produced information.

5.3.2 NON-REFERENTIAL NULL SUBJECTS (EXPLETIVES)

Overt expletives are not allowed in Standard Spanish and even though constructions with overt expletives have been reported for Dominican Spanish (Bullock & Toribio, 2009; Martínez-Saenz, 2011), this is not a common property of this language. Wayuunaiki does not have overt expletive pronouns either (Álvarez, personal communication, March 4, 2019). Verbs used to describe the weather and other predicates which license overt expletives in [-null subject] languages like English, do not license them in Spanish. In the case of Wayuunaiki, there are no overt expletives and these types of clauses are expressed with the stative conjugation and a zero pronoun. Weather statements in this indigenous language are expressed with a verb that indicates number (singular) and gender (feminine), but not person, as in the examples (22) and (23), as well as are phrases such as *it is important.., it is necessary*…, as in examples in (24) and (25).

(22) Aisü walatshi.
a-i-sü walatshi
Ø-hurt-S.F hot
'It is hot'.

(Bustos & Ipuana, 2015, p. 93)

(23) Aisü jemiai.
a-i-sü jemiai
Ø-hurt-S.F cold
'It is cold'.

(Bustos & Ipuana, 2015, p. 93)

(24) Che'ujaasü ma'in püyaawatüin sulu'u pütchikalü.
Che'ujaa-sü ma'in pü-yaawatü-in su-lu'u pütchi-kalü.
necessary-FEM very 2s-understand-SUB 3sF-inside word-ART:FEM
'It is very necessary/important you understand the content of the word'

(Álvarez, personal communication, March 4, 2019)

(25) Che'ujaasü. püküjain süchiki jamakuwa'ipalün
che'ujaa-sü pü-küja-in sü-chiki jama-kuwa'ipalün
necessary-FEM 2s-tell-SUB 3sF-about how-form
'It is necessary you explain it'.

(Álvarez, personal communication, March 4, 2019)

5.3.3 SUBJECT/VERB INVERSION IN DECLARATIVE SENTENCES

Spanish subject-verb order is typically SVO. However, it also allows subject-verb inversion and other possibilities, such as: VSO, OVS, and OSV (Lahousse & Lamiroy, 2012). In the case of Wayuunaiki, each conjugation has its exclusive subject-verb order, and inversion is not as free as in Spanish (Álvarez, 2017). In the analytical conjugation, the dominant order is VSO as in (26), which is the unmarked order in Spanish with unaccusative verbs (Lubbers Quesada, 2015, p. 39). Wayuunaiki can also have in this conjugation a topic + VO order, for example in (27), in which the subject is fronted when the clause is topicalized or left-dislocated.[20]

[20] In Spanish, the inversion of the subject, although apparently free, is conditioned by syntactic-semantic rules, the Spanish verb inflection, pragmatic, lexical, and phonetic principles (Liceras, 1994, pp. 219-220).

(26) O'unajachi taya sünain alapajaa nümaa tashi.
a-'una-ja-chi taya sünain alapajaa nü-maa t-ash-i.
0-go-FUT.M I towards wake 3S-with 1S-blood.M
'I am going to a wake with my dad'.

(Bustos & Ipuana, 2015, p. 7)

(27) Taya, o'unajachi sünain alapajaa nümaa tashi.
Taya a-'una-ja-chi sünain alapajaa nü-maa t-ash-i.
I 0-go-FUT.M towards wake 3S-with 1S-blood.M
'As for me, I am going to a wake with my dad'.

(Álvarez, personal communication, March 4, 2019)

In the synthetic conjugation, the subject is a prefix incorporated to the verb, which means that a pronominal subject is not phonetically realized, thus, the order would be represented as ØVO, as shown in (28). In this conjugation there is no possibility of inversion.

(28) Tousajeechi Kamiirü.
Ta-ousa-jee-chi Kamiirü.
1S-kiss-FUT-M Kamiirü.
'I will kiss Camilo'

(Bravo, 2005, p. 87)

5.3.4 The use of null and overt subjects in Spanish

Null and overt subjects in Spanish are determined by many factors. In variationist studies, several predictors of overt SPP expression have been found, for example: person/number of the verb, tense, mood, aspect (TAM), switch reference, priming effects, among others.[21] However, Torres-Cacoullos and Travis (2018, p. 74) claim that the previous predictors must be analysed in a broader perspective to include cognitive, interactional, and communicative elements. From a discourse-pragmatic approach, SPP choice is constrained by factors such as "cohesion, coherence, saliency, presupposition, pragmatic weight, and cognitive status" (Lubbers Quesada, 2015, p. 262). The distribution of null and overt pronominal subjects in Spanish also depends on the PAH

[21] A more detailed analysis of these variables is provided on pp. 105-110.

strategies to resolve anaphoras and on the OPC properties to understand and interpret the co-reference of subjects.

Additionally, in Spanish the overt pronouns are said to disambiguate a referent or convey emphasis. This emphasis contrasts the information and focalizes the element, as in (29), in which the pronoun *ellos* [they] is required to emphasize or contrast that a person different than "Ana ate the cake".

(29) Ana hizo el biscocho y ellos se lo comieron
Ana baked$_{[3p\text{-}s]}$ the cake and they ate$_{[3p\text{-}p]}$ it
'Ana baked the cake and they ate it'.

Unlike overt subjects, null subjects suggest no changes in the discourse topic. When new information is introduced in the discourse, it is appropriate to use an overt subject (Pladevall, 2013), albeit it does not happen systematically because native Spanish speakers produce null subjects in [+ topic shift] contexts as well (Liceras, Alba de la Fuente & Martínez-Sanz, 2010). Nevertheless, Torres-Cacoullos and Travis (2018, p. 76) suggest that the contrastive use of overt/null pronouns is also linked to the presence of specific constructions or contrastive elements such as: *mismo, pero, mas*, etc. [himself/themself, but, nonetheless, etc.].

Concerning Wayuunaiki, as far as we know, there are no studies on the constraints affecting the distribution of null and overt subject pronouns in this language. In fact, the information about the use of the pronominal subject in Wayuunaiki is very limited, and we are relying on secondary sources.[22] We have found, however, that omission of the subject is possible in contexts in which there is no change of subject reference, but this is probably limited to one preceding clause. On

[22] We should acknowledge that the shortage of studies on the distribution of null and overt subject pronouns in Wayuunaiki and our limited knowledge of Wayuunaiki is a disadvantage when it comes to producing a more complete comparison of the factors that constraint the distribution of null and overt subject pronouns and the overt SPP expression in this language as it compares to Spanish. This first examination aims to open the door to more studies on this specific topic in Wayuunaiki.

the contrary, overt pronouns seem to be expressed in all contexts, including in topic continuity, see (M) in the dialogue *The visit*, in which the pronoun *taya* [I] is kept even if it is the same subject as the one pronounced in the previous sentence.

Three properties of Spanish have been analysed and compared to see the way they function in Wayuunaiki. Wayuunaiki has a double conjugation system and this could have implications regarding the distribution of null and overt pronominal subjects in Guajiro Spanish. With respect to the analytical conjugation, Wayuunaiki is a [-Null Subject] language like English is, but seems to have more reliance on discourse mainly when the subject has been mentioned in the previous sentence. However, in the synthetic conjugation Wayuunaiki is a [+Null Subject] language. We have also seen that Spanish and Wayuunaiki share two properties; they both have null expletives and they both can have VS order, albeit, in the case of Wayuunaiki, only in the analytical conjugation.

The dual conjugation system of Wayuunaiki, which is additionally linked to whether the verb is stative, active transitive, or active intransitive, creates an important difference between the two languages since Spanish has just one conjugation system and the expression of the subject is not bound to any verb class but to constraints of varied nature, as summarized in Table 5.4.

Table 5.4. Summary of subject realization according to the properties of the Wayuunaiki and Spanish conjugations

Wayuunaiki		Spanish
Analytical conjugation	**Synthetic conjugation**	**All conjugations**
[- Null subject language]	[+ Null subject language]	[+ Null subject language]
Omission of the subject restricted to contextual conditions	Subject always null	The subject is overt/null depending on a wide number of factors, including discourse-pragmatic, syntax-pragmatics, cognitive, communicative, and interactional constraints
• Active verbs (transitive and intransitive) • Stative verbs	• Active transitive verbs	• Any verb

5.4. The distribution of null and overt pronominal subjects in Bilingual and Monolingual Spanish in Relation to Stative and Active Transitive/Intransitive Verbs

The comparative description depicted in Table 5.4 will be our point of departure to investigate whether there is a relationship between the distribution of NOS in Guajiro Spanish, and the stative, active transitive, and active intransitive nature of Wayuunaiki verbs in the case of the Spanish bilingual group. We will also investigate how the bilingual group compares to the monolingual group and whether there are differences depending on the level of proficiency in the Spanish of the bilingual participants.

To carry out this analysis, only null and overt subject pronouns in declarative sentences were selected for a total of 2448 tokens for the bilingual group and 481 for the monolingual group (a complete description of these two groups is presented in Chapter 4). Sentences with modal verbs, gerunds, and periphrasis such as *ir a trabajar* [to go to work] were not included. The selected tokens were separated into sentences containing overt or null personal pronouns, and according to verb class, the tokens were then divided into clauses with stative, active transitive, or active intransitive verbs, as in examples (30) to (35).

(30) Ellos están en la ranchería. ABR
They are$_{[3p\text{-}p]}$ in the rancheria.
'They are in the rancheria'.
[Overt subject pronoun and stative verb]

(31) Ellos reparten hamacas. ROS
They deliver$_{[3p\text{-}p]}$ hammocks.
'They deliver hammocks.'
[Overt subject pronoun and active transitive verb]

(32) Yo trabajo por aquí. NAR
I work$_{[1p\text{-}s]}$ for here.
'I work here'.
[Overt subject pronoun and active intransitive verb]

(33) Ø viven en la ranchería. GEI
Ø live$_{[3p\text{-}p]}$ in the rancheria.
'They live in the rancheria.'
[Null subject pronoun and stative verb]

(34) Ø bordo mis mantas.[23] CAT
Ø embroider$_{[1p\text{-}s]}$ my dresses
'I embroider my dresses'.
[Null pronoun and active transitive verb]

(35) Ø participamos en cualquier evento. MIG
We participate$_{[1p\text{-}p]}$ in any event.
'We participate in any event.'
Null pronoun and active intransitive verb

Verbs were coded as active intransitive or active transitive depending on how they were used by the participants. For example, *yo hablo mucho* [I speak a lot] was coded as active intransitive, but *yo hablo el Wayuunaiki* [I speak the Wayuunaiki language] was coded as active transitive. Instances of stative verbs were: *tener, ser, vivir, morir, nacer, sufrir, estar, saber, cumplir,* etc. [to have, to be, to live, to die, to be born, to suffer, to be, to know, to achieve or accomplish, etc.]. Examples of active transitive verbs were: *decir, poner, buscar, comprar, vender,*

[23] A manta is a Wayuu type of dress used by the women of this Aboriginal group.

preguntar, etc. [to say, to put, to search, to buy, to sell, to ask, etc.]. Cases of active intransitive verbs were*: ir, salir, viajar, volver, trabajar,* etc. [to go, to leave, to travel, to return, to work, etc.].

5.5 Research Questions

The research questions that we would like to answer are the following:

1) How does the verb type affect the rates and the distribution of the null and overt subject pronouns of the bilingual and monolingual Spanish groups, depending on whether the verb is stative, active transitive, or active intransitive?
2) For the bilingual group, are the rates and distribution of the null and overt subject pronouns related to Spanish proficiency?

5.6 Hypotheses

1) If the Wayuunaiki conjugation system plays a role in the bilingual speakers' rate of overt pronouns realization and the distribution of null and overt subject pronouns in Spanish, there should be a higher rate of overt subject pronouns with stative and active intransitive verbs, since these verbs obligatorily select the analytical conjugation too and therefore require an overt pronoun. On the other hand, the active transitive verbs should display a lower rate of overt subject pronouns, as they can select the synthetic or the analytical conjugation, which implies that whenever an active transitive verb is used with the synthetic conjugation the subject must be null. In order to assess this hypothesis, we will compare the rates of null and overt pronominal subjects produced for each type of verb by the intermediate and advanced Spanish groups. These rates will be compared to those produced by the monolingual group as well.
2) If Wayuunaiki, the dominant language of the bilingual speakers, plays a role in the distribution of the null and overt subject pronouns in Spanish, the Spanish intermediate

speakers should have higher rates of overt subject pronouns with stative -especially- and active intransitive verbs vis-à-vis the Spanish advanced speakers since the intermediate group may not have acquired the discourse-pragmatic and the other constraints that are linked to the presence or absence of overt subject pronouns in Spanish. Additionally, the intermediate and advanced group could display a different choice of the type of verbs that promote overt pronominal subjects in Guajiro Spanish. This hypothesis will be assessed by comparing the rate of overt pronominal subjects in the Spanish intermediate and the Spanish advanced groups for each type of verb.

5.7 RESULTS

The research questions of this study were related to (i) the verb types and its association with the rate and the distribution of null and overt pronominal subjects, and (ii) the effect of Spanish proficiency in the bilingual group and how the results relate to the monolingual Spanish cohort.

The analysis of the data shows that the intermediate Spanish proficiency group produced the following order with respect to the rate of null and overt subject pronouns: (1) stative verbs, (2) active intransitive verbs, and (3) active transitive verbs (See Table 5.5). This implies that this group produces more overt pronominal subjects with stative verbs than with the other two verb classes. The rate of overt subject pronouns with stative verbs was 41% vs. 36% for the active intransitive verbs and 32% for the active transitive verbs. Regarding the overall rate of overt subject realization, this group had a rate of usage of 36%. This association between verb type and the overt realization of the pronominal subjects is statistically significant, as confirmed by a Chi-square test: X^2 (3, N = 901) = 6.40, $p < .05$.

Table 5.5. Distribution of null and overt subject pronouns by verb type for the intermediate group

Verb type	Total N of tokens	Overt pronouns	Null pronouns
Stative verbs	N= 345	N= 140 (41%)	N= 205 (59%)
Active intransitive verbs	N= 155	N= 56 (36%)	N= 99 (64%)
Active transitive verbs	N= 401	N= 127 (32%)	N= 274 (68%)
Total N of tokens	N= 901	N= 323 (36%)	N= 578 (64%)

Similar results were found in the case of the advanced group. For these speakers, the ordrer was identical: (1) stative verbs, (2) active intransitive verbs, and (3) active transitive verbs. Similarly, more overt subject pronouns occurred with stative verbs (39%), than with active intransitive verbs (33%) and active transitive verbs (32%). The last two types of verbs had almost identical rates of usage and hence the difference between these two groups is less clear-cut. This group had an overall rate of overt subject pronouns usage of 35%, (Table 5.6). The Chi-square test performed on the relationship between verb type and the realization of pronominal subjects showed that it was also significant X^2 (3, N = 1547) = 8.76, $p < .05$.

The intermediate and advanced groups are similar with respect to the expression of the personal pronouns by verb class. In both cases, overt personal pronouns are preferred with stative verbs, 41% for the intermediate group and 39% for the advanced group, followed by active intransitive and active transitive verbs. Also, the overall rate of overt subject pronouns use is nearly identical for both the intermediate and advanced groups (36% vs. 35%).

These results confirm the first hypothesis which stated that more overt pronouns should be expected with stative, especially; and with active intransitive verbs if the Wayuunaiki conjugation system was playing a role in the production of null and overt subject pronouns by the bilingual speakers, as these types of verbs are conjugated with the analytical conjugation, which requires

overt pronouns to be expressed. In the case of the advanced group, nonetheless, the difference between the active intransitive and the active transitive was only 1%.

Table 5.6 Distribution of null and overt subject pronouns by verb type for the advanced group

Verb type	Total N of tokens	Overt pronouns	Null pronouns
Stative verbs	N= 702	N= 275 (39%)	N= 427 (61%)
Active intransitive verbs	N= 242	N= 81 (33%)	N= 161 (67%)
Active transitive verbs	N= 603	N=190 (32%)	N= 413 (68%)
Total N of tokens	N= 1547	N= 546 (35%)	N= 1001 (65%)

In relation to the effect of Spanish proficiency, any effect of this variable must be ruled out as the intermediate and the advanced groups behaved almost identically regarding the overall distribution of NOS and the expression of the subject by verb class.

In the case of the Spanish monolingual group, 69% of the tokens were null pronouns and 31% of the tokens were realized overtly, as shown in Table 5.7. This result is in line with the results of the variationist analysis of the overt SPP expression (cf. chapter 6), in which the monolingual group was less productive regarding overt SPPs than the bilingual group.

Table 5.7. Distribution of null and overt subject pronouns by verb type for the monolingual speakers

Verb type	Total N of tokens	Overt pronouns	Null pronouns
Active intransitive verbs	N= 95	N= 39 (41%)	N= 56 (59%)
Stative verbs	N= 209	N= 67 (32%)	N= 142 (68%)
Active transitive verbs	N= 177	N= 43 (24%)	N= 134 (76%)
Total N of tokens	N= 481	N= 149 (31%)	N= 332 (69%)

When the distribution of the null and overt pronominal subjects is analysed by verb class, the rate of overt pronouns with active intransitive verbs was the highest (41%), followed by stative verbs (32%) and active transitive verbs (24 %). The Chi-square test performed shows that the association between verb type and the variable realization of the pronominal subjects was significant X^2 (3, N = 481) = 8.32, $p < .05$.

These results show that the monolingual and the bilingual groups pattern differently and only coincide with respect to the class of verb (the active transitive verbs) that disfavours the use of overt personal pronouns the most. Regarding the order by verb class, the monolingual group produced the following: (1) active intransitive verbs, (2) stative verbs, and (3) active transitive verbs.

5.8 DISCUSSION

Our first research question was geared towards uncovering the rate and the overall distribution of overt and null subject pronouns by verb type for the bilingual and the monolingual groups.

It was hypothesized that there would be a higher rate of overt pronominal subjects with stative and active intransitive verbs in the bilingual group. This hypothesis has been confirmed since the rate of overt pronominal subjects with stative verbs was 41% for the intermediate group

and 39% for the advanced group. For the active intransitive verbs, the rate was 36% for the intermediate group and 33% for the advanced group. These results were predicted since more overt pronouns were expected for the stative verbs and the active intransitive verbs as these verbs use the analytical conjugation, for which, as we have seen above, Wayuunaiki obligatorily requires overt subject pronouns.

It was also found that the monolingual group presented different results in the preference for the expression of pronominal subjects vis-à-vis verb type. There was a 41% rate of overt pronouns with the active intransitive type, followed by the stative verbs with 32% and by active transitive verbs with the lowest rate (24%). In Table 5.8, we can see the that the bilingual and monolingual groups pattern differently as the bilinguals produce more overt pronominal subjects with stative verbs. This can be accounted for by the fact the only available conjugation for the stative verbs in Wayuunaiki is the analytical one which requires an overt subject pronoun. The monolingual group, however, produced more overt subject pronouns with active intransitive verbs. This could suggest a semantic difference, in which active verbs in Spanish are preferred to produce overt pronominal subjects. However, more investigation needs to be done in this specific regard.

Table 5.8. Summary of the distribution of null and overt subject pronouns by verb type for the monolingual and bilingual groups

Verb type	Native Spanish	Verb type	Advanced group	Intermediate group
Active intransitive verbs	N= 39 (41%)	**Stative verbs**	N= 275 (39%)	N= 140 (41%)
Stative verbs	N= 67 (32%)	**Active intransitive verbs**	N= 81 (33%)	N= 56 (36%)
Active transitive verbs	N= 43 (24%)	**Active transitive verbs**	N=190 (32%)	N= 127 (32%)
Total N of tokens	N= 149 (31%)	**Total N of tokens**	N= 546 (35%)	N= 323 (36%)

Regarding the second research question, when the bilingual groups are compared by Spanish proficiency, the result is that the advanced group and the intermediate group behave similarly regarding the overt realization of the personal pronoun, distribution of the subjects, and the effects of the type of verbs, which shows that Spanish proficiency does not play a role in the rate and distribution of the null and overt personal pronouns realized by these Wayuunaiki-Spanish bilinguals.

5.9 SUMMARY

The aim of this chapter was to investigate whether the Wayuunaiki's dual conjugation system and verb class had an impact on the Guajiro Spanish variety, specifically in relation to the distribution of NOS. The results of this investigation suggest that the Wayuunaiki verb class distinction might be playing some role in the bilingual group. This claim is supported by the fact that overt pronominal subject expression is preferred with stative verbs by the intermediate and advanced bilingual groups, the verb classes that select overt pronouns in Guajiro Spanish was similar for both groups of bilingual speakers, whereas in the monolingual group overt pronominal subjects realization was preferred with active intransitive verbs and the order of the type of verbs

was slightly different from the one produced by the bilingual cohort. The orders of type of verbs established here for the bilingual and monolingual group were based on rates of pronominal subjects realization, and we acknowledge that rates can vary due to all kinds of non-linguistic reasons, however, if we can show that there is a relationship between the type of verb and the rate of null and overt pronouns which the bilinguals (but not the monolinguals) use, there is no reason why we cannot attribute that relationship to the fact that in Wayuunaiki the conjugation linked to stative and active intransitive verbs requires the obligatory realization of subject pronouns.

There is also an important question that we will leave for further research and it is whether the rate of pronominal subjects produced by the monolingual group reflects the existence of a possible universal tendency which would characterize [null subject] languages. That is, whether monolingual Spanish and other null subject languages prefer overt personal pronouns expresión wth active intranstive verbs than with any other type of verbs.

