

Electronic Thesis and Dissertation Repository

7-7-2021 11:00 AM

Variable Subject Pronoun Expression in the Spanish of Londombia: A study of language contact in Canada

Angelica Hernandez Constantin, *The University of Western Ontario*

Supervisor: Heap, David, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in French

© Angelica Hernandez Constantin 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Anthropological Linguistics and Sociolinguistics Commons](#)

Recommended Citation

Hernandez Constantin, Angelica, "Variable Subject Pronoun Expression in the Spanish of Londombia: A study of language contact in Canada" (2021). *Electronic Thesis and Dissertation Repository*. 8020. <https://ir.lib.uwo.ca/etd/8020>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

According to the extended projection principle, the presence of a subject in a clause is mandatory (Chomsky 1982). Overt presence of a subject, however, is not always required and languages vary in their use of null and overt subjects. Languages such as English require that a subject is overtly expressed (1a), rendering phrases with a null subject ungrammatical (1b), while languages like Spanish allow for use of both overt (2a), and null subjects (2b).

- (1) a. She wants bread.
b. * \emptyset wants bread.
- (2) a. Ella quiere pan.
b. \emptyset quiere pan.
“(She) wants bread”

The variable use of Spanish subject personal pronouns (SPPs) has been studied in monolingual and bilingual populations. Studies show that populations differ in the frequency with which overt SPPs are used. In addition studies of Spanish-English bilingual populations have observed higher frequency of use of overt SPPs as a function of speakers' length of residence in the U.S. (Montrul 2004; Otheguy and Zentella 2012). It is suggested that the observed higher frequency is due to the situation of contact with English which is a non-null subject language (Silva-Corvalán 1994). However, other studies suggest that the higher frequency of overt SPP usage may be due to contact with Caribbean Spanish varieties in the U.S. (Flores-Ferrán 2004). Despite the fact that the variable expression of Spanish SPPs has been studied in bilingual populations, most of these studies are centered on populations in the U.S. Since the Spanish-English contact situation in Canada differs from that in the U.S. according to various socio-historic factors, and critically due to the difference in the presence of Caribbean Spanish varieties, studying the variable expression of Spanish SPPs in Canada is an important contribution to the discussion on the variable expression of SPPs and its use in bilingual populations.

In this dissertation, the results of an investigation regarding the variable use of SPPs in two generations of Colombian-Canadian Spanish speakers ($N_{1Gen} = 10, N_{2Gen} = 10$) living in London, Ontario are reported. A total of 2366 tokens were extracted from the sociolinguistic interviews and used to calculate frequency of use of overt SPPs for each generation, and to determine the social (generation, age, gender, and interview modality), and linguistic factors (pronoun person and number, switch reference, semantic verb type, clause negation, position of pronoun in relation to verb, verb tense, verb mood, and clause type) that condition variable use of SPPs in this population. In addition, this study adopts an embedded mixed-methods approach by also considering the data from sociolinguistic interviews from a qualitative perspective. This allows me to provide valuable contextual information for the quantitative analyses and to explore whether the attitudes, language use habits, and ties to cultural identity of Colombian speakers align with factors known to favour heritage language maintenance across generations.

Keywords: Language contact and change, bilingual and heritage speakers, Spanish, Subject pronouns, Canadian studies, and sociolinguistics.

Summary for Lay Audiences

Although all languages require the use of subjects for tensed clauses (Chomsky 1982), languages vary in how they use subject pronouns. For instance, while languages like English require that subjects are expressed (1a), rendering phrases with non-expressed subjects ungrammatical (1b), in languages like Spanish subjects can be expressed (2a) or they can be omitted (2b).

- (1) a. She wants bread.
b. * \emptyset wants bread.

- (2) a. Ella quiere pan.
b. \emptyset quiere pan.
“(She) wants bread”

In Spanish this variable use of subject pronouns has been studied in monolingual (Cameron 1992; Orozco 2015) and bilingual populations (Montrul 2004; Otheguy et al. 2007), and studies have shown that the frequency of use of null vs. overt subject pronouns varies between varieties of Spanish. In bilingual populations, an increase in the use of expressed subject pronouns has been documented in second generation immigrant speakers of Spanish (Otheguy et al. 2007). However, it is debated whether this effect is due to contact with English or with other varieties of Spanish, which also show a higher use of expressed subject pronouns, such as Caribbean varieties of Spanish (Flores-Ferrán 2004).