Finally, we would like to conclude by stating that the results of this chapter represent a contribution to the field of descriptive linguistics in general and language contact in particular because, as far as we know, the effect of transitivity and stativity on the distribution of null and overt pronominal subjects has not been investigated in the L2 Spanish of speakers whose L1 is an indigenous language.

6. OVERT SUBJECT PERSONAL PRONOUN EXPRESSION (SPP) IN GUAJIRO SPANISH

6.1 INTRODUCTION

In this chapter, we analyse the variable realization of pronominal subjects in Guajiro Spanish and in a Spanish native benchmark group from the variationist sociolinguistics perspective. This chapter is organized as follows: First, we present an overview of the overt subject personal pronoun expression in Spanish. Second, we review studies on overt SPP expression in monolingual and bilingual Spanish. Third, we present the research questions and the hypothesis. Fourth, the results of the study are examined. Finally, we discuss the results of the study and offer a summary of the chapter.

6.2 OVERT SUBJECT PERSONAL PRONOUN EXPRESSION IN SPANISH

Overt SPP expression in Spanish can be overt or null, as we mentioned in the properties of the Null Subject Parameter discussed in the previous chapter. This means the subject can be present or absent in a finite clause as illustrated in examples (36a) and (36b).

(36) a. Yo como cereal con leche todos los días.
b. Ø como cereal con leche todos los días.
Ø eat$_{[1p.s]}$ cereal with milk everyday.
'I eat cereal with milk everyday'.

Although in some contexts one of the options is categorical, variable overt SPP expression is not free, but is constrained by certain grammatical, semantic, and discourse-related factors such as: *grammatical person/number*, *TAM morphology*, *priming effects*, *switch reference,* and *lexical semantics of the verb*.[24]

This alternation in the expression of the overt SPP in Spanish has been widely studied in relation to internal and external language factors, cross-dialectally and in language contact situations (e.g., Silva-Corvalán, 1982; Bentivoglio, 1987; Cameron, 1993, 1995; Lowther, 2004;

[24] We explain these contexts in more detail in the section devoted to the envelope of variation on pages 87-89.

Flores-Ferrán, 2004, 2005, 2007; Hurtado, 2005; Travis, 2007; Otheguy, Zentella & Livert, 2007; Orozco & Guy, 2008; Torres-Cacoullos & Travis, 2010; Carvalho & Child, 2011; Sánchez Arroba, 2011; De Prada, 2009; Martínez-Sanz, 2011; Otheguy & Zentella, 2012; Holmquist, 2012; Abreu, 2012; Shin, 2012, and Michnowicz, 2015, *inter alia*).

Although overt SPP expression in Spanish has been extensively examined in monolingual Spanish, less attention has been paid to bilingual contexts. In the latter case, these studies have mainly concentrated on English-Spanish scenarios, (Flores-Ferrán, 2004; Otheguy, Zentella & Livert, 2007, Otheguy & Zentella, 2012, among others). However, most recently, there have been studies on Spanish in contact with Catalan (De Prada, 2009, 2015), Spanish in contact with Portuguese (Carvalho & Child, 2011; Carvalho & Besset, 2015) and Spanish in contact with Yucatec Maya in Yucatán, Mexico (Michnowicz, 2015).

The underlying expectation in the aforementioned studies is that Spanish in contact with a non-pro-drop or a partial-null subject language, as in the case of Brazilian Portuguese, should exhibit some contact-induced influence from the other language. This influence should be patent by means of a higher rate of overt SPP expression in the Spanish contact variety vis-à-vis the overt SPP rate of a monolingual Spanish benchmark and, more importantly, by the language varieties under analysis having a different patterning in the direction of the effects of the independent variables promoting overt SPP realization. (Torres-Cacoullos & Travis, 2010; Abreu, 2011).[25],[26]

[25] Liceras, Fernández & Alba de la Fuente (2012) found that in bilingual child language acquisition, bilingual English-Spanish children did not display a higher production rate of overt Spanish subjects than the monolingual counterpart. They argued that even if subjects are obligatory in English, this is a core syntactic operation of that language which is not supposed to contribute to interlinguistic influence on the child's other language.

[26] Since some overt SPP expression is permitted in Spanish, some questions arise: what is the standard baseline and what does constitute an increase on overt SPP expression? (Poplack & Levey, 2010, p. 395). This is solved by having an appropriate comparison variety. In our case, by comparing the Guajiro Spanish group to the monolingual group of La Guajira.

By presenting and analysing data from Guajiro Spanish and comparing it to a local vernacular Spanish benchmark, this study seeks to shed light on how overt SPP expression is realized in this variety of Spanish. At the same time, this study looks to expand the research and coverage of this topic in other language contact situations and language pairs, a suggestion expressed by Carvalho, Orozco, and Shin (2015) who claim that research on Spanish overt SPPs expression in contact language situations has been primarily conducted on the English-Spanish pair.

On another front, our interest in studying overt SPP expression in Guajiro Spanish is justified by the fact that Spanish is a pro-drop language and Wayuunaiki has a dual conjugation system. In Wayuunaiki, the pronominal subject is null in the synthetic conjugation: the slot for the subject is zero and this can only be retrieved from the pronominal prefix incorporated to the verb (Matera, 2001), as in (37). In the analytical conjugation, the subject is post-verbally realized and is compulsory, as in (38). It can be null; however, this is limited to situations where the subject can be retrieved from the context (Álvarez, 2004).

(37) Tek**aa**pa.
ta-ka-a-pa
1S-eat-PROG-COMPLETED
'When I am eating'.

Álvarez (2007, p. 138)

(38) Aya'lajüshi taya piama shikii paa'a(irua).
a-ya'laja-shi taya piama shikii paa'a-irua
0-buy-M I two ten cow-PLU
'I bought twenty heads of cattle'.

Álvarez (2017, p. 256)

Our study is also pertinent to see if there is any contact-induced change in Guajiro Spanish. The expectation of this study is two-fold: to uncover the rate of overt SPP usage and the different constraints affecting this variable in Guajiro Spanish and the Spanish monolingual baseline group. A higher rate of overt SPPs is expected for the bilingual group as compared to the monolingual

benchmark, as well as a different direction of the effects of the constraints. This expectation is motivated by the Wayuunaiki conjugation systems and the discourse-pragmatic nature of the overt SPPs in Spanish, which is regulated at the syntax/pragmatics interface, an area that has been argued to be vulnerable to cross-linguistic influence and variability (Sorace, 2005; Filiaci, 2010, p. 171). Furthermore, Torres-Cacoullos & Travis (2018, p. 74) propose that the expression of overt SPPs in Spanish is not merely a syntax/pragmatics phenomenon, since other factors such as socio-cognitive and communicative knowledge are also at play when the overt pronouns are used to create contrast or emphasis, resolve ambiguity or determine the effect of the verb tenses. In this sense, the interpretation of the null and overt pronouns in pro-drop languages like Spanish or Italian is a complex operation (Sorace et al., 2009).

6.3 Previous Studies on Overt SPP Expression in Monolingual Spanish

Research carried out to date has found mixed results regarding the direction of the effects of the variables impacting overt SPP expression in Spanish. These variables are normally classified as, grammatical, discursive, semantic, and social.[27] As to the grammatical variables, *TAM morphology, lexical semantics of the verbs, grammatical person/number,* and *clause type* are the most frequent significant factors reported in the literature for overt SPP expression in Spanish, as summarized by Carvalho, Orozco and Shin (2015).

[27] In this study, we analyse only grammatical, discursive, and semantic variables. Although we coded for social variables such as gender, sex, literacy, and onset of learning of Spanish, they were left out in the final analysis of this chapter. There were a lot of interactions among the groups, for example, the group of *young speakers* were also the group who contained the majority of the literate participants, *the early learners of Spanish* group were also the group who had the majority of *advanced speakers*. Regarding the monolingual group, there were only 5 participants, and for the literacy group there was one illiterate participant, and there were no illiterate females or young speakers. Also, a factor group such as literacy is not supposed to influence overt SPP expression, since the determinants of overt SPP realization are community norms and not an idealized literate benchmark.

Table 6.1. Studies reporting the grammatical variables affecting overt SPP expression in Spanish

TAM	Silva-Corvalán, 1982
	Bentivoglio, 1987
	Cameron, 1993
	Lowther, 2004
	Hurtado, 2005
Grammatical Person/Number	Cameron, 1993
	Flores-Ferrán, 2004
	Hurtado, 2005
	Otheguy et al., 2007
	Orozco & Guy, 2008
	Carvalho & Child, 2011
	Holmquist, 2012
Lexical semantics of the verbs	Bentivoglio, 1987
	Hurtado, 2005
	Travis, 2007
	Torres-Cacoullos & Travis, 2010, 2011
	Otheguy & Zentella, 2007, 2012
	Orozco & Guy, 2008
Clause type	Abreu, 2009
	Orozco & Guy, 2008

In relation to discursive variables, switch reference is one on the most robust variables contributing to overt SPP expression in Spanish. According to this predictor, overt SPPs are more frequently expressed in switch reference contexts or clauses where there is a change in the referential subjects of two consecutive clauses than in co-referential environments, where the same subject is kept in two or more continuous clauses, as illustrated by the example in (30), -repeated here as (39) for convenience- in which there is a change from *yo* [I] to *tú* [you].

(39) **ELI**: yo piensa [sic], es claro, **yo** pienso porque **tú** quiere(s) tener una vida así libre.	**ELI**: I think, it is clear, I think about the reason you want to have a free life.

Table 6.2. Studies reporting discourse-related variables affecting overt SPP expression in Spanish

Switch reference	Bentivoglio, 1987
	Cameron, 1993,1995
	Travis, 2005
	Lowther, 2004
	Flores-Ferrán, 2004
	Otheguy et al., 2007
	Orozco & Guy, 2008
	Torres-Cacoullos & Travis, 2010
	Sánchez Arroba, 2011
	Carvalho & Child, 2011
	Holmquist, 2012

Another less common variable found to have a significant effect on overt SPP expression is that of *priming effects* (see Travis, 2007). For priming effects, the expectation is that a pronoun has more chances of being overtly realized when the previous clause is also expressed with an overt SPP.

Torres-Cacoullos and Travis (2018, p. 74) revisited some of the previous factors: subject continuity, ambiguity resolution, priming effects, tense and temporal sequence, etc. and they argue that overt SPP expression must be investigated considering cognitive, communicative and interactional constraints that also affect this issue. Their proposal is that the variables interact with others. For example, verb morphology interacts with temporal sequencing and the effects of semantic classes of the verb must be considered in relation to particular constructions such as *yo creo* [I think] or quotative *decir* [to say].[28]

6.4 Overt SPP Expression in Bilingual Settings

Studies on overt SPP expression in Spanish contact situations address the influence of the other language on Spanish by means of an increase in the rate of overt SPP expression and

[28] In this chapter, however, we are not taking into account their proposal and we are analyzing these variables in the way they have been traditionally investigated.

differences in the ranking of constraints. A contact hypothesis based on rates of overt SSP use has been dismissed by Flores-Ferrán (2004, p. 69) for Puerto Ricans in New York. In that study, although there was an increase in overt SPP expression, Puerto Ricans living in New York City patterned similarly to their counterparts in San Juan regarding person/number and switch reference factors. For the Portuguese-Spanish language contact situation in the town of Rivera (on the Uruguay-Brazil border), Carvalho & Child (2011) also reported that a contact hypothesis was not supported as bilinguals and monolinguals expressed overt SPP at similar rates and the direction of the effects of the independent predictors hypoththesized to constraint overt SPP was comparable for the Spanish of the bilingual and the monolingual speakers.[29] This was also the case for monolingual and bilingual Spanish in Mexico. Monolinguals in Mexico City expressed the overt SPP with a rate of 22% and bilingual Yucatec Maya-Spanish speakers of Yucatán, Mexico with 20%. Additionally, the bilingual speakers patterned similarly to other studies in relation to internal factors such as *person/number, switch reference, TAM,* and *reflexivity* (Michnowicz, 2015).

On the other hand, Otheguy et al. (2007) found evidence of English influence on the Spanish of New York. In their study, the children of immigrants presented a significant change in the rate of overt SPP realization, particularly the bilingual English-Spanish speakers born in New York, which the researchers interpreted as the effect of the English-Spanish contact situation. Silva-Corvalán (1994) mentions that Spanish speakers in contact situations with English exhibit a different preference in the expression of overt SPPs. While Spanish speakers in contact with English expressed more overt SPPs with 3rd person singular, speakers that were not exposed to English showed a higher rate of overt SPP expression in 1st person singular.

[29] These studies assume that Brazilian Portuguese is in the process of becoming a non-null subject language.

In summary, there have been conflicting results about whether Spanish in contact is affected by other languages. In relation to the overt SPP expression rate, conflicting results are expected, since rates of overt SPP expression are variable across Spanish varieties (Mayol, 2012, p. 421) and these rates can fluctuate for any kind of non-linguistic factors. For this reason, the patterning of factors is the important reference point for assessing contact-induced change. Carvalho, Orozco and Shin (2015) affirm that bilingual and monolingual varieties are, for the most part, affected by the same grammatical conditioning. Bilingual varieties, however, may present slight changes in their grammar due to contact-induced change and these changes are detectible because of differences in the underlying grammar of the bilingual variety as compared to a monolingual benchmark (Thomason, 2001). In this respect, the present study also looks to contribute to this discussion by addressing overt SPP expression in Guajiro Spanish and the Spanish vernacular benchmark to which the Wayuunaiki-Spanish speakers are habitually exposed.

6.5 Research Questions

To carry out the present study, we take into consideration the results obtained from previous research and our particular context. Thomason (2001, p. 76-77) suggests that among the linguistic predictors of contact-induced change are the typological distance between languages and markedness. Spanish and Wayuunaiki behave differently with regard to the issue under investigation in this chapter as the former is an Indo-European language and the latter is an Aboriginal language. Spanish is a pro-drop language and Wayuunaiki has a dual conjugation system, which means that in the synthetic construction it behaves like a null-subject language and in the analytical construction like a non-pro-drop language.

In addition, the Wayuunaiki-Spanish contact situation has a history of more than 500 years. This protracted contact, along with extensive bilingualism and high level of knowledge of the

languages involved, are widely believed to favour contact-induce change (Poplack & Levey, 2010, p. 411). Although stable, the Wayuunaiki-Spanish relation is asymmetrical. Wayuunaiki native speakers are more likely to be bilingual than native Spanish speakers are. In this investigation, the following research questions are addressed:

1) What is the rate of overt SPP expression for the Wayuunaiki-Spanish bilingual speakers and how does this compare to the rate of overt SPP expression of the Spanish native benchmark?
2) What are the significant linguistic variables contributing to overt SPP expression in the bilingual and the monolingual groups?
3) How does the constraint hierarchy of the bilingual speakers and the monolingual Spanish benchmark group pattern?

6.6 HYPOTHESES

1) Spanish overt SPP expression is high in Caribbean dialects and low in Mainland South American dialects (Cameron**,** 1994; Otheguy et al., 2007)**.** A rate of 45% was reported for Puerto Rico (Cameron, 1993) and (Flores-Ferrán, 2004) and 51% for Dominican Spanish (DS) (Martínez-Sanz, 2011), whereas in mainland dialects the frequencies are lower: Colombia 24%, Mexico 19%, and Ecuador 33%, (Otheguy et al., 2007). In this regard, we expect the bilingual and monolingual groups to have rates of overt SPP expression close to that of Caribbean dialects, since Guajiro Spanish is influenced by this variety. However, for the Guajiro Spanish group, the rate should be higher, particularly if Wayuunaiki's dual conjugation system has an influence on bilingual Spanish.[30]

[30] Nonetheless, it must be considered that DS Spanish is losing its status of a typical pro-drop language and in this way may be different from other Spanish varieties (Toribio, 2000).

2) For Colombian Spanish, Hurtado (2005) found *switch reference*, *textual and morphological ambiguity* and *verb class* and Orozco & Guy (2008) found *person/number*, *switch reference*, *TAM*, *verb class* and *type of clause* to be significant independent variables. If bilingual and monolingual varieties of Spanish are frequently conditioned by the same grammatical variables (Carvalho, Orozco & Shin, 2015), we should find that overt SPP expression is affected by the same factor groups in the Wayuunaiki-Spanish and the monolingual Spanish participants.

3) By analysing the constraints ranking and, more specifically the direction of effects, we want to uncover how the bilingual and monolingual grammars pattern. We don't expect to find substantial differences in the bilingual and monolingual grammar regarding this issue. However, we may find some permutations in the direction of the effects of the constraint rankings. Thomason (2001) says that bilingual varieties may exhibit minor changes in their grammar due to contact-induced processes. Similar patterns will indicate that the speakers of both varieties share the same grammar, whereas non-trivial differences in the ranking of constraints will support the argument that the Wayuunaiki-Spanish bilinguals and the monolingual Spanish speakers have a different underlying system for the variable expression of pronominal subjects.

6.7 METHODOLOGY

6.7.1 THE ENVELOPE OF VARIATION

Following previous studies (Silva-Corvalán, 1982; Cameron, 1993, 1995; Otheguy et al., *2012; inter alia*), the dependent variable under study in this chapter concerns the presence or absence of a subject personal pronoun and the reference is [+human] (*yo, tú, él/ella, usted, nosotros/as, ustedes, ellos/ellas*) [I, you, he she, we, you and they] in finite clauses. This means lexical or syntactic contexts where either a null or an overt subject pronoun could have potentially occurred. Non-human and inanimate subjects were not included in our study because there were only a few of them and there are different views on what to include. In Otheguy and Zentella (2012), inanimate subjects were included only if they were overt whereas Carvalho and Besset (2015) make a distinction between tangible and intangible inanimate (e.g., *la ciudad* [the city], *el país* [the country]) subjects and only included those that were tangible.

Contexts in which there is limited or no variability in the expression of overt SPPs were excluded from the analysis. These contexts were:

A. *Chronological or time-period sentences and finite sentences about the weather. This type of sentence does not have a denotational subject.*

(40) **Interviewer:** ¿y a los cuántos años vino, vino para Colombia? **MAR**: uff **Ø** hace años, cuando esta(ba) joven todavía.	**Interviewer:** At what age did you come, did you come to Colombia? **MAR:** uff**,** many years ago, when I was still a young woman.

B. *Subject-headed relative clauses. Although in some sentences there could be an overt subject, these kinds of sentences are almost categorically expressed with null subjects.*

(41) **ALB**: sí, ellos en este momento (es)tán estudiando, uno que **Ø** hace primero, segundo, así.	**ALB**: Yes, at this moment (they) are studying, there is one who is doing first (grade), second (grade), and so.

C. *Existential verbs with 'haber' [to be/to exist]. There is no a denotational subject or a pronoun in this type of sentences.*

(42) **RAM**: **Ø** hay un jagüey que se llama "4 vía(s)", una huerto [sic] al lado del jagüey

RAM: There is a reservoir named "4 ways". There is a vegetable garden next to the reservoir.

D. *Non-personal subjects. Sentences lacking an animate or denotational subject were excluded.*

(43) **MIG**: un accidente, le tropezó un carro, **eso** fue en Venezuela.

MIG: An accident, he was run over by a car, that happened in Venezuela.

E. *Verbs with uno(s), una(s) [one, ones] as their subject. These sentences are considered impersonal and are outside of the envelope of variation.*

(44) **SAR**: y **uno** empieza a cuidarse.

SAR: and one starts to take care of oneself.

F. *Clauses with impersonal SE. These sentences are out of the envelope of the variable since they are impersonal and are expressed without an overt SPP.*

(45). **NER**: sí **Ø** se celebra así, con parranda.

NER: Yes, they are celebrated with a big party.

G. *Clauses where the subject is a noun. This study focus on the expression of overt pronominal subjects and therefore finite sentences with a noun phrase as the subject are excluded.*

(46). **LUZ**: **una persona** cuando se muere.

LUZ: When a person dies.

H. *General or impersonal clauses in third-person plural. In these sentences, the subject is undetermined and consequently these types of sentences were left out of the analysis.*

(47) **ISR**: la chicha **Ø** lo preparan, **Ø** lo molen los maíz y **Ø** lo cocinan y lo **Ø** echan agua pa(ra) que no, pa(ra) que ponga así clarito, así **Ø** lo preparan la chicha.

ISR: To prepare the maize drink, they grind the corn, they cook it and pour water into it to make it clear, that's the way they make the maize drink.

I. *Clauses where the overt SPP co-occurs with 'mismo(s)' [himself/theirself]. Phrases with mismo require the compulsory realization of the overt SPP.*

(48) **ABR**: y él mismo lo dice al pueblo

ABR: And he himself says it to the people.

J. *Hesitations, repetitions, and incomplete clauses were not included. Only sentences with a subject and a finite verb are considered. Cut-off sentences or with the pronoun alone are not considered full sentences and consequently they were left out of the study.*

(49) **MIS**: ellos viven en el monte, **cuando yo**, cuando yo nací, ellos viven aquí en Uribia, Fuerza Wayuu.

MIS: They live in the countryside, when I, when I was born, they lived here in Uribia, in the Fuerza Wayuu district.

Sentences that were included are of the following types:

K. *Lexically marked contrast clauses with words such as:, 'pero' [but], 'mas' [but], 'sin embargo' [however], etc.*

(50) **SAR**: pero **Ø** no estoy de acuerdo que la persona que se haya ido a estudiar a una universidad fuera de aquí o aquí mismo en Colombia, que se le olvidó el wayuunaiki, su idioma.

SAR: But I do not agree with a person who left to study at a university out of here or here in Colombia, and she forgets the Wayuunaiki language, her mother tongue.

Although some researchers like Flores-Ferrán (2004) have left these clauses out, Otheguy et al. (2007) consider they should be included since it is possible to find contrastive clauses with a null subject.

L. *Clauses where a null SPP or an overt pronominal subject, pre-verbal or post-verbal, would have been possible, as in (51) and (52).*

(51) **YOE**: pues yo estoy estudiando en la universidad de Riohacha.

YOE: Well, I am studying at a university in Riohacha.

In (51), the subject is overt and pre-verbally realized. However, the sentence could have been produced with a null subject as well. In (52), the subject is also overt, but is pronounced post-verbally.

(52) **GEI**: hay que sacar excusa primo, porque sin excusa queda mal, le **dije yo**.

GEI: You need to have an excuse, without excuses you are in trouble, I told him.

6.8 Independent Variables

- **Grammatical variables:** grammatical person/number, reflexive use of the verb, tense, aspect and mood (TAM), and clause type.
- **Discourse variables:** coreferentiality and priming effects.
- **Semantic variables:** lexical semantics of verbs.

6.8.1 Factor groups

I. GRAMMATICAL PERSON AND NUMBER: for this factor group, we coded for the following persons: first (yo), second (*usted* and *tú*: without making any difference between specific and non-specific *tú*), and 3rd person singular (él/ella). 1st person plural (nosotros), 2nd person plural (ustedes), and 3rd person plural (ellos/ellas). This variable will inform which grammatical persons contribute to the overt SPP expression of the Wayuunaiki-Spanish bilinguals and how the results relate to parallel findings for the monolingual baseline variety.

As in other studies (Cameron, 1992; Flores-Ferrán, 2004; Orozco & Guy, 2008; among others), for this variable we expect a higher realization of overt SPPs in singular persons than in plural for both groups. If the bilingual speakers are different from the monolingual Spanish speakers, the constraint hierarchy for this variable should exhibit a different pattern in the bilingual and monolingual cohorts. Otheguy and Zentella (2012) found a preference for 2nd person singular and 1st person singular among Caribbean Spanish speakers in New York. In Colombian Spanish, Orozco and Guy (2008) found the following order: 1st person, 3rd person, and 2nd person singular. This means that comparing this variable cross-dialectally is difficult since there have been found disparate results across the board (Carvalho & Besset, 2015, p. 151).

II. LEXICAL SEMANTICS OF VERBS: following Otheguy and Zentella (2012), the verbs in this factor group were classified as follows: mental verbs for verbs of mental activity like

pensar [to think], *aprender* [to learn]; estimative verbs for verbs expressing opinions and judgements like *entender* [to understand], *creer* [to believe], *considerar* [to consider], *suponer* [to suppose]; stative verbs for verbs in which the subject is in a non-dynamic process, including *encontrarse, estar, ser* [to be]*, tener* [to have]*, vivir* [to live]*, morir* [to die]*, durar* [to last]*, olvidarse* [to forget]; and external activities verbs implying an activity, in movement, state of rest, or in development, including *comprar* [to buy]*, escribir* [to write]*, obligar* [to oblige]*, ir* [to go]*, poder* [to be able to]*, ver* [to see]. For this variable, there are conflicting results as to what kind of verbs promote overt SPP expression, however, several studies have found that cognitive-psych verbs favour overt SPP realization (Bentivoglio, 1987; Travis, 2007). The argument is that verbs expressing the speakers' opinions and verbs related to speech acts benefit overt SPP expression. If this argument holds true across varieties, then the Wayuunaiki speakers are also expected to pattern similarly to Spanish monolinguals regarding this variable.[31]

III. COREFERENTIALITY INDEX: this group has two levels, *switch reference* for a change in the subject and *coreference* when the subject is the same. Switch reference is one of the most robust variables contributing to overt SPP expression in Spanish (Carvalho, Orozco & Shin, 2015). The expectation for this variable is that overt pronouns are expressed more frequently when there is a change or switch reference across two consecutive overt SPPs. We expect this variable to be also significant in this study. If the bilingual speakers have acquired the discourse constraints on overt SPP, we will find similar patterning between the Guajiro Spanish and monolingual groups. If the bilingual speakers have not acquired the discourse constraints of this variable, the

[31] Though beyond the scope of the present chapter, a future study could add a new factor group "stative/active verbs" to find out how if they promote or not overt SPP expression in Guajiro Spanish. This is also an issue in which generative linguistics and variationist sociolinguistics can be mutually informative, since we have already done this study from the Generative Grammar perspective.

direction of the effects in this factor group will be different for the monolingual and bilingual speakers.

(53) **ELI**: yo piensa [sic], es claro, yo pienso porque tú quiere(s) tener una vida así libre.	**ELI**: I think, it is clear, I think about the reason you want to have a life without trouble.

In (53), *yo piensa* [I think] was coded as *coreference* or *no switch,* since is the same subject as in the previous clause. In the case of *tú quiere(s) tener* [you want to have] we coded the subject as a *switch* because there is a change from *yo* [I] to *tú* [you].

IV. REFLEXIVE USE OF THE VERB: this factor group has two levels. Reflexive (R) and non-reflexive (X) verbs. This variable has been significant in Otheguy et al. (2007) and Carvalho and Child (2011). In these two cases, non-reflexive verbs favour overt SPP expression. The reason is that the use of an overt pronoun with a reflexive verb may be redundant (Carvalho & Child, 2011). If this hypothesis holds true, we expect to find similar results in this study.

(54) **MIS**: pero se murieron viejitos.	**MIS**: They died when they were very old.
(55) **ROS**: él nació ahí también donde yo vivo.	**ROS**: He was born there, in the same place where I live.

In (54), the verb *se murieron* [they died] was coded as reflexive. In (55) the verb *nació* [he was born] and *vivo* [I live] were coded as non-reflexives.

V. TAM (TENSE, ASPECT AND MOOD): we coded the following distinct tenses: present, preterit, conditional, periphrastic future, future indicative, present subjunctive, past subjunctive, and perfect forms.

(56) present: **ALB:** y también **habla** el español.	**ALB**: And he also speaks Spanish.
(57) preterit: **ANA: me quedé** sin nada.	**ANA**: I was left without anything.

(58) imperfect: **ALB: yo hablaba** las dos lenguas.	**ALB**: I spoke both languages.
(59) conditional: **GEI: yo debería** bailar.	**GEI**: I should dance
(60) periphrastic future: **CAT: me voy a dormir.**	**CAT**: I am going to sleep
(61) future indicative: **GEI: adivinaremos** (este resultado).	**GEI**: We'll guess (this result)
(62) present subjunctive: **FAB: trabaje** hasta que consiga pa'l pasaje.	**FAB**: Work until you find the money to pay for your bus fare
(63) past subjunctive: **FAB:** cuando ellos **llegaran**, yo tenía que estar listo.	**FAB**: When they arrived, I had to be ready.
(64) perfect forms: **CAT:** aquí me **he quedado** viviendo.	**CAT**: I have been living here.

With this variable, we want to see if there is any specific tense positively contributing to the overt SPP expression. We expect the imperfect or the present tense to be the most important contributors. In the case of the imperfect tense, there are three competing conjugations that are potentially ambiguous, *yo, él/ella, usted* [I, he/she, you] (and in some cases *tú* [you]) if we consider that /s/ tends to be variably aspirated or deleted by the Wayuunaiki speakers and in Caribbean Spanish in general. Ambiguity is not only confined to the imperfect, in the present tense /s/ is also variably weakened in vernacular speech and /n/ suffers a process of velarization (Lipski, 1986).

TAM has been significant in Abreu (2009), Claes (2011), and Prada Pérez (2009) and this suggests that ambiguous forms motivate more overt SPP use. However, this increase in overt SPP expression in the case of the imperfect must not be necessarily taken as an adjustment in the language to compensate ambiguity, since speakers also have contextual and probabilistic information to understand what their conversational partners are trying to mean (Claes, 2011).

VI. PRIMING EFFECTS (or the production of more overt SPPs triggered by the presence of a previous overt SPP): For this variable, a previous overt SPP was coded as (P) and previous null subject as (N). Priming effects have been found to have an effect on overt SPP expression in Spanish (Cameron & Flores-Ferrán, 2004; Travis, 2007; Torres-Cacoullos & Travis, 2010; Carvalho & Child, 2011; Abreu, 2012). We expect this variable to be significant in the Guajiro Spanish and the monolingual Spanish group if priming effects are influencing overt SPP choice in our study.

(65) **ZUR**: y entonces y él le habla el español y yo le hablo el wayuunaiki pa(ra) que él entienda y las dos idiomas [sic].	**ZUR**: so, he speaks Spanish to him and I speak Wayuunaiki to him so he can understand both languages.

In example (65) the speaker expresses three consecutives overt SPPs. In this excerpt there are three pronominal subjects: *él*, *yo, é*l [he, I, he], which were coded thus as the previous pronoun was overt. The overt SPP of the first clause is connected to a sentence previously uttered by the speaker, in which the subject was also overt.

VII. CLAUSE TYPE: the operationalization of clause type aims to find which type of clause motivates the overt expression of overt SPPs. Results from other studies (Abreu, 2009; Otheguy & Zentella, 2012, *inter alia*) suggest that overt SPPs are more favoured in main clauses and less likely in dependent clauses. If the bilingual speakers of this study follow the same trend and share a similar grammar to that of the monolingual Spanish benchmark, main clauses will also have a positive effect in the production of overt SPPs in Guajiro Spanish. This variable has three levels and are coded as follows: main clauses, subordinate clauses, and coordinate clauses.

(66) Main clause: **JOS:** tengo 5 hijos.	**JOS:** I have 5 children.
(67) Subordinate clause: **GEI:** mi mama no se dio cuenta que yo estaba allí.	**GEI:** My mom did not realize that I was there.

(68) Coordinate clauses: **MIG:** se lo venden a él y lo compra. **MIG:** It is sold to him and he buys it.

6.9 Results

6.9.1 Overall rates of overt SPP expression for the bilingual and monolingual groups

The overall rate or the total percentage of expressed pronouns is an indicator that helps us to compare the bilingual participants to the Spanish native group and uncover their pattern of usage of null and overt pronominal subjects.

Table 6.3. Overall distribution of null and overt subjects for the bilingual and monolingual groups

Subject	Bilinguals		Monolinguals	
Null	N= 2166	62%	N= 478	68%
Overt	N= 1327	38%	N= 227	32%
Totals	N= 3493	100%	N= 705	100%

This analysis shows a rate of 38% of overt SPP expression for the bilingual group and 32% for the monolingual group. The difference in overt SPP expression among the two groups is significant, χ^2 (1, N = 4198) = .003671, $p < .05$. This result means that there is a lack of convergence between the two varieties under analysis in terms of the rates of overt SPP expression. However, overt SPP expression rates can fluctuate due to any kind of non-linguistic factors. For example, Orozco and Guy (2008) found a rate of 35.7% for Barranquilla Spanish, a variety very close to La Guajira. Reaching therefore conclusions about the differences between two varieties of Spanish solely on the base of overt SPP rates of usage is problematic. Overall rates vary significantly even among monolingual Spanish varieties (Torres-Cacoullos & Travis, 2015). This variation in overt SPP rates may also be motivated by different data collection methods or the communicative situation (conversations, interview-based corpus, sociolinguistic interviews) and how the data is handled and coded (Carvalho & Besset, 2015). Having a benchmark group is

therefore necessary to make comparisons. In our case, the data was obtained and organized in the same way for both groups, which makes the groups suitable for the comparative analysis. Nevertheless, this partial result must be complemented with a detailed analysis of the independent variables motivating overt SPP expression in both groups, providing a more "penetrating characterization of variable structure" (Erker and Guy, 2012, p. 546).

6.9.2 Factor groups contributing to overt SPP expression in the bilingual and monolingual data

The first run of Varbrul GoldVarb X helped us uncover some problems in the data. There were some empty cells and knockouts.[32] Consequently, the following variables were recoded. The variable *grammatical person/number* was regrouped to include *él/ella* [he/she] and *usted* [you] in the same level, as all these pronouns are conjugated with the 3rd person singular. Additionally, *ustedes* [you] was included with *ellos/ellas* [they] in a new constraint, as all these pronouns are conjugated with the 3rd person plural.[33]

The factor group TAM was also reconfigured. *Periphrastic future*, *future indicative*, *conditional*, *present* and *past subjunctive,* and *perfect* forms were included in a new factor group with the label *other tenses/moods*. This restructuring was necessary as some constraints did not have enough tokens to carry out the analysis. *Present indicative*, *preterite indicative,* and *imperfect indicative* kept their independent levels.

[32] **Knockouts** mean that a cell in the analysis has either a 0 per cent or a 100 per cent value, therefore the data is not variable, and a variable rule analysis cannot be run (Tagliamonti, 2006, p. 152).

[33] While the verb form is the same in both cases, from a theoretical point of view, merging *usted/es* with 3rd person forms is problematic because of their referential information. However, having *usted/es* in a separate group was not possible because there were very few tokens of these forms. Many cells in Varbrul were empty for *usted/es*, and this is why we either had to group them or to delete these tokens. As we assumed that the results would not radically change because of this operation, we decided to group the tokens.

The group *clause type* was also recoded into two levels: *main clauses* versus *other clauses*, which included subordinate and coordinate clauses. Finally, the *group lexical semantics of the verbs* was modified to incorporate mental and estimative verbs in the same group.

After regrouping and recoding the independent variables, we ran a new fixed-effects regression analysis in Varbrul Goldvarb X and the following factor groups came back significant: *grammatical person/number*, *priming effects, switch reference,* and *reflexive use of the verb* (only for the bilingual group). These results permitted us to obtain the factor groups that are significantly contributing to overt SPP expression in Guajiro Spanish and the monolingual group, as seen in Table 6.4 and Table 6.5.

Table 6.4. Factor groups contributing to the probability of overt SPP expression in the Wayuunaiki-Spanish bilinguals

Total N of tokens		3493		
N and % of overt subjects		1327 (38%)		
Corrected mean[34]		0.37		
Factor group	**Factors**	**Varbrul Weight**	**% of the data**	**N**
Grammatical person/number				
	Tú	.68	60	61/102
	Yo	.61	48	707/1462
	Él/ella/usted	.47	34	286/839
	Ellos/ellas/ustedes	.39	27	181/684
	Nosotros	.32	23	92/406
	Range[35]	36		

[34] **Corrected mean** is "the overall likelihood that speakers will choose the variant selected as the application value" (Bayley, 2013, p. 93) In our study, the application value was the overt pronominal subject. The corrected mean was 0.37 for the bilingual speakers and 0.31 for the monolingual group, this indicates that the bilingual will choose an overt SPP 37% of the time and the monolingual group 31% of the time.

[35] **Range** is the number "obtained by subtracting the lowest factor weight from the highest factor weight" (Tagliamonte, 2006, p. 251). For example, in the factor group or the variable *grammatical person/number*, we subtracted .32 (the factor weight for *nosotros*) from .68 (the factor weight from *tú*) and we obtained .36, (cf. Table 6.5). The range or magnitude of effect tells that in this factor group, the strongest constraint is *tú* and the weakest constraint is *nosotros*. The range also helps to determine which factor group is most significant and which factor group

Priming effects				
	Previous pronoun was overt	.68	56	695/1238
	Previous pronoun was null	.40	28	632/2255
	Range	28		
Coreferentiality index				
	Switch reference	.58	43	530/1236
	Same reference	.46	28	797/2257
	Range	12		
Reflexive use of the verb				
	Non-reflexive use	.51	39	1236/3180
	Reflexive use	.41	29	91/313
	Range	10		
$P = 0.007$				
Log likelihood -2095.195				

Varbrul did not select the following factor groups as significant: *TAM, clause type and lexical semantics of the verbs*.

The results displayed in Table 6.4 and Table 6.5 (below) indicate that, for the bilingual group, the most important factor as gauged by the range value is *grammatical person/number* followed by *priming effects*, *switch reference*, and *reflexive use of the verb*. For the monolingual group the results are similar (except for *reflexive use of the verb*, which was selected as non-significant by the variable rule program in the step-up step-down procedure in the Spanish group).

These findings indicate then the same independent variables are selected as significant in the comparison groups and the factor groups operate in the monolingual and bilingual grammars in the same order of significance. Although the range value of the variables differs across the two

is least significant (Tagliamonte, 2006, p. 253). For the bilingual group, the variable *grammatical person/number* is the most significant (range 36) and the variable *reflexive use of the* verb is the least significant (range 10).

independent analyses, the ranking of the variables is similar in the two groups under analysis, which is the key heuristic for comparison.

Table 6.5. Factor groups contributing to the probability of overt SPP expression in the Spanish monolinguals

Total N of tokens		705		
N and % of overt subjects		227 (33%)		
Corrected mean		0.31		
Factor group	**Factors**	**Varbrul Weight**	**% of the data**	**N**
Grammatical person/number				
	Tú	.73	52	26/50
	Él/ella/usted	.61	44	69/158
	Yo	.51	32	101/311
	Nosotros	.34	27	13/48
	Ellos/ellas/ustedes	.33	18	25/138
	Range	40		
Priming effects				
	Previous pronoun was overt	.66	47	104/219
	Previous pronoun was null	.43	25	123/486
	Range	23		
Coreferentiality index				
	Switch reference	.62	41	116/284

	Same reference	.42	36	111/310
	Range	20		
Reflexive use of the verb				
	Non-reflexive use	[ ]	33	208/636
	Refkexive use	[ ]	29	20/69
	Range			
P = 0.011				
Log likelihood= -403.349				

Varbrul did not select the following factor groups as significant: *TAM, clause type, lexical semantics of the verbs and reflexive use of the verb*.

To have a thorough analysis of the issue we are investigating, we still need to inspect the constraint hierarchy of each factor group to see how they behave and compare in the two groups of speakers under examination.[36] Analysing how the factors are ordered within the factor groups will permit us to assess the possible effects of language contact, if the factor groups display different constraint rankings there is an indication the underlying grammar of the monolingual Spanish and the bilingual speakers is not the same.

[36] **The hierarchy of constraints** is the ranking of the factors in a factor group as obtained by organizing them according to their Varbrul weight, being the factor with the closest weight to 1.00, the first in the ranking, and so on.

6.9.3 GRAMMATICAL PERSON/NUMBER

The results for the factor group *Grammatical person/number* reveal that for the bilingual group, an overt SPP is more likely to occur with the pronoun *tú* [you] than with any of the other personal pronouns. The second place in the hierarchy of constraints for this factor group is for *yo* [I] and the third place is for *ellos/ellas/ustedes* [they/they/you].

Table 6.6. Grammatical person/number and expression of pronominal subjects in bilingual and monolingual Spanish

Rank	Bilingual group	Varbrul[37] weight	%	N	Monolingual group	Varbrul weight	%	N
1	Tú	.68	60	61/102	Tú	.73	52	26/50
2	Yo	.61	48	707/1462	Él/ella/usted	.61	44	69/158
3	Él/ella/usted	.47	34	286/839	Yo	.51	32	101/311
4	Ellos/ellas/ ustedes	.39	27	181/684	Nosotros	.34	27	13/48
5	Nosotros	.32	23	92/406	Ellos/ellas/ Ustedes	.33	18	25/138
Log likelihood[38] = -2095.195					Log likelihood = -403.349			
Significance = 0.007					Significance = 0.011			

Having *tú* [you] as the main constraint could be due to the fact that /s/ is variably deleted or aspirated in Guajiro Spanish and the speakers are avoiding ambiguity with other conjugations like the 3rd person singular. This hypothesis was proposed by Hochberg (1986 in Flores-Ferrán, 2004) to account for overt SPP expression for Puerto Rico Spanish and contended by Cameron

[37] **Varbrul weight** is a number that "provides a numerical measure of the strength or influence of each factor, relative to other factors in the same group, on the occurrence of the linguistic variable under investigation. Values range between 0 and 1.00. A value, or weight, between .50 and 1.00 indicates that the factor favors use of a variant relative to other factors in the same group" (Bayley, 2013 p. 92). If we analyse the factor group *reflexive use of the verb* in Table 6.8, we have that an overt SPP is more likely to appear with a non-reflexive verb and is unlikely to occur with a reflexive verb. In this sense, factor weights are probability co-efficients, they don't provide statistical significance.

[38] The **log likelihood** is "a measure of the fit of the model to the data" (Tagliamonte, 2006, p. 224). The closer the figure to zero, the better the model.

(1995), who said in his comparative study of Puerto Rico Spanish and Madrid Spanish that the difference in overt SPP rate of expression for *tú* [you] lies in the specificity of this pronoun. According to Cameron (1995, p. 328) in San Juan, Puerto Rico, "non-specific *tú*, with respect to specific *tú*, correlates to an increase in pronominal subject expression".[39]

1st person singular is the second most important constraint for the factor group *grammatical person/number*. This result can be explained by the use of interviews as the data collection methodology. Travis (2007) explains that when speakers are the most prominent person in the conversation, the participants tend to systematically realize overt pronouns when talking about themselves. Carvalho and Child (2011) suggest that an interview is not a "natural" situation, and this could trigger a tendency to express more overt SPPs in 1st person. Additionally, the participant would tend to realize more overt pronouns when talking about themselves to emphasize their presence and points of views. However, Torres-Cacoullos and Travis (2018, p. 74) argue that the higher expression of overt pronominal subjects with *yo* [I] is related as well to some lexical particular constructions like *yo creo* [I think].