In this dissertation, I report on the results of my investigation on the variable expression of subject pronouns in two generations of Colombian Spanish speakers living in London, Ontario ($N_{1Gen} = 10, N_{2Gen} = 10$). A total of 2366 tokens from sociolinguistic interviews are used to calculate the frequency of use of expressed subject pronouns for each generation, and to determine the social (generation, age, gender, and interview modality), and linguistic factors (grammatical person and number, switch reference, semantic verb type, clause negation, position of pronoun in relation to verb, verb tense, verb mood, and clause type) that condition the variable expression of subject pronouns in this population. In addition, this study uses a mixed-methods approach by considering quantitative and qualitative data. The use of qualitative data provides valuable contextual information for the quantitative analysis and allows me to investigate whether the attitudes, language use habits, and ties to cultural identity of Colombian speakers align with factors known to favour heritage language maintenance across generations.

Dedication

The one who plants trees, knowing that he will never sit in their shade, has at least started to understand the meaning of life.

-Rabindranath Tagore

Con mi mayor admiración, esta tesis está dedicada a tod@s l@s inmigrantes de primera generación en Canada, pero especialmente a l@s inmigrantes de primera generación de la comunidad Colombiana de London, Ontario.

(With my utmost admiration, this thesis is dedicated to all first generation immigrants in Canada, but especially to the first generation immigrants from the Colombian community of London, Ontario.)

Acknowledgements

The completion of this work would not have been possible without the support of many individuals.

I would like to thank my supervisor, Dr. David Heap, for all his helpful comments and guidance, for the innumerable hours of discussion, which I will deeply miss, and for his patient and unwavering support throughout my Ph.D. program. I also would like to offer my sincere gratitude to those who offered their time and provided valuable feedback on my work, including Dr. Tania Granadillo, Dr. Jeff Tennant, and Dr. Joyce Bruhn de Garavito from Western University, as well as to Dr. Naomi Nagy from the University of Toronto.

My sincere thanks also go to Western University, but more importantly, to the staff and linguistics faculty of the French and Hispanics Studies departments. Throughout my time at Western, I had the opportunity to work alongside many experts in different fields and to develop deep, life-long friendships with students, staff, and faculty alike. These bonds provided much needed support to push forward at many points throughout my program, even while in the midst of a global pandemic.

I also would like to recognize and thank Santiago Tellez Alvarez for his assistance. His work on this project was valuable in ensuring timely completion of the transcriptions. It surely would have taken much longer to complete the transcription work without his collaboration. Additionally, a special thanks is extended to the Colombian community of London, Ontario. I am especially grateful to those who volunteered their time to participate in this study. Having grown up in London, Ontario, I am familiar with the community and I have a deep admiration for all members of this community and the strength with which they face a myriad of struggles every day.

Finally, but no less importantly, I want to thank my family and friends. The unconditional support I received from them was instrumental and without it, this project may not have been completed. I could easily spend multiple pages thanking my friends at Western for the numerous chats about this project and for reminding me to take a coffee break when needed, and my family for their support, but I will reserve these thanks for private meetings (over more coffee). That being said, a big special thank you needs to be expressed to my husband Dr. Chris Plyley who not only was witness to all of my struggles, but who also patiently allowed me to discuss my ideas in length and provided technical support when LaTeX was not cooperating well with me.

Contents

Abstract	i
Summary for Lay Audiences	ii
Dedication	iii
Acknowledgements	iv
List of Figures	x
List of Tables	xii
List of Appendices	xiv
1 Introduction	1
1.1 General overview and research objectives	2
1.2 Language Contact	5
1.2.1 Types of Language Contact	6
1.2.2 Contact at the level of society	8
1 Community language varieties	12
2 Heritage speakers and society	14
3 Language shift and language death	16
1.2.3 Contact at the level of the individual	18
1 Language transfer effects	18
2 Heritage speakers and bilingualism	22
3 Language attrition and language loss	25
1.2.4 Minority language maintenance	26
1 Macro-level factors	28
2 Micro-level factors	33
1.3 Spanish-English Language Contact in North America	35
1.3.1 Spanish-English contact in the United States	36
1.3.2 Spanish-English contact in Canada	38
1.4 The Null-Subject Parameter	41
1.4.1 Null and non-null subject languages	41
1.4.2 Non-compliance with the NSP	44
1 Caribbean Spanish	44
2 English subject pronoun omission	48