With respect to the monolingual group, the analysis also returned the factor group *grammatical person/number* as the most significant for these speakers. However, the bilingual and monolingual groups diverge in the constraint hierarchy. Although they both had *tú* [you] as the pronoun most likely to be overtly expressed, there is a subtle permutation in the ranking of constraints as generated by the monolingual group having *él/ella/usted* [he/she/you] in the second place in the hierarchy and *yo* [I] in the third place. This could suggest the control group are telling more 3rd person narratives than 1st person ones vis-à-vis the bilingual group. Tagliamonte (2006,

[39] In our study we did not investigate this difference. Further research could investigate this and whether *tú* in other Null Subject Languages such as as Portuguese, Italian and Greek has the same special status it has in Spanish and whether specifity is also correlated with an increase in overt SPP expression, independently of how the verb form is realized in each language.

p. 241) argues that in these situations "the underlying mechanism producing the surface form is different", and that may be the case for the monolingual and bilingual groups of this investigation, at least for this specific factor group.

The monolingual and bilingual group coincide in the fact that singular pronouns are realized more frequently than plural pronouns. This result was expected because cross-dialectally and in language-contact situations overt SPPs are more likely expressed with singular referents (Carvalho and Besset, 2015, p. 152).

The factor group *grammatical person/number* has also been selected as significant in other studies (Flores-Ferrán, 2004; Otheguy et al*,* 2007; Orozco & Guy, 2008; Carvalho & Child, 2011; Abreu, 2012; Holmquist, 2012; and Shin, 2012). In the case of Orozco & Guy (2008) for Barranquilla Spanish, 1st person singular was the main constraint. However, 2nd person *tú* and *usted* [you and you formal] ranked third.

In Barranquilla Spanish, 1st person plural *nosotros* [we] ranked last, similar to the results presented here for the bilingual group, implying that this is the person that least favours an overt SPP. A possible explanation for this result may be the fact that *nosotros* [we] uses a conjugation with a highly salient morphological ending (-mos), so using an overt pronoun in 1st person plural may look redundant (Orozco & Guy, 2008).

2nd person singular and 1st person singular are the persons that contribute the most to overt SPP expression across the Spanish varieties. Mayol (2012) says that 2nd person is a pronoun constantly showing a high rate of overt SPP expression in Spanish. In this sense, the direction of the constraints of the Wayuunaiki speakers for this factor group is not *sui generis* and confirms a tendency to select these persons as the main contributors of overt SPP expression for the group *grammatical person/number*. A similar direction is found, for example, in Castañer Spanish in

Puerto Rico (Holmquist, 2012), as *tú* [you] was also the first constraint and *yo* [I] the second most important constraint.

6.9.4 PRIMING EFFECTS

For priming effects, we only took into account the previous inflected verb. Priming effects can be coded in different ways, for example, Carvalho and Child (2011) coded up to three consecutives occurrences of an inflected verb. However, following Travis (2007) and Torres-Cacoullos and Travis (2010), we coded only for the immediate previous token.

In this study, the second most important factor group is *priming effects* or the "tendency of linguistic forms to occur together within a stretch of discourse" (Carvalho & Besset, 2015, p. 157), this variable was significant for the monolingual and bilingual groups and was the second most important in both cases. The effect was identical in both groups, an overt SPP has a greater likelihood of being overtly realized when the preceding clause has also an overt SPP. When the previous clause has a null subject, an overt SPP is disfavoured in the next clause.

Table 6.7. Priming effects and expression of pronominal subjects in bilingual and monolingual Spanish

Rank	Bilingual group	Varbrul weight	%	N	Monolingual group	Varbrul weight	%	N
1	Previous pronoun was overt	.68	56	695/1238	Previous pronoun was overt	.66	47	104/219
2	Previous pronoun was null	.40	28	632/2255	Previous pronoun was null	.43	25	123/486
Log likelihood = -2095.195					Log likelihood = -403.349			
Significance = 0.007					Significance = 0.011			

The factor *priming effects* or *perseveration effects* has been identified as a significant factor in other varieties as well (Cameron & Flores-Ferrán, 2004; Travis 2007; Torres-Cacoullos &

Travis, 2010; Carvalho & Child, 2011; Abreu, 2012). This factor has also been considered a reason for the differences on the rates of overt SPP expression among different dialects (Mayol, 2012). Cameron and Flores-Ferrán (2004, p. 41) argue that the "expression of subject pronouns leads to more pronouns, and expression of null subjects leads to more nulls". They say that given two sentences, the trigger of the first one spreads to the next one. Therefore, a sentence with a null subject will spread a similar form to the following sentence and a sentence with an overt SPP will spread this overt pronoun to the coming one.

The priming effects factor was also significant in English-Spanish bilinguals of Florida (Abreu, 2012), Spanish-Portuguese bilinguals of Rivera, Uruguay (Carvalho & Child, 2011), and Spanish in contact with Portuguese (Carvalho & Besset, 2015). The fact that priming effects are significant and the direction of the effects of this factor group is the same cross-linguistically lends support to the idea that the constraints at work for this factor group are universal and not a property of a specific language (Tagliamonte, 2006, p. 240).

6.9.5 COREFERENTIALITY INDEX

This factor is the third most important group. It has the same effects for the bilingual and monolingual group. A switch in reference favors the appearance of an explicit SPP in Guajiro Spanish and in the control group. If the reference is the same in the previous clause, consequently the expression of an overt subject is disfavoured. These results indicate then that both groups pattern in the same direction regarding this specific factor group.

Table 6.8. Coreferentiality index and overt SPP expression in bilingual and monolingual Spanish

Rank	Bilingual group	Varbrul weight	%	N	Monolingual group	Varbrul weight	%	N
1	Switch reference	.58	43	530/1236	Switch reference	.62	41	116/284
2	Same reference	.46	28	797/2257	Same reference	.42	36	111/310
Log likelihood = -2095.195					Log likelihood = -403.349			
Significance = 0.007					Significance = 0.011			

The so-called continuity of reference, discourse connection or switch reference, is one of the strongest variables boosting overt SPP expression in Spanish. It has proven to be significant in New York City Puerto Ricans (Flores-Ferrán, 2004), Cuba (Sánchez Arroba, 2011), Puerto Rico (Holmquist, 2012) Colombia (Hurtado, 2005) and bilingual Spanish (Carvalho & Child, 2011; Abreu, 2012; Carvalho & Besset, 2015). Shin and Otheguy (2009, p. 111) say that continuity of reference is a variable "that was found to work the same way in all varieties of Spanish where subject pronouns have been studied". In this sense, the Wayuunaiki speakers and the monolingual Spanish speakers of this study pattern like other contact and non-contact Spanish varieties and this also supports the universality of this variable.

6.9.6 Reflexive use of the verb

This factor group was only significant for the bilingual group. The ranking of constraints shows that an overt SPP is more likely to occur with a non-reflexive verb than with a reflexive verb. The result is as expected, as it is considered that the reflexive pronouns add more referential information, making the use of the overt SPP redundant (Carvalho & Child, 2011).

Although this factor group did not return significant for the monolingual group, the direction of the effects is the same as the one found for the bilingual cohort. This suggests these constraints work similarly in both grammars.

Table 6.9. Reflexive use of the verb and expression of pronominal subjects in bilingual and monolingual Spanish

Rank	Bilingual group	Varbrul weight	%	N	Monolingual group	Varbrul weight	%	N
1	Non-reflexive use	.51	39	1236/3180	Non-reflexive use	[ ]	33	208/636
2	Reflexive use	.41	29	91/313	Reflexive use	[ ]	29	20/69
Log likelihood = -2095.195								
Significance = 0.007								

Regarding other language contact situations, this factor group was also significant in Spanish-Portuguese speakers with similar results, that is, non-reflexive use of the verb promotes more overt SPP expression than the reflexive use of the verb (Carvalho & Child, 2011). As of the moment of writing this dissertation and according to Carvalho, Orozco, and Shin (2015), this variable has never been selected as significant in monolingual settings, which is the same result found for the monolingual group of this study.

6.10 Discussion

The first objective of this research was to elucidate the rates of overt SPP expression for the bilingual and monolingual speakers. We found a significant difference among the groups as the bilingual group produced overt SPPs at a rate of 38% and the monolingual group a rate of 32%. This partial result indicates a lack of convergence between the groups. Nonetheless, the rates of overt SPP expression alone do not present the whole picture, since they can fluctuate based on any kind of non-linguistic factors and the rate discrepancies are not a reliable diagnostic to establish

any underlying differences, in fact, the rate divergences may be hiding subjacent parallelisms in the structure of the variability in the two comparison groups.

Reviewing the factor groups is more revealing of what happens between the grammars under comparison. For this reason, the second and third research questions were aimed at identifying and characterizing the factors that condition the variable use of overt SPPs in the bilingual and monolingual groups. The Varbrul analysis produced four significant factor groups for the bilingual cohort and three factor groups for the monolingual participants (Tables 6.4 and Table 6.5).

In the case of the bilingual speakers, the strongest factor was *grammatical person/number*, followed by *priming effects*, then *coreferentiality index,* and finally *reflexive use of the verb*. For the monolingual group, the strongest factor was *grammatical person/number*, followed by came *priming effects*, then *coreferentiality index.* This means that for both groups, the expression of an overt SPP is also highly conditioned by the grammatical person and number of the pronoun while *priming effects* and *coreferentiality index* are secondary.

Finally, we analysed the direction of effects of the constraint hierarchies. The greatest difference in the ranking of constraints was found in the factor group *grammatical person/number*. *Tú* [you] was the first constraint for both groups, however, *yo* [I] ranked second for the bilinguals and third for the monolinguals. Additionally, the constraint *él/ella/usted* [he/she/you] favoured overt SPP expression in the monolingual group, but not for the bilingual speakers. This subtle permutation in the ranking of the constraints suggests the underlying systems of the monolingual and bilingual groups differ in this specific window of the grammar. This difference may be motivated by a distinct preference of the speakers to tell the narratives during the interview. While

the bilingual participants could have told more narratives with *yo* [I], the monolingual speakers favoured narratives with *él/ella/usted* [he/she/you].

Regarding the priming effects and coreferentiality, the directions of the effects were similar across both group of speakers. This suggests the grammars of the bilingual and monolingual groups are basically the same regarding these factors, plus these factors operate in a universal way across the varieties of Spanish. Regarding reflexivity, this factor group did not have any effect for the monolingual group but seems to be operative in the bilingual group. Nevertheless, the direction of effects is similar for both cohorts.

6.11 SUMMARY

The aim of this chapter was two-fold: (i) obtaining and comparing the rates of usage of overt pronominal subjects for Guajiro Spanish and the Spanish benchmark groups and; (ii) uncovering the constraints and their direction of effects in the two cohorts of speakers under investigation.

The variable grammars of the Wayuunaiki-Spanish bilinguals and the Spanish monolinguals exhibited some differences: i) the rate of overt SPP expression is different for each group; ii) there is a quantitative difference in the number of significant factor groups found, as the bilingual group added the variable *the reflexive use of the verb*; iii) the direction of the effects presented a subtle difference in the case of the group *grammatical person/number*. In the rest of the variables, the direction of the effects is similar for the bilingual and monolingual groups.

The findings of this chapter allow us to propose hence that the two varieties under comparison are independent codes. Nonetheless, for the portion of the grammar studied here, the underlying systems of overt SPP expression are not radically dissimilar even if there are some divergencies; however, the convergencies are more obvious; the two varieties under scrutiny

presented parallel hierarchy of constraints, the chief heuristic to tell if there are non-trivial differences.

The same factor groups or independent variables were selected as having a significant effect on overt SPP for the bilingual and monolingual speakers and the constraint hierarchies were very parallel across the two comparison groups. This tell us the two Spanish varieties under analysis share basically the same underlying constraints that promote overt SPP expression, even if their overall rates of overt SPP usage differ.

In a broader view, the results of this chapter support as well the argument that the same linguistic variables operate similarly across the diverse varieties of Spanish: bilingual and monolingual equally. Carvalho, Orozco and Shin (2015) affirm that overt SPP expression in Spanish is affected frequently by the same grammatical conditioning and this is confirmed by the findings discussed here.

7. SUBJECT- VERB MISMATCHES IN GUAJIRO SPANISH

7.1 INTRODUCTION

In this chapter, we analyse the subject-verb agreement mismatches produced by the 27 bilingual speakers of Wayuunaiki-Spanish interviewed in the Province of La Guajira, Colombia (see Chapter 4). Additionally, as a control group, we have five monolingual Spanish speakers from the same area. For the purpose of this study, a mismatch is understood as a sentence that does not conform to the vernacular benchmark or everyday speech of the community under analysis, in which the subject-verb agreement between person and number is normally realized as in Standard Spanish but is variable for *tú* [you]. For example, in (69) the /s/ is not realised and in (70) it is pronounced in liaison with the adjacent vowel.

(69) **Interviewer:** qué bien, y algo que se fabrique localmente aquí en Maicao o aquí solo se venden cosas que vienen de otras partes.	**Interviewer:** How good, is there something made locally here in Maicao or do you just sell stuff made somewhere else?
YOE: no, de aquí de Maicao, Maicao no, lo único que te puedo decir e(s) la(s) venta de la(s) mochilas y de las manta(s) que vienen de aquí, propio de Maicao, porque tú te puede(s) decir de la ropa, uno la compra aquí en Maicao, pero viene, ya tú sabe(s), fabricada de otras partes, como las chanclas…	**YOE:** No, from Maicao, from Maicao no, the only things that I can say that are made here are the Wayuu bags and manta dresses. The rest of the clothing and flip-flops, for example, are sold here, but they are made somewhere else.
(70) **Interviewer:** ¿las hamacas? ¿qué es eso?	**Interviewer:** Hammocks? what is that?
YOE: pues como la que tú ves allá, la que tú ves ahí.	**YOE:** Well, like the one you see there, the one you see there.

Because of this, the sequences with *tú* [you] are not included because they are subject to the elision or weakening of /s/ in this variety (Montes, 1982), and it will be difficult here to determine whether the tokens produced with the personal pronoun *tú* [you] are instances of incomplete Spanish acquisition or a replication of the Spanish vernacular to which bilingual

speakers are routinely exposed. Having excluded tokens with *tú* [you], we will focus exclusively in the analysis of sequences that point unambiguously to problems related with the acquisition of features, as in examples (71a) to (75a).[40]

In example (71a), there is a mismatch between the subject pronoun *yo* [I], that represents a 1st person singular and the verb *sabe* [he/she knows] that is marked as a 3rd person singular.[41] In (72a), the mismatch is between the 3rd person singular and the verb in 3rd person plural. In (73a), there is also a mismatch, but in this case, it works at two levels. There is not only a mismatch at the level of person, but also at the level of number. *Nosotros* [we] is a 1st person plural subject pronoun, whereas the verb marks for a 3rd person singular. In (74a), the mismatched sentence is also affected at two levels, 2nd person plural clashes in person/number with 3rd person singular. In (75a), 3rd person plural clashes with 3rd person singular, here the mismatch is at the level of number since both the subject and the conjugation coincide in person. All the previous mismatched sentences are expected to be as in (71b) - (75b).

(71)	a.	Yo	no	sabe	escribir.	1S ≠ 3S (single mismatch)	(JOS)
		PRO.**1S**	not	know.**3S**	write.INF		
	b.	Yo	no	sé	escribir.	1S = 1S (match)	
		PRO.**1S**	not	know.**1S**	escribir.		
		'I don't know how to write'.					

(72)	a.	(Él)	no	ayudan.	3S ≠ 2P/3P (single mismatch)	(ANE)
		PRO.**3S**	not	help.**2P/3P**		
	b.	(Él)	no	ayuda	1S = 1S (match)	
		PRO.**3S**	not	help. **3S**		
		'He doesn't help'.				

[40] Unambiguity is problematic for 2P/3P as well. The velarization and weakening of /n/ have been also reported for Colombian Caribbean Spanish (Flórez, 1963, p. 272; Canfield, 1981, p. 7; Montes, 1982, p. 45; Aleza, 2010, p. 90). Lipski (1986, p. 141) suggests that the velarization of /n/ is a natural step that eventually leads to the elision of this consonant and to the nasalization of the preceding vowel.

[41] This conjugation also marks 2nd person singular (formal).

(73) a. Nosotros viene de allá. 1P ≠ 3S (double mismatch) (NAR)
PRO.**1P** come.**3S** from There
b. Nosotros venimos de allá. 1P = 1P (match)
PRO.**1P** come.**1P** from There
'We come from there'.

(74) a. Ustedes cuando describe una persona. 2P ≠ 3S (double mismatch) (ZUR)
PRO.**2P** when describe.**3S** a person
b. Ustedes cuando describen una persona. 2P = 2P (match)
PRO.**2P** when describe.**2P** a person
'When you describe a person'.

(75) a. (Ellos) sí habla español. 3P ≠ 3S (single mismatch) (ANI)
PRO.3**P** yes. speak.**3S** Spanish
b. (Ellos) sí hablan español. 3P = 3P (match)
PRO.3**P** yes. speak.**3P** Spanish
'They do speak Spanish'.

As mentioned, mismatched sequences with 2P/3P are problematic since the bilingual speakers may be replicating a natural vernacular process as in examples (84, token #12) and (85, token #30), in which these two monolingual Spanish speakers dropped the /n/ sound. Nonetheless, we are analysing these mismatches because, as it is shown in the results section, the higher percentage of mismatches of the bilingual group for 2P/3P overrides the possibility that all these sentences are purely a replication of the Spanish vernacular. Additionally, the Spanish monolingual group only produced 2 tokens of /n/ deletion or 0.5% of this kind of instances within the whole universe of their data.

The mismatches of this study are analysed within the Generative Grammar framework, specifically the theory of features. The chapter is organized as follows. First, a brief discussion of the relationship between features and agreement is presented. Second, we review previous studies done on mismatches in Guajiro Spanish. Third, a brief overview of subject-verb agreement in Wayuunaiki is discussed. Fourth, we develop the research questions and the hypotheses of this

study. Fifth, we present and analyse the mismatches that occur in the bilingual data. Finally, the results of the analysis are discussed and summarized.

7.2 The Theory of Features

Features, specifically morphosyntactic features in this case, are the minimal building units necessary to generate lexical categories and the abstract constructs used by linguists to generalize and theorize about the different facts of language. Features have been metaphorically referred to as the atoms of language by Adger (2003) and by Liceras, Zobl, and Goodluck (2008) as the DNA (base pairs) of human languages, the functional categories being the genes that give shape to the structure of a specific language. Commonly known features (and their values) are number (singular, plural, etc.); person (first, second, third); gender (masculine, feminine, etc.); case (nominative, accusative, etc.); and aspect (perfective, imperfective) (Corbett, 2012).

The notion of features has been part of the Generative Grammar ever since the inception of this theory, however, the approaches to features have varied over time. In Chomsky (1965), features expressed properties of lexical items: selectional [± abstract] and subcategorial [± transitive] (Liceras, Zobl & Goodluck, 2008). In Chomsky (1970), features were used as binary values to define syntactic categories: [± noun], [± verb], [± adjective], and [± preposition] (Liceras, Zobl, & Goodluck, 2008). Features such as number, person, gender, etc. were defined as the minimal units which make up functional categories such as determiners, auxiliaries, pronominals, etc. In the 1990's, features became central to the Minimalist Program (Chomsky, 1993; 1995) and now play an important role in the basic operations of Merge, Agree, and Move (Liceras, Zobl & Goodluck, 2008). Thus, within this framework, the functional categories are made up of features, and the feature(s) of a given constituent must be checked or valued against those of another

constituent to establish a relationship, this checking mechanism triggers, for example, the formal operation of Agree (Chomsky, 2001).

Features have become invaluable to describe and analyse linguistic phenomena and to compare languages since they allow us to generalize and put forward explanations about linguistic facts that are sometimes realized differently across the languages of the world.

7.3 On Agreement

Agreement is the process by which the grammatical features or values of two different constituents are matched. This asymmetrical relationship holds when a constituent lacking a certain feature (the probe) matches its value on another constituent (the goal). These features could be the aforementioned gender, person, or number, etc. and its corresponding values.

In the specific case that we discuss in this chapter, subject-verb agreement in Spanish is realized when the verb agrees in person and number with its related overt or null subject in finite clauses (Zagona, 2002), as seen in (76) and (77) below. When this requirement is not met, mismatched sentences like those as in (78) and (79) are generated.

(76) (Yo) llego temprano. (RAM)
(PRO.**1S**) arrive.**1S** early
'(I) drink black tea'.

(77) Ustedes traen su comida. (GEI)
PRO.**2P** bring.**2P** you food
'You bring your food'.

(78) Yo sabe leer. (ADE)
PRO.**1S** know.**3S** read
'I know how to read'.

(79) Nosotros habla wayuunaiki. (ANI)
PRO.**1P** speak.**3S** Wayuunaiki
'We speak Wayuunaiki'.

In example (76), the constituents (a null subject and the verb *llego* [I arrive]) agree in number (singular) and person (1st). In (77), the constituents agree in number (plural) and person (2nd). On the other hand, in (78), there is a mismatch between the subject pronoun and the conjugated verb as the pronoun is marked as 1st person singular while the verb is marked for a 3rd person singular. In this case the mismatch lies with the feature person since the verb does not value a 1st person singular as required by the expressed subject. In (79), the mismatch is due to both the person and the number features since the subject is marked as 1st person plural and the verb person affix is Ø, which corresponds to 3rd person singular.

What we have seen in the cases of (78) and (79) is that the probe (the subject) cannot copy the features corresponding to the subject pronoun into the goal (the verb). The probe-goal relationship is one in which the features of the probe are specified first and then are those of the goal, so that the probe and the goal match their identities. When this matching identity fails, subject-verb agreement mismatches are generated (Chomsky, 2000).