1.5	Spanish Subject Pronouns	50
1.5.1	Non-variable subject expression	51
1	Categorically Null Subjects	51
2	Categorically Overt Subjects	53
1.5.2	Variable Subject Pronoun Expression	54
1	Spanish Variable Subject Expression in monolingual contexts	54
2	Spanish Variable Subject Expression in bilingual contexts	63
2	Methodology	69
2.1	Overview and research questions	69
2.2	The data	71
2.2.1	London, Ontario : Londombia	71
2.2.2	The CoSLO corpus	74
2.3	Data collection	76
1	Participant recruitment	76
2	Participant eligibility	77
3	Data collection materials and protocol	78
4	Other methodological considerations	80
2.4	Transcription and coding	81
2.4.1	Overview	81
2.4.2	The envelope of variation	82
1	Inanimate subjects	84
2.5	Factors considered: Variable expression of Spanish SPPs	85
2.5.1	Linguistic Factors	85
1	Subject pronoun type (grammatical person and number)	86
2	Verb form (tense and mood)	89
3	Semantic content of the verb	91
4	Switch reference	92
5	Clause polarity (Negation)	93
6	Clause type	93
7	Position of pronoun in respect to verb	94
2.5.2	Social factors and extra-linguistic factors	96
1	Immigration generation	98
2	Gender and socioeconomic status	99
3	Language contact intensity (Length of residence, bilingual proficiency, and language use habits)	102
4	Interview modality	108
2.6	Factors considered: Spanish language maintenance	109
2.6.1	Language maintenance factors	109
1	Use of Spanish and perceived importance of Spanish for speakers	111
2	Attitudes towards inter-generational language transmission	113
3	Attitudes towards linguistically endogamous/exogamous relationships	114
4	Access to language resources and community involvement	115

	5	Cultural association to Hispanicity and/or Colombia (2nd Generation speakers)	116
2.7		Data Analysis	118
	2.7.1	Quantitative data analysis	119
	2.7.2	Qualitative data analysis	120
	1	Qualitative study reliability measures	123
3		Results	126
3.1		Quantitative analysis: Variable subject pronoun expression in Londombia . . .	126
	3.1.1	Considering both generations	126
	1	Grammatical person and number	128
	2	Verb mood	131
	3	Switch reference	132
	4	Clause polarity	132
	5	Other factors - Non-significant or not selected as significant conditioning factors	133
	3.1.2	First Generation Speakers	135
	1	Grammatical person and number	136
	2	Switch reference	138
	3	Other factors - Non-significant or not selected as significant conditioning factors	139
	3.1.3	Second Generation Speakers	142
	1	Grammatical person and number	144
	2	Verb Mood	145
	3	Interview mode	146
	4	Switch reference	147
	5	Clause polarity	148
	6	Other factors - Not included in the multivariate logistic regression model	149
	3.1.4	Summary of logistic regression analysis results	152
	3.1.5	Generation analysis	152
	1	Pronoun position in relation to verb	154
	2	Individual speakers' considerations and generation analysis	155
	3.1.6	Quantitative Results - Summary	157
3.2		Qualitative Analysis: Language Maintenance	160
	3.2.1	Spanish language use in various contexts in Canada	162
	1	The home environment	163
	2	Social contexts	168
	3	Professional Contexts	172
	4	Public contexts	175
	3.2.2	Attitudes towards intergenerational language transmission	179
	1	Maintenance of Spanish is important	180
	2	Language maintenance too difficult	183
	3.2.3	Attitudes towards exogamous/endogamous relationships	185
	1	Language is not a main consideration in choosing a partner	186

2	Second generation wishes to facilitate communication with family	188
3.2.4	Spanish language-learning opportunities	189
1	Few/No resources available	189
2	Resources are available	190
3	Spanish is learned in the home	192
3.2.5	Identification with cultural background (Second generation only)	193
1	More Canadian than Colombian	194
2	Other's perceptions influencing identity	195
3	Blended or divided identity	198
3.2.6	Qualitative results - Summary	199
3.3	Explanatory mixed methods analysis	202
3.3.1	Mixed methods approach to individual speaker analysis	202
3.3.2	Mixed-methods analysis - Summary	211
4	Discussion	213
4.1	Variable subject pronoun expression in Londombia across two generations . . .	214
4.1.1	Grammatical person and number	216
4.1.2	Change of referent	222
4.1.3	Verbal mood	225
4.1.4	Clause polarity (negation)	227
4.1.5	Interview modality	228
4.1.6	Intergenerational analysis of the use of Spanish SPPs	231
1	No significant change in the rate of overt SPPs between generations	231
2	Increased use of pre-verbal subjects	232
4.1.7	Complexification in the second generation group	233
4.1.8	A case for cross-linguistic influence in second-generation Colombians in Canada	235
4.2	Spanish language maintenance in Londombia	241
4.2.1	Language use patterns	242
1	The home and family environment	243
2	Social (friendship) contexts	244
3	Professional contexts (School and/or work)	245
4	Public contexts and majority community's perceptions	246
4.2.2	Attitudes towards intergenerational language transmission	249
4.2.3	Attitudes towards endogamous/exogamous marriages	252
4.2.4	Access to Spanish language resources (Heritage community impact) . .	255
4.2.5	Identification with cultural background (Second generation only)	257
4.3	Individual speaker mixed-methods analysis	259
5	Conclusions	264
5.0.1	Future research	275
	References	277