These types of morphological mismatches have been accounted for from different perspectives. The Impairment Representation Hypothesis (IRH) proposes that adult L2 learners have trouble accessing Universal Grammar (UG) functional categories, the features linked to functional categories, and the mechanisms to carry out this process (Eubank, 1993 & 1994; Meisel, 1997). Since these interlanguage (IL) grammars do not have these mechanisms and may not check features or have underspecified features, agreement mismatches are expected (Meisel, 1991; 1997) as the V(erb) features cannot be checked within the functional category Infl(ection).

On the other hand, for the Missing Surface Inflection Hypothesis (MSIH) (Prévost & White, 2000b; Prévost, 2008), (IL) grammars do have access to UG and consequently to functional categories and feature checking mechanisms. Learners may not be able to access these abstract

properties systematically, which may lead to the production of subject-verb mismatches. However, the speakers "have the unconscious knowledge of the functional projections and features underlying tense and agreement" (Prévost & White, 2000b; p. 103). That is to say, the problem lies with the production, in which the speakers may use another form instead of the expected one. It is not because the speakers do not have the knowledge of functional categories or features related to agreement or the feature-checking mechanisms, but rather because they have trouble realizing the correct overt morphology.

Another explanation for agreement errors comes from McCarthy (2004, p. 1) when she proposes the Morphological Underspecification Hypothesis (MUH). According to this hypothesis, 3rd person singular is the expected form in contexts when there is a problem of underspecification of features, that is to say, if there are problems of verb-subject agreement, the 3rd person singular is the candidate to emerge as the default conjugation because this form has zero inflection regarding person/number.

7.4 Subject-Verb Agreement Mismatches in Guajiro Spanish

The issue of subject-verb agreement mismatches in Guajiro Spanish has been barely described or explained. Ramírez Cruz (2009) describes subject-verb agreement mismatches in Guajiro Spanish, however, he does not provide a specific explanation for this phenomenon. He argues that it is difficult to attribute these sequences to language interference from Wayuunaiki and that the reason lies in the intrinsic difficulty of Spanish grammar and particularly the Spanish conjugations. Ramírez Cruz considers subject-verb mismatches to be the realization of a reduced and simplified system (the Guajiro Spanish interlanguage) that characterizes the grammar of Wayuunaiki speakers because they do not have enough knowledge of Spanish.

Pimienta (2008, p. 55) tries to provide an explanation from the Wayuunaiki perspective. She argues that these subject-verb agreement mismatches are due to the necessity of the Wayuunaiki speakers to express the gender suffixes (*shi, sü*) in Spanish that are present in the analytical conjugation of Wayuunaiki (see Table 7.1 and examples 80-82). For Pimienta, in a sentence like *yo vive* [I lives]$_{[3p.s]}$, the Wayuunaiki speakers are expressing the gender features their language has in the analytical construction by means of the 3rd person singular. This explanation goes along the same lines as Nicol and Greth's (2003) idea that bilinguals transfer agreement strategies from their L1 to their L2, making the acquisition of the L2 agreement patterns more difficult whenever the L1 and the L2 differ in how features are checked and realized.

7.5 Subject-Verb Agreement in Wayuunaiki

Full subject verb agreement in Wayuunaiki is only possible in the analytical conjugation. In this construction, the Wayuunaiki verb has features [+number] and [+gender] only in singular and lacks [person]. If a Wayuunaiki verb were conjugated along the lines of the Spanish conjugation, there would be only three forms rather than six in Standard Spanish and two in English, (Table 7.1). Nonetheless, variable deletion processes in speech can neutralize the differences between 2S and 3S as well as 2P and 3P in some varieties of Spanish.

Table 7.1. Conjugation of *ayonnajaa* 'to dance' (Álvarez, 2017, p. 67)

Spanish	Wayuunaiki	English
yo bailo	ayonnajü**shi** taya (M)	I dance
	ayonnajü**sü** taya (F)	
tú bailas/usted baila	ayonnajüshi pia (M)	you dance
	ayonnajüsü pia (F)	
él baila	ayonnajüshi nia	he dances
ella baila	ayonnajüsü shia	she dances
nosotros bailamos	ayonnajü**shii** waya (Pl.)	we dance
ustedes bailan	ayonnajüshii jia (Pl.)	you dance
ellos/ellas bailan	ayonnajüshii naya (Pl.)	they dance

In (80), the verb and the subject agree in gender and number. The subject is *Camilo*, and this is reflected in the verb having the masculine/singular suffix *-shi*. This suffix's information is also useful to know the gender of the subject personal pronoun when the subject is not a NP, as in (81) and (82).

(80) Ayonnajüshi Kamiirü.
a-yonnaja-**shi** Camilo
0-dance-M Camilo
'Camilo dances'.

(Álvarez, 2017, p. 66)

In example (81), the subject is masculine and in (82) the subject is feminine, as per the information about gender provided by the endings *-shi* and -s*ü*.

(81) Ekerot**shi** taya tepialu'um*üin.*
a-kerola-**shi** taya ta-pia-lu'u-m*üin*
0-enter-M I 1S-house-inside-towards
'I enter the house'.

(Álvarez, 2002, p. 3)

(82) O'unus***ü*** taya.
a-'una-s*ü* taya
0-go-F I
'I go'.

(Mansen & Mansen, 1984, p. 36)

7.6 THIS STUDY

The different requirements for subject-verb agreement checking in Spanish and Wayuunaiki present a potential locus of variability or a conflict site (Table 7.2). In this sense, the present study analyses and compares subject-verb agreement in the Spanish data produced by the Wayuunaiki-Spanish bilinguals and the Spanish native monolinguals.

Table 7.2. Summary of the verb features of Spanish and Wayuunaiki

Spanish	Wayuunaiki
[+person]	[-person]
[+number]	[+number]
[-gender]	[+gender]

7.6.1 SELECTION OF TOKENS

In the current study, we try to account for the mismatches between a subject personal pronoun and a conjugated verb that occurred in the data of the Guajiro Spanish speakers and relate these results to the monolingual Spanish group.

For the bilingual group, the overall number of retained tokens was 2336. From this sample, we extracted 431 mismatched sequences; 131 were produced with an overt subject pronoun and 300 were realized with a null subject pronoun.

To extract the mismatched sentences that were produced with a null subject, we selected the sentences that were unambiguous regarding the expected subject, as we could infer from the discourse context. In (84), the conversation is about the women in the rancherías. Although the Wayuunaiki speaker uses a 3rd person singular form at a times, we can assume he is referring to a plural subject *las mujeres* = *ellas* [the women = them] that has been mentioned before and about which we ask more about.

(83) **Interviewer:** Ok., ¿y las mujeres qué hacen?	**Interviewer:** Ok., and what do women do?
RAM: y la(s) mujere(s) cocinando y hacen comida, por ahí mediodía busca_ agua, de la tarde busca_ agua…	**RAM:** And the women cooking, and they prepare the food, around noon they search for water, in the afternoon they search for water…
Interviewer: ¿y qué más?	**Interviewer:** And what else?
RAM: hacen desayuno.	**RAM:** They make breakfast.

For the monolingual Spanish group, 391 tokens were selected. Two of these tokens were produced with the elision of /n/, as showed in the examples (84) and (85), in which the tokens #12 and #30 presented the omission of the nasal sound.

(84) **Interviewer:** ¿cuáles son las (mochilas) más caras?	**Interviewer:** Which are the most expensive bags?
YOE: las má(s) cara(s) son las que son abordadas [sic]	**YOE:** The most expensive ones are those which are embroidered.
Interviewer: pero ¿enteras bordadas o tienen bordados en algunas partes?	**Interviewer:** Are they embroidered everywhere or just in some places?
YOE: tiene #**12** bordados por toda(s) parte(s) y como le ponen tantas cosas que son más caras, se van encontrando a 80, 100.	**YOE:** They are embroidered everywhere, and because they add on more expensive decorations, they cost around 80, 100 pesos.

(85) **Interviewer:** ¿y cómo aprendió a hacer estas cosas?	**Interviewer:** And how did you learn how to make those things?
YAN: viendo a las paisana(s), como las hacen, y si uno no sabe algo uno le(s) pide el favor, vienen y lo ayuda #**30**, lo hacen, le explican a uno.	**YAN:** By looking at the Wayuu women, how they make them, and if one doesn't know how to do it, one can ask the Wayuu women the favor to help, and they help, they do it, they explain to us.

The tokens retained for this analysis were all in the present tense.[42] Psych verbs that require an experiencer and theme were not included (e.g., *gustar* [to like]). The tokens were in interrogative or declarative clauses with pre-verbal and post-verbal pronominal subjects, as in (86). In this example, token #101 is in present tense in a declarative sentence and the subject is in pre-verbal situation, whereas token #102 is in the present tense in a declarative sentence but the subject is in post-verbal position.

(86) **Interviewer:** Ok., y en las rancherías cómo, cómo hacen por ejemplo, ¿viven todos juntos?	**Interviewer:** Ok., in the rancherias, how, how do you, for example, do you live all together?
JOS: nosotro(s), sí, nosotro(s) vivimo(s) #101 juntos sí, poco, poco familia de la mujer, poco familia de nosotro(s), yo y familia de la mujer y todo el padre, la madre de ello(s), y tíos de ello(s), y sobrino de ello(s), allá donde vive nosotro(s) #102, uno así, uno casería [sic] así, no, no má(s)) cerquita así con ese pueblo, así, así salteado.	**JOS:** We, yes, we live together, yes, some, some family members are from my wife's side, some of us, me and my wife's family everyone: the father, her mother, her aunts, her uncles, her nieces and nephews, we live in a small village, a little bit far from the town, scattered in the area.

7.6.2 RESEARCH QUESTIONS

Even though the MSIH (Prévost & White, 2000a) does not predict agreement mismatches, when they surface, they are systematic and temporary. Although the mismatches may be present at a certain point of the acquisition process due to morphology accessibility issues and problems related to mapping morphology and syntax, they disappear overtime. Prévost and White (2000a) argue for a default form and Ferdinand (1996) suggests 3rd person singular as an "elsewhere" form that replaces the correct one when an agreement mismatch occurs in French. Similarly, McCarthy

[42] We made the decision of only including tokens in the present tense to have a homogeneous examination. We left out tokens with *tú* [you] because of the problem of /s/ deletion, but this problem doesn't exist in the simple past. Having multiple tenses at the same time would have created a conflict in the analysis and we would have needed a separate study to best discriminate the results.

(2004) proposes the 3rd person singular form as the default form for errors of underspecification of features in Spanish.

If the agreement mismatches are temporary or a problem of overt morphology realization, this issue should disappear as the bilingual speakers become more proficient in their L2. When analysing the age of acquisition and proficiency as factors affecting agreement production, Foote (2010) found that the least proficient bilingual speakers had more trouble with agreement than the more proficient bilingual speakers.

In this regard, our goal is to determine whether there is any systematicity in the data regarding the agreement mismatches produced by the bilingual speakers and whether there is any effect of the speakers' proficiency on this issue. The use of the monolingual Spanish control group helps us assess if the mismatched sequences are a diagnostic of L2 acquisition or lower proficiency in Spanish as the native group is not expected to exhibit the behaviour of the L2 group. Therefore, we propose the following research questions:

1) Which is the subject personal pronoun that displays the highest percentage of subject-verb mismatches?
2) Does proficiency correlate with the production rate of mismatches?
3) Are there more mismatches with null or with overt subject pronouns?
4) How does the grammar of the bilingual speakers compare to the grammar of the monolingual Spanish group regarding subject-verb agreement?

7.6.3 HYPOTHESES

1) According to McCarthy (2004) *-mos* (1P), *-n* (2P/3P) and *-o* (1S) are the most highly specified vocabulary items, and in the competition for vocabulary insertion, these forms are

barred.[43] Therefore, the 3rd person singular will emerge as the preferred form in mismatched sequences because this is the least marked item. Consequently, the affixal forms that correspond to 1st, 2nd person formal plural, 3rd person plural, and 1st person singular will present more mismatches than 3rd person singular, which in turn is expected to be the default form and most of the mismatched sequences should have it as the most frequently supplied non-expected conjugation. We should therefore find forms such as: *yo come* [I eats]*, nosotros come* [we eats]*,* and *ellos come* [they eats].

2) Language proficiency is an important determinant in the successful production of a foreign language. Consequently, the advanced speakers will produce fewer mismatches than the intermediate speakers. We expect the advanced group to produce fewer mismatches, since they are supposed to have a more native-like knowledge of Spanish and hence a better knowledge of the Spanish conjugation system.

3) An overt subject is more transparent when it comes to inferring what the referent of the verbal external argument is. Consequently, more mismatches will be produced with null subject sentences since it may be more difficult to match a verbal affix with a pronoun that is not overtly expressed.

4) The bilingual grammar is expected to have mismatched sequences. However, this is not expected for the monolingual grammar. Consequently, we expect to find mismatched sentences in the bilingual group and no mismatched instances for the monolingual group.

[43] *-mos* (1P), *-n* (2P/3P) *-o* (1S) are specified for person/number, whereas 3rd person singular lacks features. Also, in terms of person/number, plural is more marked than singular and 1P is more marked than 2P and 3P. In turn, 2P is more marked than 3P.

7.7 RESULTS

In this section, we analyse the mismatches found in the data in order to establish whether there are specific patterns or regularization tendencies. We first present a general description and then a more in-depth analysis by segmenting the data by groups: intermediate vs. advanced speakers and sequences with null vs. overt subject pronouns. At the end of the section, we present the data for the monolingual group.

7.7.1 MISMATCHES IN THE BILINGUAL DATA

There are 431 mismatched sentences in the overall bilingual data with 18% of the tokens presenting non-target-like forms (cf. Table 7.3). The subject with the most problems was *ellos/ellas* [they] with 27% mismatches, followed by *yo* [I] with 21% and *nosotros* [we] with 18%.

These results were expected since the personal pronouns that have a conjugation with highly specified items will suffer more in case of underspecification of features. In the competition for vocabulary insertion if the correct form is not inserted, then the least marked form will emerge and the conjugations for *ellos/ellas* [they], *yo* [I] and *nosotros* [we] are the expected candidates to present more problems.

Table 7.3. Mismatched sequences produced by the bilingual participants

Pronoun	Total N of tokens	Total N of mismatches[1]	%
ellos/ellas	625	171	27%
Yo	861	185	21%
Nosotros	307	55	18%
Ustedes	13	2	15%
él/ella	475	17	4%
Usted	55	1	2%
Totals	**2336**	**431**	**18%**

[1]mismatches = produced person verbal affixes that correspond to a different person

As expected, rather than the production of a more complex form for the mismatched sequences, this contact language seems to display a simplification of the Spanish set of personal

affix markers. This simplification is clear as the 3rd person singular marker stands out as the default conjugation for most of the mismatched sentences independently of the subject personal pronoun (Table 7.4). 3rd person singular is the preferred supplied form for the mismatched sequences with 91% of the reported occurrences. This means that out of 10 non-target-like sentences, at least 9 mismatched sentences were produced with a 3rd person singular form.

We must acknowledge that a subset of the mismatched sequences may be just a replication of the native group elision of /n/, as we found in the systematic study of this phenomenon in the monolingual control group. Nevertheless, the low percentage (0.5%) of elision of /n/ in the monolingual group and the high percentage (91%) of mismatched sequences with 3rd person singular in the bilingual group support the argument that the bilingual group is still using a 3rd person singular as default form due to underspecification of features.

Table 7.4. N and % of verbal affixes found in mismatched sequences

Verbal affixes supplied as incorrect conjugations	N	%
3S/2S (formal usted)	390	91%
3P/2P (formal ustedes)	31	7%
1S	5	1%
1P	5	1%
Totals	**431**	**100%**

Although in this study there is a rather consistent use of 3rd person singular as a substitute for the other person/number markers, it is also true that a few other mismatches not involving a 3rd person singular surface in the data as well. The other forms for the mismatched sentences were 3rd person plural, 1st person singular, and 1st person plural. However, the production of these forms was very low and accounted for less than 10 % of the incorrect forms produced.There were found 41 tokens (9%) of the non-standard occurrences reported with a form other than a 3rd person singular, such as those combinations found in (87) to (91).

(87) Ella dan muchas fotocopias. (GEI)
PRO.**3S** gave.**3P** many photocopies
'She gave out many handouts'.

(88) Nosotros no son evangélicos. (ISR)
PRO.**1P** NEG are.**3P** evangelical Christians
'We aren't evangelical Christians'.

(89) (Él) viven con otra mujer (ANE)
(PRO.**3S**) live.**3P** with another woman
'He lives with another woman'.

(90) ¿Usted no conozco? (CRI)
PRO.**2S** NEG know.**1S**
(formal)
'Don't you know?'

(91) Yo no hablamos español. (JOS)
PRO.**1S** NEG speak.**1P** Spanish
'I don't speak Spanish'.

In (87), (88), and (89), the supplied affix form corresponded to a 3rd person plural/2nd person plural formal (*da-n*, *vive-n* and *son*), however, the pronouns *él/ella* [he/she] require a 3rd person singular and *nosotros* [we] requires a 1st person plural.[44] The number of tokens in which the supplied incorrect form was this agreement marker accounts for 7% of the overall number of non-target-like sequences, making this form the second most frequently found, yet still very far from the identified default form.

In (90), the produced conjugation corresponds to a 1st person singular, which clashes with the subject pronoun *usted* [you] that needs the verb in 2nd person singular (formal). There are only

[44] 3rd person singular was the default form for the pronouns other than *él/ella*. What is then the preferred form for *él/ella* in problematic sentences? We found that most of the mismatched sequences (13 out of 17) with these pronouns were realized with a 3rd person plural/2nd person plural formal. In this case, the speakers are also using a 3rd person form as a replacement in a mismatched sequence, but in plural. Although plural is more marked than singular, what is important in this case is that the speakers still get *person* right.

5 instances of this type of mismatches and just 1% of the overall mismatched tokens contain a 1st person singular as the unexpected form.

In (91) the supplied affix should have been 1st person singular (habl-o), since it has to agree with the pronoun *yo* [I], yet the pronoun received a 1st person plural (habla-mos). Instances in which a 1st person plural is the supplied incorrect form ascends to 5 tokens or 1% of the overall mismatches.

7.7.2 Analysis of the data by Spanish proficiency

In Table 7.5, we present the mismatches produced by each participant and according to their Spanish proficiency. The intermediate group produced 332 tokens (66%) of mismatched sequences, whereas the advanced group produced 99 mismatched sentences (34%).

Table 7.5. Mismatched sequences produced by each participant and according to Spanish proficiency[45]

Informant	Spanish Level	Total N of tokens	Total N of mismatches	%
EST	I	35	18	51%
BEN	I	93	13	14%
ISR	I	55	13	24%
JOS	I	134	49	37%
NAR	I	56	14	25%
CRI	I	35	17	49%
ROS	I	88	15	17%
MAI	I	99	30	30%
ANE	I	148	75	51%
MAR	I	59	31	53%
NER	I	50	10	20%
ANI	I	86	44	51%
MIS	I	55	3	5%
TOTAL		**993**	**332**	**33%**
ZUR	A	66	2	3%
CAT	A	133	3	2%
LUZ	A	129	12	9%
ELI	A	70	8	11%
ALB	A	87	6	7%
RAM	A	38	15	39%
SAR	A	129	2	4%
ORL	A	69	5	2%
GEI	A	288	7	2%
MIG	A	95	9	9%
ABR	A	41	5	12%
FAB	A	50	4	8%
CON	A	97	7	7%
ADE	A	51	14	27%
TOTAL		**1343**	**99**	**7%**

[45] There is some variation in the intermediate group data. The participants who produced more mismatches were also those who were illiterate and learned Spanish later in their lives. The interaction tells us that literacy and onset of learning of Spanish could also be some other variables to study in future research. They could help us understand what extralinguistic factors are at play regarding the issue of subject-verb agreement.

When the subject-verb mismatches are analysed by Spanish proficiency, 33% of the data produced by the intermediate group contains a non-target-form (Table 7.6). While for the advanced group, 7% of the data contains mismatched sequences (Table 7.7).

Table 7.6. Production of subject-verb mismatches by the intermediate speakers

Pronoun	Total N of tokens	Total N of mismatches	%
ellos/ellas	226	105	46%
Yo	408	164	40%
Nosotros	128	47	37%
él/ella	203	15	7%
Usted	24	1	4%
Ustedes	4	0	0%
Total	**993**	**332**	**33%**

In both cases, the intermediate and the advanced group selected the same personal pronouns that require a highly specified form as the pronouns with more problems to realize the correct subject-verb agreement. The intermediate group selected *ellos/ellas* [he/she] with 46% and the advanced group *ustedes* [you] with 22%, which share the same morphological ending *-n*.

Table 7.7. Production of subject-verb mismatches by the advanced speakers

Pronoun	Total N of tokens	Total N of mismatches	%
Ustedes	9	2	22%
ellos/ellas	399	66	17%
Yo	453	21	5%
Nosotros	179	8	4%
él/ella	272	2	1%
Usted	31	0	0%
Totales	**1343**	**99**	**7%**

An analysis of the number of mismatches segmented by Spanish proficiency shows the intermediate group produced more mismatches (N= 332) than the advanced group (N=99), and this difference is significant X^2 (2, N = 2336) = 257.725, p > .05. The difference in the numbers and percentage of mismatched sequences produced by the intermediate and the advanced Spanish groups may be explained in terms of a less automatic internalization of the Spanish language by

the least proficient group, which is less able to confront the processing demands that can cause problems in the computation of subject-verb agreement (Nicol & Greth, 2003).

7.7.3 Analysis of the data by the status of the subject pronoun: null vs. overt

When the data of the participants is globally analyzed we found out that 16% of the overt subject pronouns presented mismatches. In the case of the null subjects pronouns there was 19% of instances with problematic agreement (cf. Table 7.8).

Table 7.8. Mismatches by subject expression in the overall data

Pronoun status	Overall # of tokens	Mismatches	%
Overt	796	131	16%
Null	1540	300	19%
Total	2336	431	18%

The specific analysis of the mismatched data tells us that most of the mismatches (300 or 70%) were produced with a null subject while 131 tokens or 30% of the mismatches were realized with an overt pronoun, as shown in Table 7.9.

Table 7.9. Mismatches by subject expression in the overall mismtached data

Null	**Overt**	**Totals**
N= 300 (70%)	N= 131 (30%)	N= 431

When the production of mismatches is analysed in relation to proficiency in Spanish (Table 7.10), it can be observed that the intermediate and advanced groups produced more mismatched sequences with a null subject pronoun than with an overt pronominal subject. This result confirms the hypothesis that more mismatched sentences would be produced with null subjects because of the difficulty to match a verb with an absent subject.

The fact that the Wayuunaiki speakers have trouble producing more of the expected sequences with null subjects shows they have a problem at the syntactic level, in which verbal agreement takes place according to the Generative Grammar. Within this framework, null subjects

in Spanish are theoretically explained from two perspectives. According to Alexiadou and Anagnostopoulou (1998), null subjects are licensed in Spanish because inflection is [+affixal] and morphologically rich, and the Spanish person verbal affixes are part of the numeration and with pronoun-like features. In Spanish verbal agreement, affixes have a categorical [+D] feature which is merged with the verb and have semantic content, making the subject redundant. For Holmberg (2005) and Sheehan (2006), overt/null subjects are possible in Spanish because zero pronouns are related to the phonological form (PF), and because of how the Extended Projection Principle (EPP) checking is realized. In Spanish, EPP checking follows three steps: (i) null subjects rise to Spec TP, which in turn can be realized or not; (ii) the verb and morphological agreement is moved to T; and (iii) the uninterpretable features in T and the interpretable features in V are valued (Liceras & Fernández, 2016, p. 7).

Table 7.10. Spanish proficiency and mismatches by null and overt pronouns

Spanish proficiency	Total N of tokens	Total N of mismatches	Mismatches with null subjects	Mismatches with overt subjects
Intermediate N= 13	N= 993 (42%)	N= 332 (33%)	N= 225 (68%)	N= 107 (32%)
Advanced N= 14	N= 1343 (58%)	N= 99 (11%)	N= 73 (72%)	N= 26 (26%)

However, the Guajiro Spanish speakers may not have a grammar where the subject affix markers bear the feature [+D] as it is the case with native or monolingual Spanish group. In this sense, this variety of Spanish is not a classic Null Subject Language, but a partial Null Subject Language in which subjects are identified with overt pronouns or the context, this may explain why the Guajiro Spanish speakers produced a higher number of subject verb mistmaches with null subjects than with overt prononominal subjects. The fact that Wayuunaiki is also a null partial language could support the argument that Guajiro Spanish is not a classic Null Subject Language like Spanish or Italian.

7.7.4 THE MONOLINGUAL GROUP

Regarding subject-verb agreement in the monolingual group, the monolingual Spanish speakers produced two tokens in which /n/ was omitted. These sentences were produced with *ellos/ellas* [they] and represents just 0.5 % of the tokens of the monolingual data. This result rules out the influence of a vernacular process in the bilingual data for this particular person/number, since the percentage of mismatched sequences in the bilingual data was considerably higher (91%) (cf. Table 7.4).

Table 7.11. Non-expected sequences produced by the monolingual participants

Pronoun	Total N of tokens	Total N of mismatches	%
ellos/ellas	143	2	0.5%
yo	145	0	0%
nosotros	39	0	0%
ustedes	2	0	0%
él/ella	42	0	0%
usted	20	0	0%
Total	**391**	**2**	**0.5%**

7.8 DISCUSSION

In this chapter, we have tried to provide an account of the subject-verb mismatches produced by the Guajiro Spanish speakers and to compare their results to a reference group of monolingual Spanish participants. With respect to the four research questions that were formulated, the results are the following.

The first question intended to elucidate which was the subject with the highest percentage of mismatches. For the bilingual group, we found that *ellos/ellas* [they] presented the highest percentage of non-target-like sentences (27 %), followed by *yo* [I] with 21% and *nosotros* [we] with 18%. When the bilingual group is segmented by proficiency, the same pattern is found for the intermediate group. The advanced group, however, selected *ustedes* [you] as the pronoun with

the highest percentage of mismatches. Nonetheless, this result is not problematic since these *ellos/ellas/ustedes* share the same verb ending *-n*.

The results for the bilingual group were as expected, as the most highly specified items — *-mos* (1P), *-n* (2P/3P), and -o (1S) — presented more mismatches than 3rd person singular. *Ellos/ellas* [they] had the highest production rate of mismatched sentences, followed by *yo* [I] and *nosotros* [we]. A particular order was not predicted in this specific analysis, however, the conjugation belonging to *nosotros* [we] did not come before *yo* [I]. We would have expected *nosotros* to be in the second place because this pronoun also requires a highly specified conjugation: plural is more marked than singular. However, *yo* [I] emerged as the second most problematic pronominal subject. This could be because *nosotros* [we] uses a conjugation with a highly salient morphological ending (Orozco & Guy, 2008), which is more noticeable for the L2 learners and thus easier to remember and learn. According to Blyth (2005, p. 217), a form that is perceptually salient is easier to be noticed in the input by the learners and thus a better candidate to be retained and learned. For example, *-mos* is easier to perceive than just *-o* or *-n*.

This research question also uncovers that the most frequently supplied non-expected conjugation was 3rd person singular. This form was produced in 91% of the cases of mismatched sequences independently of the pronominal subject. McCarthy (2004) proposes that 3rd person singular is in fact the expected form in subject-verb mismatches in Spanish. She argues that 3rd person singular is an unmarked form and consequently has underspecified features. Therefore, if the correct form does not surface during lexical insertion, the default form takes over, which is predicted to be 3rd person singular as per the MUH of McCarthy.

The second research question was linked to the speaker's proficiency in Spanish. The data showed that the advanced group produced 99 tokens or 7% of mismatched sequences while the

intermediate group produced 332 tokens or 33% of mismatched sequences from their respective totals. These results give support to the idea that language proficiency is a key factor that helps the speakers overcoming the difficulties posed by the overt morphology of Spanish verbs. There is a relationship between degree of proficiency and number of mismatches such as the higher the proficiency the less deviance from the subject-verb agreement requirements of Spanish, which results in fewer subject-verb mismatches. This result also supports the argument of the MSIH, which argues for a temporary stage of the mismatched forms, as the problematic sequences tend to disappear or are less frequent with a higher proficiency of Spanish.

Also, in terms of proficiency in Spanish, the results uncover a different sensitivity to the problem under investigation: the shallower processing resources of the intermediate group may lead to retrieval problems and therefore to the production of more non-matching agreement markers by these speakers. The Shallow Structure Hypothesis (SSH) proposes that L2 speakers are less sensitive to structural requirements because their syntactic representation is less comprehensive than that of native speakers (Clahsen & Felser, 2006a, 2006b, 2006c, 2018).

The third question was intended to uncover the influence the lexical realization of the pronoun had on the mismatches. 300 (70%) of the mismatches occurred with a null subject and 131 (30%) were realized with an overt pronoun. This supports the argument that when the speakers can correlate the conjugation with an overt pronoun, it is easier for them to produce the required morphology or to copy the subject pronoun features onto the conjugation. The tendency found in the overall data was also confirmed in the bilingual by-group analysis, in which the advanced and intermediate group produced more mismatched sequences with null subjects. This could be accounted for by the fact that non-phonetically realized features are more problematic when it comes to the realization of agreement and checking operations. In contact language situations and

in spontaneous speech, subject-verb agreement becomes vulnerable and in null subject languages like Spanish, the subject verb-agreement is not necessarily enforced if the subject is not phonetically pronounced (Grinstead, De la Mora, Pratt & Flores, 2009).

The last question deals with the comparison of the bilingual and monolingual grammar regarding the subject-verb agreement production. In monolingual Spanish data, two tokens in which the /n/ is dropped were found. As argued before, these tokens represent a feature of the vernacular Spanish used as a benchmark for this study, and not a matter of mismatched sequences. In Colombian Caribbean Spanish, /n/ is velarized, a necessary step before the sound is completely dropped. The presence of just two tokens shows, however, that the process may be still in development or not entirely widespread. More quantitative and phonetic studies about this specific issue are necessary to uncover how extensive and distributed the phenomenon is.

Moreover, the monolingual data does not present variability across all personal subject pronouns nor with all the informants. In the bilingual data, all informants produced mismatched sentences, every pronominal subject exhibited non-vernacular sequences, and the supplied incorrect form could be any of the conjugations (although there is a strong preference for 3rd person singular). In the monolingual Spanish data, the omission of the verbal affix was only found with *ellos/ellas* [they] and the only possible supplied form is a 3rd person singular.

Based on these results, it can be said that the monolingual and bilingual grammars are behaving differently regarding the topic under analysis.

7.9 SUMMARY

The main objective of this chapter was to provide an analysis of subject-verb mismatches found in Guajiro Spanish, an issue that had been briefly described in the literature, but for which more research was needed. In this sense, this study sheds new light on this problem, specifically for this variety of Spanish. The results drawn from the data allow us to offer a new account of this characteristic of Guajiro Spanish by uncovering specific patterns suggesting these speakers tend to use a default form, the 3rd person singular, which bears no affix marker. This default marker is used with all the subject pronouns.

Intermediate proficiency in Spanish leads to the production of more subject-verb mismatches. This result suggests that subject-verb agreement is a computation that is particularly difficult for speakers of an L2 (Hoshino *et al*., 2010), albeit this difficulty decreases as language proficiency increases. Most of the mismatches were realized with a null subject pronoun, which gives support to the idea that matching features to an absent pronoun is more difficult than matching features with an overt subject pronoun.

Finally, we propose that the bilingual and monolingual grammars are distinct. The monolingual group presented sentences in which /n/ was omitted with *ellos/ellas* [they], and we argued that this is due to natural vernacular process typical of this variety of Spanish spoken close to the Caribbean. The bilingual group, however, exhibited mismatched sequences with all subject pronouns and the supplied non-expected conjugations were not only in 3rd person singular but also in 3rd person plural/2nd person plural formal, 1st person singular and 1st person plural. In the case of the bilingual group, the problem stems from problematic acquisition of the subject-verb feature checking process required to produce the correct overt morphology of the Spanish conjugations.

Furthermore, if Guajiro Spanish subject affix markers don't bear the feature [+D] proposed by Alexiadou and Anagnostopoulou (1998), this variety uses other ways to identify the subjects: context or the overt realization of the subject itself.

8. CONCLUSIONS

The aim of this dissertation was to find out what linguistic outcomes have been produced by the protracted contact between two very dissimilar languages, Wayuunaiki and Spanish. In this sense, two grammatical issues were analysed. The first grammatical topic was the distribution of null and overt subject pronouns, an analysis carried out using the Null Subject Parameter and the variable overt expression of the pronominal subject using the variationist theory. The second topic was the subject-verb agreement issue, which was analyzed with the theory of features.

The results found for the two issues under study support the argument that Wayuunaiki has an effect on some portions of the Guajiro Spanish grammar, particularly when the stative verbs motivate a higher realization of overt pronominal subjects, as we learned in chapter 5. In the variationist analysis in chapter 6, the results support the argument that the bilingual and monolingual underlying grammars did not differ. Although the groups presented different rate of overt SPPs usage, the factor groups did not differ greatly, actually the comparison groups' structures were parallel, and the differences were minimal, the Guajiro Spanish speakers's underlying structure in this chapter was very similar to the native Spanish group. In the analysis of the subject-verb agreement in chapter 7, it was found that the bilingual and monolingual grammars were different regarding this topic. The bilingual speakers, when having issues to retrieve the appropriate conjugation, resort to a default conjugation, to wit, the 3rd person singular. However, this phenomenon was not found in the monolingual group. Finally, we would like to underline that the Spanish conjugations are a learnability issue that may be challenging for Wayuunaiki-Spanish teachers (Liceras, Méndez & Moreno, 2017), given the fact that if these bilingual teachers are not dominant in the Spanish standard variety, they could possibly exposed their students to the subject-verb mismtaches that are present in their Guajiro-Spanish variety.

On another note, we want to underline the contributions of this dissertation to the field of linguistics, and especially to the study of Spanish in contact with an indigenous language of Colombia and Venezuela. In the case of Guajiro Spanish, this dissertation is one of the few available that specifically investigate the Spanish produced by these speakers. The studies of the distribution of null and overt subjects and the overt expression of the pronominal subject presented here are the first ones, as far as we know, carried out for this variety of Spanish and with the use of the Null Subject Parameter of the Generative Grammar and the quantitative methodology of the variationist sociolinguistics. Regarding, the study of the subject-verb mistmaches, our research is a new contribution to the study of this problem, which was not previously researched from the theory of features within the Generative Grammar. Although the topic had interested some researchers, the explanation provided with the use of the theory of features will advance the understanding of this issue found in Guajiro Spanish.

During the writing of this dissertation we also found out that there are still a lot of opportunities to study from different angles the topics we covered here. For example, the study of overt SPP expression by verb type (stative versus active) can be complemented by asking the participants to translate from Wayuunaiki to Spanish specific sentences containing stative and active verbs. This could help to uncover patterns of production or omission of the pronominal subjects with each type of verb. Besides, more research needs to be done to confirm if /s/ is deleted in second person *tú* in Guajiro Spanish because is a vernacular process or a language acquisition problem. For a future variationist study on overt SPP expression and for the thesis in general, the Spanish monolingual cohort needs to have a bigger number of participants. This situation did not allow us, for instance, to produce the analysis of social variables on the quantitative study, and to carry out the study of the subject-verb agreement issue with the use of variation theory.

Finally, we would like to mention that the study of Spanish in contact with Wayuunaiki is incipient and more studies need to be done. Also, having a better knowledge of Wayuunaiki is an advantage for future studies as there is limited information about this language.

9. REFERENCES

Abreu, L. (2012). Subject Pronoun Expression and Priming Effects among Bilingual Speakers of Puerto Rican Spanish. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, ed. Kimberly Geeslin and Manuel Díaz-Campos, 1-8. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2651.

Abreu, L. (2009). Spanish subject personal pronoun use by monolinguals, bilinguals and second language learners. Gainesville, FL: University of Florida dissertation.

Adger, D. (2003). *Core syntax: A minimalist approach*. Oxford: Oxford University Press.

Adger, D. (2006). Combinatorial variability. *Journal of Linguistics*, 42(3), 503-530. https://doi.org/10.1017/S002222670600418X

Adger, D., & Smith, J. (2005). Variation and the minimalist program. In L. M. E. A. Cornips, L. & K. P. Corrigan (Eds.), *Current Issues in Linguistic Theory* (Vol. 265, pp. 149-178). https://doi.org/10.1075/cilt.265.10adg

Adger, D., & Trousdale, G. (2007). Variation in English syntax: theoretical implications. *English Language and Linguistics*, 11(2), 261-278. https://doi.org/10.1017/S1360674307002250

Aikhenvald, A. (1999) The Arawak language family (chapter 3), in *The Amazonian Languages*, ed. Dixon & Aikhenvald. Cambridge: Univesity Press.

Alexiadou, A., & Anagnostopoulou, E. (1998). Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16. 491–539

Aleza, M. (2010). Morfología y sintaxis. Observaciones gramaticales de interés en el español de América. Parte segunda: fenómenos gramaticales. En M. Aleza & J. Enguita (coords.) (2010). *La lengua española en América: normas y usos actuales*. Valencia: Universidad de Valencia.

Álvarez, P. (2000). *Aspectos lingüísticos de la escritura del español por estudiantes guajiros* (Unpublished master´s dissertation). Universidad del Zulia, Venezuela.

Álvarez, J. (2002). Definitud y foco en las conjugaciones subjetiva y objetiva del Guajiro. *Revista Latinoamericana de Estudios Etnolingüísticos*, 10, 24-42.

Álvarez, J. (2004). Raices y sufijos temáticos en la morfología verbal del guajiro/Wayuunaiki. *Antropólogica*, 102:27-98.

Álvarez, J. (2004). Raíces y sufijos temáticos en la morfología verbal del guajiro/Wayuunaiki. *Antropólogica* 102:27-98.

Álvarez, J. (2005). Comparative constructions in Guajiro/Wayuunaiki. *Opción*, 21 (47), 9-36.

Álvarez, J. (2007). Vocalic Mora Augmentation in the Morphology of Guajiro/Wayuunaiki. *Lingua Americana*, 20, 119-142.

Álvarez, J. (2017). *Manual de la lengua wayuu. Karalouta atüjaaya saa´u wayuunaikikuwa'ipa*. Organización Indígena de La Guajira YANAMA.

Ángel Rodríguez, D. M. (2014). El desplazamiento lingüístico: Corolario de la transculturación en la Baja Guajira. Interacción, 12, 169-189. https://doi.org/10.18041/1657-7531/interaccion.0.2324

Arellano, F. (1987). *Una introducción a la Venezuela prehispánica: culturas de las naciones indígenas venezolanas*. Caracas: Universidad Católica Andrés Bello.

Atencio, Atencio, M. (2014). *Rasgos distintivos del español en contacto con el wayuunaiki en la parroquia Andrés Bello del municipio La Cañada de Urdaneta del estado Zulia, Venezuela* (Unpublished master´s dissertation). University of Amsterdam, Amsterdam.

Baker, M. C. (2008). The macroparameter in a microparametric world. In T. Biberauer (Ed.), *Linguistik Aktuell/Linguistics Today* (Vol. 132, pp. 351-373). https://doi.org/10.1075/la.132.16bak

Barbiers, S. (2014). Syntactic doubling and deletion as a source of variation. In M. C. Picallo (Ed.), *Linguistic Variation in the Minimalist Framework* (pp. 197-223). https://doi.org/10.1093/acprof:oso/9780198702894.003.0010

Barrera Monroy, E. (1990). La rebelión Guajira de 1769: algunas constantes de la Cultura Wayuu y razones de supervivencia. *Credencial Historia No. 6*. Recuperado de http://www.banrepcultural.org/biblioteca-virtual/credencial-historia/numero-6/la-rebelion-guajira-de-1769

Bayley, R. (2013). The quantitative paradigm. In J. K. Chambers & N. Schilling (Eds.), *The handbook of language variation and change*, 85-107. Malden, MA: Wiley-Blackwell.

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

Blyth, C. (2005). From Empirical findings to the teaching of aspectual distinctions. In *Tense and Aspect in Romance Languages: Theoretical and applied perspectives*, ed. D. Ayoun and R. Salaberry, 211-52. Amsterdam: John Benjamins.

Borer, H. (1984). *Parametric syntax, case studies in semitic and romance languages*. https://doi.org/10.1515/9783110808506

Bravo, M. (2005). El Orden de los Constituyentes en wayuunaiki/guajiro, Msc dissertation. La Universidad del Zulia, Marcaibo, Venezuela.

Bullock, B., & Toribio, A.J. (2009). Reconsidering Dominican Spanish: Data from the rural Cibao. *Revista Internacional de Linguistica Iberoamericana* 14, pp. 49- 73.

Bustos, Y., & Ipuana, M. (2015). Waashajaai Wanüiki (Hablemos nuestra lengua). Yasir Bustos: Colombia.

Camacho, J. (2016). The Null Subject Parameter revisited. The evolution from null subject Spanish and Portuguese to Dominican Spanish and Brazilian Portuguese. In *the Morphosyntax of Portuguese and Spanish in Latin America*, by Mary A. Kato and Francisco Ordóñez (Eds.). New York: Oxford.

Camacho, J. (2013). *Null subjects*. Cambridge: Cambridge University Press.

Cameron, R., & Flores-Ferrán, N. (2004). Preservation of subject expression across regional dialects of Spanish. *Spanish in Context*. 1:1, 41-65.

Cameron, R. (1993). Ambiguous agreement, functional compensation, and non-specific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5:305–334.

Cameron, R. (1995). The scope and limits of switch reference as a constraint on pronominal subject expression. *Hispanic Linguistics* 6/7, 1-27.

Candelier, H. (1994). *Riohacha y los indios guajiros* (Ecoe Ediciones). Universidad de Texas.

Canfield, D. L. (1981). *Spanish Pronunciation in the Americas*. Chicago: University of Chicago Press.

Carminati, M. N. (2002). *The Processing of Italian Subject Pronouns*, Ph.D. dissertation, University of Massachusetts at Amherst, Amherst (MA), GLSA publications.

Carnie, A. (2012). *Syntax: A Generative Introduction*, third edition. Wiley-Blackwell.

Carvalho, A. M., & Child, M. (2011). *Subject pronoun expression in a variety of Spanish in contact with Portuguese* (J. Michnowicz & R. Dodsworth, Eds.), Recuperado de www.lingref.com/cpp/wss/5/abstract2502.html

Carvalho, A. M., Orozco, R., & Shin, N. L. (2015). *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Georgetown University.

Carvalho, A. M., & Bessett, R. M. (2015). Subject Pronoun Expression among Spanish-Portuguese Bilinguals. In *Subject Pronoun Expression in Spanish: A Cross-dialectal perspective*, edited by Carvalho, A. M.; Orozco, R. & Lapidus, N. Washington, 143-165. D.C.: Georgetown University Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), Readings in English transformational grammar (pp. 184–221). Waltham, MA: Ginn-Blaisdell.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Holland ; Cinnaminson, [N.J.]: Foris Publications

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from building 20. Essays in linguistics in honor of Sylvain Bromberger* (pp. 1–52). Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist program*. Cambridge, Mass: The MIT Press

Chomsky, N. (2001). Derivation by phase. In J. Kenstowicz (Ed.), Ken Hale: *A life in language* (pp. 1–52). Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries. R. Martin, D. Michaels & J. Uriagereka, (Eds.), Step by step: *Essays in minimalist syntax in honor of Howard Lasnik*, 89-156. Cambridge, Mass.: MIT Press.

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.

Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics* 27, 3-42. doi:10.1017.S0142716406060024

Clahsen, H., & Felser, C. (2006b). Continuity and shallow structures in language processing. *Applied Psycholinguistics* 27, 107-126. doi:10.1017.S0142716406060206

Clahsen, H., & Felser, C. (2006c). How native-like is non-native language processing? *Trends in Cognitive Sciences* 10 (12), 564-570. doi:10.1016/j.tics.2006.10.002

Clahsen, H., & Felser, C. (2018). Some notes on the Shallow Structure Hypothesis. *Studies in Second Language Acquisition*, 40(3), 693-706. doi:10.1017/S0272263117000250

Constitución de la República Bolivariana de Venezuela. (1999, diciembre 30). *Gaceta oficial de la República Bolivariana de Venezuela, 36860*.

Constitución política colombiana (1991). Asamblea Nacional Constituyente, Bogotá, Colombia, 6 de Julio de 1991.

Corbett, G. G. (2012). *Features.* Cambridge: Cambridge University Press.

Cornips, L., & Corrigan, K. (2005). Toward an integrated approach to syntactic variation:A retrospective and prospective synopsis. In L. M. E. A. Cornips, L. & K. P. Corrigan (Eds.), *Current Issues in Linguistic Theory* (Vol. 265, pp. 1-30). https://doi.org/10.1075/cilt.265.10adg

D'Alessandro, R. (2015). Null Subjects. In: A. Fábregas, J. Mateu & M. Putnam (Eds.), *Contemporary Linguistic Parameters* (201-226). London: Bloomsbury Press.

De La Pedraja, R. (1981). La Guajira en el Siglo XIX: Indígenas, contrabando y carbón. *Revista Desarrollo y Sociedad*, (6), 329-359. https://doi.org/10.13043/dys.6.5

De Prada Pérez, A. (2011). *Subject expression in Minorcan Spanish: Consequences of contact with Catalan*. UMI dissertation Publishing.

Declaración de la Nación Wayuu desde el territorio ancestral. (2009, noviembre 1). Recuperado 26 de octubre de 2018, de Viajeros4x4x4 website: https://viajeros4x4x4.com/2009/11/01/declaracion-de-la-nacion-wayuu-desde-el-territorio-ancestral-colombia/

Departamento Administrativo Nacional de Estadística. (2018). *Censo nacional de poblacion y vivienda*. Bogotá, Colombia.

Domínguez, C. L., & Mora, E. (1998). *El habla de Mérida*. Mérida: Universidad de Los Andes.

Erker D. & Guy G. (2012). The role of lexical frequency in syntactic variability: Variable subject pronoun expression in Spanish. *Language* 88: pp.526-557.

Embick, D. (2008). Variation and morphosyntactic theory: competition fractionated. *Language and Linguistics Compass*, *2*(1), 59-78. https://doi.org/10.1111/j.1749-818X.2007.00038.x

Embick, D. (2008) 'Variation and morphosyntactic theory: Competition fractionated' In *Language and Linguistics Compass* 2/1.

Etxebarria Arostegui, M. (2012). Bilingüismo y realidad sociolingüística de la lengua del grupo Wayuu en el Caribe Colombiano. *Anuario del Seminario de Filología Vasca Julio de Urquijo: International Journal of Basque Linguistics and Philology*, 46 (2) pp. 271-293

Eubank, L. (1993/94). On the transfer of parametric values in L2development. *Language Acquisition* 3, 183–208.

Ferdinand, A. (1996). *The development of functional categories: The acquisition of the subject in French*. The Hague: Holland Academic Graphics.

Ferguson, C. A. (1971). *Language structure and language use: essays*. Stanford, Calif: Stanford University Press.

Filiaci, F. (2010). Null and overt subject biases in Spanish and Italian: A cross-linguistic comparison. In C. Borgonovo, M. Espanol-Echevarria, & P. Prevost (Eds.), *Selected Procedings of the 12th Hispanic Linguistics Symposium* (pp. 171–182). Cascadilla Press.

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* 16:49–73.

Flores-Ferrán, N. (2005). La expresión del pronombre personal sujeto en narrativas orales de puertorriqueños de Nueva York", en L. Ortiz López & M. Lacorte (Eds.), *Contactos y contexto lingüísticos: el español en los Estados Unidos y en contacto con otras lenguas*. Frankfurt am Main/Madrid: Vervuert/Iberoamericana. Pp. 119-129.

Flores-Ferrán, N. (2007). Los Mexicanos in New Jersey: Pronominal expression and ethnolinguistic aspects. *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*, (Eds.), por J. Holmquist, A. Lorenzino, & L. Sayahi, 85-91. Somerville, MA: Cascadilla Proceedings Project.

Flórez, L. (1963). El español hablado en Colombia y su atlas linguístico. *Thesaurus*, 18, 268-356.

Foote, R. (2010). Age of acquisition and proficiency as factors in language production: Agreement in bilinguals. *Bilingualism: Language and Cognition* 13 (2), 99-118.

García, E., López, B., & Agudo, X. (1996*). La cosmovisión Wayuú a través de los textiles*. (Fundación Carlos Eduardo Frías). Caracas.

Gobernación de la Guajira. (2012). *Plan de Desarrollo Departamento de la Guajira 2012-2015 "La Guajira Primero"*. Riohacha.

González R., O. (2006). *Uso y tratamiento del wayuunaiki en dos instituciones etnoeducativas wayuu* (Master´s thesis). Universidad Mayor de San Simón, Cochabamba, Bolivia.

González, O. (1973). *Los guajiros: una cultura indo-hispana*. Universidad Central de Venezuela, Caracas.

Grinstead, J., De la Mora, J., Pratt, A., & Flores, B. (2009). Temporal interface delay and root nonfinite verbs in Spanish-speaking children with specific language impairment. In J. Grinstead (Ed.), *Hispanic child languages: Typical and impaired development*, pp. 239–263. Philadelphia: John Benjamins.

Halle, Morris & Alec Marantz. 1993. 'Distributed Morphology and the Pieces of Inflection.' In *The View from Building 20*, ed. Kenneth Hale and S. Jay Keyser. MIT Press, Cambridge, 111-176.

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, *22*(1), 27–62.

Hazen, K. (2011). Labov: language variation and change. In P. Kerswill, B. Johnstone, & R. Wodak (Eds.), *The SAGE Handbook of Sociolinguistics* (pp. 24-39). https://doi.org/10.4135/9781446200957

Henry, A. (1995). *Belfast english and standard english: dialect variation and parameter setting*. Oxford University Press.

Henry, A. (2003). "Variation and Syntactic Theory". In *The Handbook of Language Variation and Change*. Edited by J. K. Chambers, P. Trudgill & N. Schilling-Estes.

Holmberg, Anders. (2005). Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36. 533–564.

Holmquist, J. (2012). Frequency rates and constraints on subject personal pronoun expression: Findings from the Puerto Rican highlands. *Language Variation and Change*, 24(2), 203-220. https://doi.org/10.1017/S0954394512000117

Hoshino, N., Dussias, P. E., & Kroll, J. K. (2010). Processing subject-verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition*, 13, 87 -98.

Hudson, R. A. (1996). *Sociolinguistics*. Cambridge: Cambridge University Press.

Hurtado, L. M. (2005). Condicionamientos sintáctico-semánticos de la expresión del sujeto en el español colombiano. *Hispania*, 88(2), 335-348.

Instituto Nacional de Estadística. (2011). *XIV Censo Nacional de Población y Vivienda*. Caracas, Venezuela.

Jacobson, P. F. (2012). The effects of language impairment on the use of direct object pronouns and verb inflections in heritage Spanish speakers: A look at attrition, incomplete acquisition and maintenance. *Bilingualism: Language and Cognition*, 15(1), 22–38.

Jonhson, M. & Barnes, S. (2013). Haya vs. Haiga: An Analysis of the Variation Observed in Mexican Spanish Using a Mixed Effects. *Selected Proceedings of the 6th Workshop on Spanish Sociolinguistics*, (Ed.), A. M. Carvalho & S. Beaudrie, 32-40. Somerville, MA: Cascadilla Proceedings Project.

Kiesling, S. F. (2011). *Linguistic variation and change*. Recuperado de http://site.ebrary.com/id/10477150

Knabenschuh de Porta, S. (1998). Hacia una sistematización semántica de la conjugación en Guajiro: tempus y aspecto. *Paramillo* No. 17: pp. 387-448.

Kroch, A. (1994). Morphosyntactic Variation. K. Beals et al., Eds. *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*.

Kroch, A. (1989). Function and grammar in the history of English: Periphrastic do, in *Language Change and Variation* Edited by R. W. Fasold & D. Schiffrin, pp. 133-172.

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, 45(4), 715-762. https://doi.org/10.2307/412333

Labov, W. (1972). Some principles of linguistic methodology. *Language in Society*, 1(1), 97-120.

Labov, W. (1972b). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1984): Field methods of the project on linguistic change and variation. In: Baugh, J. and Sherzer, J. (Eds.), *Language in use: readings in sociolinguistics*, 28-54. Englewood Cliffs, NJ: Prentice-Hall

Lasnik, H., & Lohndal, T. (2010). Government–binding/principles and parameters theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1), 40-50. https://doi.org/10.1002/wcs.35

Lesley, M. (1980). *Language and social networks*. Oxford: Basil Blackwell

Liceras, J. M. (1994). La teoría sintáctica y los juicios de gramaticalidad: en torno a la posposición del sujeto en español. *Revista Canadiense de Estudios Hispánicos 18*: 219-255.

Liceras, J. M. (2010). Second language acquisition and syntactic theory in the 21st century. *Annual Review* of *Applied* Linguistic*s 30*: 266-287

Liceras, J. M. & Fernández-Fuertes, R. (2016). Subject omission/production in child bilingual English and child bilingual Spanish: The view from linguistic theory. *Probus* 29 (1).

Liceras, J. M., Zobl, H. & Goodluck, H. (2008). Formal features in linguistic theory and learnability: the view from second language acquisition. In Liceras, J. M., Zobl, H., & Goodluck, H. (Eds.), *The role of formal features in second language acquisition*. New York: Erlbaum.

Liceras, J. M., Anahí A. de la Fuente & Cristina Martínez Sanz (2010). *The distribution of null subjects in non-native grammars: Syntactic markedness and interface vulnerability*. In

Michael Iverson et al. (ed.), Proceedings of the Mind/Context Divide Workshop, 84–95. Somerville, MA: Cascadilla Projeedings Project.

Liceras, J. M., Raquel Fernández Fuertes & Anahí Alba de la Fuente. (2012). Subject and copula omission in the English grammar of English-Spanish bilinguals: on the issue of directionality of interlinguistic influence. *First Language 31(1-2):* 88-115.

Liceras, J. M., Mendez-Rivera, N. & Moreno-Mancipe, L. (2017). Perspectives on L2 Near-nativeness: linguistic, psycholinguistic, sociolinguistic and pedagogical approaches (pp. 53-74). In J. de Dios Martínez Agudo (ed.). Native and Non-Native Speakers in English Language Teaching: Perspectives and Challenges for Teacher Education. Berlin: De Gruyter Mouton.

Lipski, J. M. (2010). Spanish and Portuguese in contact. In R. Hickey (Ed.), *The Handbook of Language Contact* (pp. 550-580). https://doi.org/10.1002/9781444318159.ch27

Lipski, J. M. (1986). Reduction of Spanish word-final /s/ and /n/. *Canadian Journal of Linguistics* 31, no. 2, 139-156.

Lowther, K. (2004). First person subject pronoun expression in the Spanish of Tucson. Divergencias. *Revista de estudios linguísticos y literarios*. Volumen 2 n° 2.

Lubbers Quesada, M. (2015). *The L2 acquisition of Spanish subjects: Multiple perspectives* (Vol. 50). Berlin: Walter de Gruyter.

Mansen, K., & Mansen, R. (1984). *Aprendamos guajiro: Gramática pedagógica de guajiro*. Bogotá: Editorial Townsend.

Martínez A. (2019). La Guajira: caracterización departamental y municipal. Cerrejon: Responsible mining & Fedesarrollo. Bogota.

Martínez-Sanz, C. (2007). *Dialectal microvariation and the null subject parameter: The case of Dominican Spanish*. Ms, University of Ottawa.

Martínez-Sanz, C. (2011). *Null and overt subjects in a variable system: The case of Dominican Spanish*. Ottawa, Canada: University of Ottawa dissertation.

Matera, M. (2001). La incorporación del nombre al verbo en Wayuunaiki. *Lingua Americana*. Año V, n° 9, 69-82.

Mayol, L. (2012). An account of the variation in the rates of overt subject pronouns in Romance. *Spanish in Context.* 9:3. pp. 420–442.

McCarthy, C. (2004). Underspecification and default morphology in second language Spanish. In *Proceedings of the 29th Boston University Conference on Language Development*. Boston MA.

Meechan, M & Foley, M. (1994). On resolving disagreement: linguistic theory and variation - there's bridges. *Language Variation and Change*. 6, *1*63-85.

Meisel, J. (1991). Principles of Universal Grammar and strategies of language learning: some similarities and differences between first and second language acquisition. In Eubank, L., (Ed.), *Point counterpoint: Universal Grammar in the second language*. Amsterdam: John Benjamins, 231–76.

Meisel, J. (1997). The acquisition of the syntax of negation in French and German: contrasting first and second language development. *Second Language Research.* 13, 227–63.

Mejía, P. (2011). *Situación sociolingüística del wayuunaiki: Ranchería El Pasito* (Unpublished master´s dissertation). Universidad Nacional de Colombia, Bogotá.

Michnowicz, J. (2015). Subject Pronoun Expression in Contact with Maya in Yucatan Spanish. In A. M. Carvalho, R. Orozco & N. Lapidus Shin (Eds.), *Subject Pronoun Ex-pression in*

Spanish: A Cross-dialectal Perspective. 103-122. Washington, DC: Georgetown University Press.

Montalbetti, M. (1984). *After Binding. On the Interpretation of Pronouns. Cambridge*, MA: MIT Ph. D. dissertation.

Montes, J. (1982). El español de Colombia: Propuesta de clasificación dialectal. *Thesaurus*, 37, 23-92.

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

Moreno, L. (2019). *Subject Pronominal Expression In Uwa Spanish*. Ottawa, Canada: University of Ottawa dissertation.

Nicol, J., & Greth, D. (2003). Production of subject–verb agreement in Spanish as a second language. *Experimental Psychology*, 50, 196–203.

Olza, J., & Jusayú, M. Á. (1978). *Gramática de la lengua guajira*. Caracas: Universidad Católica Andrés Bello, Centro de Lenguas Indígenas.

Oquendo, L. (2003). Algunas variaciones lingüísticas en las lenguas en contacto: Guajiro-español. *Letras*, (67), 143-160.

Orozco, R., & Guy, G. (2008). *El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana?* In M. Westmoreland, & J. A. Thomas (Eds.), Selected proceedings of the 4th International Workshop on Spanish Sociolinguistics (pp. 70-80). Somerville, MA: Cascadilla Press.

Otheguy, R., & Zentella, A. C. (2012). *Spanish in New York: language contact, dialectal leveling, and structural continuity*. Oxford; New York: Oxford University Press.

Otheguy, R., Zentella, A., & Livert, D. (2007). Language and dialect contact in New York: Toward the formation of a speech community. *Language* 83(4): 770–802.

Palacios Alcaine, A., & Calvo-Pérez, J. (Eds.). (2008). *El español en América: contactos lingüísticos en Hispanoamérica* (1. Ed). Barcelona: Ariel.

Paredes Ferrer, Á. (2006). *Alonso de ojeda: el descubridor de Colombia*. Editorial Planeta Colombiana: Colombia.

Parrot, J. (2007). *Distributed Morphological Mechanisms of Labovian Variation in Morphosyntax*. Georgetown University: Doctoral Dissertation.

Pérez van-Leenden, F. J. (1998). *Wayuunaiki: estado, sociedad y contacto*. Maracaibo: Universidad del Zulia.

Pérez, L. A. (2006). Los wayuu: tiempos, espacios y circunstancias. *Espacio Abierto*, 15(1 y 2), 403-426.

Perlmutter, D. (1971). *Deep and surface constraints in syntax*. New York: Holt, Rinehart and Winston.

Perrin, M. (1995). *Los practicantes del sueño: el chamanismo wayuu*. Caracas: Monte Ávila Editores Latinoamericana.

Picallo, C. (2014). Introduction: Syntactic variation and Minimalist inquiries. In *Linguistic variation in the minimalist framework*. Edited by Picallo, M Carme. Oxford University Press.

Pimienta Prieto, M. M. (2008). El bilingüismo entre los wayuunaiki hablantes y la interferencia como producto del contacto de lenguas. *Entretextos*, 2.

Pineda Carnacho, R. (Ed.). (2005). Trayectoria y desafío de la política de lenguas indígenas en Colombia. In A. Kowii, *Identidad lingüística de los pueblos indígenas de la región*

andina (Universidad Andina Simón Bolívar / Instituto Ítalo-Latinoamericano / Abya-Yala, Vol. 1, pp. 29-52). Quito.

Pladevall Ballester, E. (2013). Adult L2 Spanish development of syntactic and discourse subject properties in an instructional setting. *RAEL: Revista Electrónica de Lingüística Aplicada*, 12, 111–129.

Polo Acuña, J. (2005). *Etnicidad, conflicto social y cultura fronteriza en la Guajira* (Unknown edition). Bogotá: Universidad de Los Andes.

Polo Acuña, J. (2011). Los indígenas de la Guajira en la independencia de las provincias caribeñas de la Nueva Granada: una aproximación. *Memoria y Sociedad*, 15(30), 21-37.

Polo Acuña, J. T. (2010). *Indígenas, poderes y mediaciones en la Guajira en la transición de la Colonia a la República* (1750-1850) (PhD. dissertation, Universidad Central de Venezuela). Recuperado de http://saber.ucv.ve/bitstream/123456789/5181/1/Tesias%20H2009%20P6.pdf

Poplack, S. (1993). Variation theory and language contact: concepts, methods and data. In D. R. Preston (Ed.), *American Dialect Research* (pp. 251-286). https://doi.org/10.1075/z.68.13pop

Poplack, S. & Meechan, M. 1998 How languages fit together in code-mixing. *International Journal of Bilingualism* 2: 127-138.

Poplack, S. (2018). *Borrowing: Loanwords in the Speech Community and the Grammar*. https://doi.org/10.1093/oso/9780190256388.001.0001

Poplack, S., & Tagliamonte, S. A. (2001). *African American English in the diaspora*. Malden, MA: Wiley-Blackwell.

Poplack, S., & Levey, S. (2010). Contact-induced grammatical change: A cautionary tale. In *Language and Space: An International Handbook of Linguistic Variation*, P. Auer and J. Schmidt, (Eds.), Walter de Gruyter: Berlin.

Prévost, P. (2008). Knowledge of morphology and syntax in early adult L2 French: Evidence for the Missing Surface Inflection Hypothesis. In Liceras, J. M., Zobl, H., & Goodluck, H. (Eds.), *The role of formal features in second language acquisition*. New York: Erlbaum.

Prévost, P., & White, L. (2000a). Accounting for morphological variability in second language acquisition: Truncation or missing inflection? In M.-A. Friedemann & L. Rizzi (Eds.), *The acquisition of syntax* (pp. 202–235). London: Longman.

Prévost, P., & White, L. (2000b). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 103–134.

Ramírez Boscán, K. (Ed.). (2007). *Desde el desierto - notas sobre paramilitares y violencia en territorio wayuu de la media guajira (Cabildo Wayúu Nóüna de Campamento)*. Maicao, Colombia.

Ramírez Cruz, H. (2009). La inconcordancia de género y número en el contacto de lenguas. *Forma y Función*, Vol. 22 # 2. Pp. 165-195. Bogotá, Colombia.

Regúnaga, A. (2005). El masculino como género marcado en la lengua wayuu (guajiro). *Anuario de la Facultad de Ciencias Humanas*, Universidad Nacional de la Pampa. (7) 231-243.

Rizzi, L. (2014). On the elements of syntactic variation. In M. C. Picallo (Ed.), *Linguistic Variation in the Minimalist Framework* (pp. 12-35). https://doi.org/10.1093/acprof:oso/9780198702894.003.0002

Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris.

Roy, J. (2008). *Extending Variable Rule Analysis*. Manuscript from CVC talk.

Sánchez Arroba, M. (2011). Factores pragmáticos asociados con la expresión del sujeto pronominal en el español de Cuba. *XVI congreso de la ALFAL*, Alcalá.

Sankoff, D. (1988). Variable rules. In U. Ammon, N. Dittmar, & K. J. Mathheier (Eds.), Berlin Sociolinguistics: An *International Handbook of the Science of Language and Society*, Vol. 2. Walter de Gruyter. 984-997.

Sankoff, D. (1982). Sociolinguistic method and linguistic theory. In L.J. Cohen, et al. (Eds.), *Logic, Methodology, Philosophy of Science VI*. Amsterdam: North Holland and Warsaw. Polish Scientific: 677-689.

Sessarego, S. (2014). *The Afro-Bolivian Spanish determiner phrase: a microparametric account*. Columbus: The Ohio State University Press.

Sheehan, M. (2006). *The EPP and null subjects in Romance*. Newcastle: Newcastle University PhD dissertation.

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*. Madrid: Iberoamericana.

Shin, N. L. (2012). Variable Use of Spanish Subject Pronouns by Monolingual Children in Mexico. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium* (Ed.), K. Geeslin & M. Díaz-Campos, 130-141. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2661.

Silva-Corvalán, C. (1982). Subject expression and placement in Mexican American Spanish. In J. Amastae & L. Elías-Olivares (Eds.), *Spanish in the United States: Sociolinguistic aspects*, 93-120. Cambridge/New York: Cambridge University Press.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles. Oxford studies in language contact*. Oxford: Clarendon.

Smith N. & Cormack A. (2015). Features from Aspects Via the Minimalist program to Combinatory Categorial Grammar. In Á. Gallego & D. Ott (eds.). *50 Years Later: Reflections on Chomsky's Aspects*. (pp 233-248). Cambridge, MA: MITWPL.

Snyder, W. (2001). On the Nature of Syntactic Variation: Evidence from Complex Predicates and Complex Word-formation. *Language*, (77), 324-343.

Sorace, A. (2005). Syntactic optionality at interfaces. In L. Cornips & K. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and the social* (pp. 46–111). Amsterdam: John Benjamins.

Sorace, A. (2011). *Pinning down the concept of "interface" in bilingualism. Linguistic Approaches to Bilingualism*. 1.1: 1-33.

Sorace, A. (2003). Near-Nativeness. In Catherine J. Doughty and Michael H. Long (Eds.), *The Handbook of Second Language Acquisition*, 130-153. Oxford: Blackwell Publishers. http://dx.doi.org/10.1002/9780470756492.ch6

Sorace, A.; Serratrice, L.; Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua* 119: 460-477.

Tagliamonte, S. (2006). *Analyzing sociolinguistic variation*. Cambridge: Cambridge University Press.

Tagliamonte, S. (2011). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Wiley-Blackwell.

Tagliamonte, S. (2013). Analysing and interpreting variation in the sociolinguistic tradition. In K. Manfred & J. Schlüter (Eds.), *Research methods in language variation and change*. Cambridge: Cambridge University Press.

Tagliamonte, S. (2013b). Comparative Sociolinguistics in *The Handbook of Language Variation and Change*. Edited by J.K. Chambers & N. Schilling. J. Wiley & Sons, Inc.

Taraldsen, T. (1978). *On the NIC, vacuous application and the that-trace filter*. MIT and Indiana Linguistics Club.

Tercer Concilio Limense (1583–1591): Edición bilingüe de los decretos. Edited by Martínez Ferrer, Luis. Translated by Gutiérrez, José Luis. Lima: Facultad de Teología Pontificia y Civil de Lima, 2017. Notes. 347 pp.

Thomason, S. (2011). *Language contact: an introduction*. Edinburhg: Edinburg University Press.

Torre Revello, J. (1962). La enseñanza de las lenguas a los naturales de América. *Thesaurus*, 1(3).

Torres-Cacoullos, R., & Travis, C. (2018). Subject Pronoun Expression: Reconsidering the Constraints. In *Bilingualism in the Community: Code-switching and Grammars in Contact* (pp. 74-110). Cambridge: Cambridge University Press. doi:10.1017/9781108235259.005

Torres-Cacoullos, R., & Travis, C. E. (2010). Variable yo expression in New Mexico: English influence? In S. Rivera-Mills & D. Villa (Eds.), Spanish of the U.S. *Southwest: A language in transition* (pp. 185-206). https://doi.org/10.31819/9783865278692-012

Travis, C. E. (2007). Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19, 101-135.

Vergara González, O. (1987). Guajiros. In *Introducción a la Colombia Amerindia*. Instituto Colombiano de Antropología, Bogotá.

Walker, J. A. (2010a). *Variation in linguistic systems*. New York/London: Routledge. https://doi.org/10.4324/9780203854204

Walker, J. A. (Ed.). (2010b). *Aspect in grammatical variation*. https://doi.org/10.1075/silv.6

Wardhaugh, R. (2006). *An introduction to sociolinguistics* (5th ed). Malden, Mass. USA: Blackwell Pub.

Weinreich, U., Labov, W., & Herzog, M. (1968). *Empirical foundations for a theory of language change*. Austin: University of Texas Press

Young, R., & Bayley, R. (1996). Appendix: varbrul analysis for second language acquisition research. In R. Bayley & D. R. Preston (Eds.), *Studies in Bilingualism* (Vol. 10, pp. 253-306). https://doi.org/10.1075/sibil.10.11you

10. APPENDIX A. INTERVIEW SCHEDULE

The Spanish spoken by the Wayuu indigenous people of Uribia, Colombian Guajira. The case of gender agreement and verb-subject agreement.

Interview schedule

Guideline questions

(Adapted from Tagliamonte, 2006 and Martinez-Sanz, 2007)

The student investigator will carry out what are usually known as 'sociolinguistic interviews. These guided interviews aim to reproduce an informal conversation between the participant and the interviewer. It has been shown (Labov, 1984) that it is in informal conversations where the speakers of a language produce their most natural speech, the one that is closer to the 'vernacular' use of their own language. In order for this to happen, it is essential that the interviewee feels comfortable at all times with the topics that are talked about and with the way the interviewer develops this guided conversation. It is also essential that the interviewee does not perceive the interviewer as imposing any value-judgment on her/him.

The main objective of a sociolinguistic interview is to record approximately an hour of speech from the interviewee in which she or he produces narratives, and to obtain the demographic data necessary to carry out the analysis of the data gathered. The interview consists in a hierarchically structured set of questions that constitute conversational modules. The interview begins with questions related to demography and community/neighbourhood, and progress into conversational modules that may elicit narratives of personal experience. Specifically, eight conversational modules have been designed for this study:

1. Demographics
2. Neighbourhood/Community and Social Practices
3. Parents and family
4. Work life
5. Traditions in the Family/Community and Folk Remedies
6. Travel
7. Uncommon Experiences
8. Language

Within each module, the questions are hierarchically organized as well. They begin with exploratory queries, designed to assess whether the interviewee is interested or willing to talk about that particular subject. If the interviewee shows an interest in the topic, the interviewer will go on and ask the more detailed questions. If not, the interviewer will move to the next module until she finds something that the interviewee enjoys talking about.

The modules and questions for this research project have been selected taking into account the specific characteristics and ways of living of the communities under investigation. For instance,

the module on Neighbourhood/Community and Social Practices has questions on the way people in the community relate to each other. These questions have been designed taking into account the fact that in these communities, especially in the rural areas, there are strong bonds not only among families, but also among neighbours and the larger community. The module on Travel has been designed bearing in mind that is fairly common for the people in these communities to immigrate or go to work outside the Guajira or the fact they move freely between Colombia and Venezuela for short or long periods of their lives. Also, the module on Traditions in the Family/Community and Folk Remedies focuses on specific practices and celebrations that take place in these communities.

The interviewer will be careful at all times with the wording she chooses for her questions. She will avoid highly direct questions as much as she can, she will ask questions in the context of providing a personal example when possible, and she will neutralize any questions that might be perceived as imposing a value-judgment on the interviewee. For instance, when asking interviewees about the amount of education they have, instead of asking directly 'How many years of education did you get?', he will ask 'Did you have the chance to go in school?', acknowledging the fact that some people have not had much opportunity to go to school. If any delicate topics come up, for instance talking about events that happened in the community such big rains or hurricanes, the interviewer will focus on the positive aspects of the community that these events brought up, such as the way people helped each other. This will be done this way to avoid triggering any unpleasant memories or events in the interviewee.

The structure of the interviews will be as follows: first, the student investigator will go over the informational letter and will reiterate the main issues related to her or his participation in this research project, namely anonymity, confidentiality, voluntary participation, right to not answer any of the questions posed by the interviewer and/or to withdraw from the project at any time, and finally right to listen to their interviews, ask the interviewer to erase any part of them, and right to access the transcripts of their interviews and the quotes from them that the student investigator will include in her doctoral dissertation. If permission to record the interview is not granted, the interview will continue but the data from this participant will not be analysed. After the participant signs the consent form or records her or his consent, the interview will begin. The student investigator will guide the conversation using the aforementioned conversational modules that have been prepared for this research project.

1. Demographics

Note: Although this module comes first in the Interview Schedule, some of these questions will probably be interwoven into the interview situation rather than asked at the outset. This will depend on how comfortable the interviewee feels with supplying personal information to the interviewer in the course of their conversation. The interviewee will have the choice to supply only the personal information she or he feels like sharing with the interviewer.

Your name is?

How old are you?

And your address is?

How long have you lived at this address/town?

Where were you born?

Where else have you lived?

Where were your parents born and raised?

Your grandparents?

Your spouse?

Are you working now? Where?

What did/do your parents do?

Your husband/wife?

How many years of school did you have a chance to finish?

What was the first job you had when you left school?

If the interview is not taking place at the informant's home:

Can you tell me about your home/*ranchería*[46]?

What kind of place is it?

How is it laid out? How many families live in your home/*ranchería*?

What do they do for living?

1. Neighbourhood/Community and Social Practices

This is such a nice/interesting/colourful neighbourhood! How long have you lived here?

What kind of people live on your street? In this area?

[46] We are using the term ranchería because it is the typy of community organization mostly found among the Wayuus. A ranchería is usually a piece of land with 3 to five shelters for equal number of families.

What did it make you live/move here?

Because of work?

Because of community roots?

To live close to your family?

How has your neighbourhood changed in your lifetime?

Do you feel that your neighbourhood is as safe as it was when you were growing up?

Why or why not?

Do people from around here drop by to visit?

Do they just walk in to sit and have coffee or something with you?

Who and what are your neighbours like?

Do people get together outside their houses?

Do you have a cold drink or a coffee in the afternoon? In the evenings?

When or in what special occasions do you drink *chicha*[47]?

Are there people around here you spend a lot of time with outside your family?

Do they live nearby? Whereabouts?

What do you do together?

Do you ever go to dances?

Where do you go?

Did you dance?

What kind of dancing did you do?

What would people wear?

[47] Local drink made out of corn.

What do you like the best about your neighbourhood/community?

What are the things that make you feel good about your neighbourhood?

Did anything really big ever happen here that you remember?

Any big rains or droughts? How often does it rain in the zone?

Did people in the neighbourhood help out? With food, clothes, place to stay?

Do you have anyone you can go to for help around here? What do you do?

Do you have anyone who you help around here? What do you do?

Do you ever get fruit or vegetables from your neighbours? Do you ever give fruit or vegetables to your neighbours?

Do you exchange food, like rice, milk or meat with your neighbours?

If you are sick who can you ask to look after the family?

Do you have to ask or people just offer?

What is the Wayuu medicine like? Do you use plants or herbs to heal?

How does a Wayuu become a folk healer?

Are dreams important to heal or for the dreamer's destiny?

What kinds of remedies do you use?

Who are the doctors or folk healers in your community?

If you needed a ride in an emergency, who would you ask?

Do you lend your car to others?

2. Parents and family

Do you have any idea how long your family has been living in this country/city/neighbourhood/*rancheria*?

Where did they come here from?

Do you remember coming to this country/city/neighbourhood? Tell me about it.

Did your parents have any ideas about what they wanted you to be?

Did you have the opportunity of going to school?

What did your parents want you to work in?

Do you have siblings? How many?

Were you close to your siblings growing up or did you fight a lot?

How about now?

Did you/Do you spend much time with your _____?

Grandparents, aunts and uncles, cousins, etc

How did you meet your wife/husband/partner?

Why did you decide to get married?

How did the marriage proposal happen?

Can you remember what you said?

Can you remember how your wife/husband reacted?

How much did you/he pay for you?

How was the arrangement between the families?

How was your wedding like? Did anything funny/interesting happen?

Did you ever live with your in-laws after you get married?

How did that work out?

Tell me about your first house together.

Who picked it out?

Where was it?

Were you excited?

How long did you live there?

3. Work life

What was your very first job?

How old were you when you started to work?

Can you remember how much you earned?

Do you remember what you were excited to spend your earned money?

What did your parents do to earn a living?

Do young people feel the same way about working that they did in your day?

What did your parents want you to do for a living?

What do you do?

What would you like to do?

4. Traditions in the Family/Community and Folk remedies

What kind of traditions can you remember growing up in your family/community?

Do you (plan to) keep these traditions alive with your own family?

Did people go to doctors in the older days? Do you go to the doctor when you are sick?

I have heard people say that in the older days, people relied on traditional medicines and natural remedies, do you remember that?

What kinds of remedies do you remember?

A lot of people talk about the *piache*[48], what was/is that?

How does he/she do?

What kind of things do they cure?

What do you do to prevent colds/flu?

Let's talk about celebrations in your family/community.

How do you celebrate Christmas?

What kind of typical celebrations do you have?

Is there any typical drink for Christmas? Could you tell me what is *chirrinchi*[49]?

What do you usually do on New Year's Eve? Do you dance a typical dance?

When do you dance the *yonna*[50]?

5. Travel

Have you had the opportunity to travel?

Where did you go? How long? Did anything interesting happen? What other places in the Guajira do you know?

[48] A piache is the Guajiro doctor or folk healer.
[49] An alcoholic beverage
[50] A Guajiro dance.

Has a communication barrier created any funny moments in any of your travels?

Did you ever lose your luggage? Miss a bus?

Where would you like to go that you've never been?

Why?

What is the funniest thing that ever happened to you when you were travelling?

Can you tell me the differences among the Upper, Middle and Lower Guajira?

6. Uncommon experiences

When people think back on their lives, there is always something that sticks out as being really unusual… did you ever have anything like that happen to you?

Sometimes in families there is someone who gets a feeling that something is going to happen, and it does happen.

Is there anybody like that in your family?

Do you remember anything like that that came true?

Have you had a revelation through a dream?

What did you dream about? What did it happen?

Have you had problems with other people?

How did you solve the situation?

Have you ever been somewhere new and know that you have been there before?

7. Language

Is it true the Wayuus prefer to speak more Spanish than Wayuunaiki?

What language do you prefer to pass to your children, Spanish or Wayuunaiki?

In your case, what language do you prefer to speak?

Here in your community, what do people think of the Spanish and the Wayuunaiki languages?

Do non-Wayuu people learn Wayuunaiki?

Do they speak it well?

How about the differences between old and young speakers when they speak Wayuunaiki? Do you sound the same as your parents? Do your parents sound the same as you? Do your kids?

Are there differences among the Wayuunaiki spoken in the Upper, Middle and Lower Guajira?

At school, do you receive classes in Wayuunaiki?

Is it difficult for you to speak Spanish?

Where and how did you learn your Spanish?

Have you ever tried to change the way you talk? What did you do?

11. APPENDIX B

The Spanish spoken by the Wayuu indigenous people of Uribia, Colombian Guajira. The case of gender agreement and verb-subject agreement.

Informational letter and informed consent form
(Adapted from Tagliamonte, 2006 and Martínez-Sanz, 2007)

Student investigator:	Thesis supervisor:
Nelson José Méndez-Rivera	Dr. Juana M. Liceras
Graduate student	Professor
Department of Modern Languages and Literatures	Department of Modern Languages and Literatures
University of Ottawa	University of Ottawa

Date: _____

Dear Madam/Sir,
I am a PhD student at the University of Ottawa, Canada working under the supervision of Professor Juana M. Liceras. I am conducting a study for my doctoral dissertation about the Spanish spoken by the Wayuu people. The goal of this project is to explore and characterize aspects of the Spanish used by the Wayuu indigenous community of Uribia.

So, I am interviewing and carrying out some linguistic experiments with people belonging to the Wayuu ethnic group like you and talking to them about their experiences, memories and stories about living here in the Guajira and particularly in Uribia. I am interested in knowing what the Wayuus think about the use and learning of the Wayuu and Spanish language in the community of Uribia and in general among the members of your group and also how they have learned Spanish and how they speak it. I hope this project will help to inform the scientific community and the people in general living outside the Guajira about the language and daily life in this region.

What I do is simply sit down and talk to people for about an hour, or as long as you feel comfortable talking to me. The conversation will take place wherever you feel more comfortable: in a coffee shop or restaurant, at the village's square, in the street, at your home or at your workplace. We will talk about topics such as your community, your neighbourhood, the Wayuu's traditions and culture, your work life and general personal experiences about growing up and living in Guajira. Also, the interview includes an experimental part, as described here. First, there will be a picture elicitation task. This task consists, on the one hand, of a series of pictures from which I (the interviewer) will ask you specific questions. For example, in a picture there will a couple of people dancing. I will ask you a question such as the following: " ¿que hacen ellos? what are they doing?" you are expected to utter a full answer, like, "Elllos bailan una danza indigena, they are dancing an indigenous dance". This first part of the photo task is composed of 40 pictures and consequently 40 questions. To elicit answers for first person singular and first person plural, I will ask you direct questions instead of showing you photos. On the other hand, the picture elicitation task is

complemented with a story telling. For this, you are shown 11 pictures and you have to construct a story based on the pictures.

Because I won't be able to remember everything everyone says exactly, I ask your permission to record our conversation. If you agree, I would like your permission to place this digital recorder near you. Everything that is said during this interview is kept entirely confidential and you will remain entirely anonymous. If I quote excerpts from this interview in my doctoral dissertation, I will make sure to leave all your personal information out of the quote so no one could guess from whom the words are coming. You can choose to listen to your interview after we do it, and you will be given the chance to review the transcripts of your interview and the quotes I will include in my PhD exam. If you want to review the transcripts and quotes, please leave with me either your mailing address or your e-mail address and I will make sure you get a copy of the materials. All data and consent forms will be kept at a locked office in the University of Ottawa and destroyed after five years. No one will have access to the data except my thesis supervisor, Juana M. Liceras, my co-supervisor Stephen Levey, and me.

Participation in this project is entirely voluntary and there is no monetary compensation; however, if you are interested in my research topic, I can give you a personal class about the topic of my study. Also, it is very unlikely that your participation in this project causes you any harm or inconvenience. However, if you feel that your participation in this project is causing you any physical o emotional inconvenience or harm, or if you feel that it is disrupting your daily routine, or simply you get bored or you regret disclosing any personal information, you may withdraw at any time if you choose. You can tell me to not record part of what you say, or to erase part of it. It is entirely up to you.

By participating in this study, apart from being of a lot of help in my preparation for my PhD exam, you will have a chance to reflect on your use of the Spanish language, your community, and your ethnic group. You will also be contributing to our understanding of languages in contact, and to our general understanding of the Spanish spoken in the Guajira.

Please return this page to the researcher if you wish to participate.

I, _____ agree to take part in a study about the Spanish spoken by the Wayuus.

I understand that to take part in the study I will participate in a recorded conversation, involving discussion of topics such as the following:

-My neighbourhood/community
-My family and our traditions
-Life in the Guajira
-My job
-Opinions about the knowledge and use of Spanish among the Wayuus

-Traditions, culture and habits of the Wayuus

I understand that the interview will take about an hour or as long as I wish. I understand that the interview can occur at a time and place that is convenient for me.

I understand that I am under no obligation to participate in this study and that

I may refuse to answer any questions, to stop the interview at any time, or withdraw from the study at any time. I understand that my participation is voluntary and there is no monetary compensation. However, if I want, I can ask the researcher to give me a class about his study topic.

I understand that I might benefit from the study by reflecting on my own use of the Spanish language. I also understand that the results from this research project may help to inform people outside the Guajira about the Spanish spoken in the region, and it may help researchers to understand how language contact works

I understand that I can listen to my interview after is finished, and that I can chose to have access to the transcripts of my interview and to the quotes from it that the student investigator may take for his PhD exam.

I understand I can also contact Professor Juana M. Liceras if I have any other questions regarding this research project.

If I have any questions regarding the ethical conduct of this study, I may contact the Protocol Officer for Ethics in Research, University of Ottawa, 550 Cumberland Avenue, Room 154, Ottawa, ON, K1N 6N5.

Email: ethics@uottawa.ca

I understand what this study involves, and I agree to participate. I have been given a copy of this consent form.
Participant's signature:

_____

Participant's full name: _____

Participant's mailing address / e-mail address (optional):

_____

_____

Student investigator's signature:

_____

Date:

_____

12. APPENDIX C

UNIVERSIDAD DE OTTAWA • DEPARTAMENTO DE LENGUAS Y LITERATURAS MODERNAS
LABORATORIO DE INVESTIGACIÓN SOBRE LA ADQUISICIÓN DEL LENGUAJE (LAR-LAB) - 2012

Cuestionario de conocimiento de lenguas

Fecha:

1. Nombre y apellidos: _____
2. Sexo: H [ ] M [ ]
3. Edad: 18-30 [ ] 31-40 [ ] 41-50 [ ] 51+ [ ]
4. Lugar de nacimiento: _____
5. Lengua primera: _____
6. Lengua dominante de la madre: _____
7. Lengua dominante del padre: _____
8. Lengua (s) que hablaba en casa cuando era niño: _____
7. Lengua (s) que habló durante los primeros tres años de su vida: _____
9. Lengua (s) que habló durante los primeros cinco años de su vida: _____
10. ¿Dónde aprendió español? _____
11. ¿A qué edad empezó a hablar español? _____
12. Lenguas que usa: _____

En casa: _____
En la ranchería: _____
En el pueblo/ciudad: _____
En la escuela/Universidad: _____
Con su pareja: _____
Con sus hijos: _____
Con la madre: _____
Con el padre: _____
Con los amigos: _____

13. ¿En qué lengua se siente más cómodo en la actualidad?
_____