Appendix A	Western Research Ethics Board Approval	292
Appendix B	Interview Guides	298
Appendix C	Language Questionnaires	305
Curriculum Vitae		308

List of Figures

3.1	Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to 3SG ($N=564$), 1SG ($N=1121$), 2SG ($N=79$), 3PL ($N=287$), and 1PL ($N=315$) grammatical person.	129
3.2	Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to singular ($N= 1764$) and plural ($N= 602$) pronoun forms.	130
3.3	Frequency of use (+/- SEM) of 2SG overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to [+ specific] ($N=38$) or [- specific] ($N=41$) pronouns.	130
3.4	Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to indicative ($N=2258$), subjunctive ($N=77$), or conditional ($N=41$) verbal mood.	131
3.5	Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to the switch reference factor ([+ continuity], $N=1267$, [- continuity], $N=1099$)	132
3.6	Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to whether the clause was [+ Negative] ($N= 292$) or [- Negative] ($N= 2074$).	133
3.7	Frequency of use (+/- SEM) of overt SPPs by first-generations speakers of Colombian Spanish in London, Ontario according to 1SG ($N=474$), 2SG ($N=50$), 3SG ($N=379$), 1PL ($N=186$), and 3PL ($N=150$) grammatical person.	137
3.8	Frequency of use (+/- SEM) of overt SPPs by first-generation speakers of Colombian Spanish in London, Ontario according to singular ($N= 897$) and plural ($N= 42$) SPP constructions.	137
3.9	Frequency of use (+/- SEM) of overt SPPs by first-generation speakers of Colombian Spanish in London, Ontario according to whether there was a continuous referent ([+ continuity], $N=712$), or a non-continuous referent ([- continuity], $N=527$).	138
3.10	Frequency of use (+/- SEM) of overt SPPs by second-generations speakers of Colombian Spanish in London, Ontario according to 1SG ($N=647$), 2SG ($N=29$), 3SG ($N=185$), 1PL ($N=129$), and 3PL ($N=137$) grammatical person.	144
3.11	Frequency of use (+/- SEM) of overt SPPs by second-generation speakers of Colombian Spanish in London, Ontario according to singular ($N= 860$) and plural ($N= 267$) pronoun forms.	145

3.12	Frequency of use (+/- SEM) of overt SPPs by second-generation Colombian Spanish speakers in London, Ontario according to indicative ($N=1079$), subjunctive ($N=22$), or conditional ($N=26$) verbal mood.	146
3.13	Frequency of use (+/- SEM) of overt SPPs by second-generation Colombian Spanish speakers in London, Ontario according to whether the interviews were in person ($N=446$), via telephone ($N=470$), or via video call ($N=211$).	147
3.14	Frequency of use (+/- SEM) of overt SPPs by second-generation speakers of Colombian Spanish in London, Ontario according to whether there was a continuous referent ([+ continuity]: $N=555$), or whether the pronoun represented a new referent or a change in referent in the discourse ([- continuity] $N=572$).	148
3.15	Frequency of use (+/- SEM) of overt SPPs by second-generation speakers of Colombian Spanish in London, Ontario according to whether the clause was [+ Negative] ($N= 957$) or [- Negative] ($N= 170$).	149
3.16	Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to in first ($N= 1239$) and second ($N= 1127$) generation speakers.	153
3.17	Frequency of use (+/- SEM) of pre-verbal SPPs by first ($N = 287$) and second ($N = 343$) generation Colombian speakers in London, Ontario.	154

List of Tables

1.1	Phonetic realizations of present indicative conjugations of the verb <i>cantar</i> (to sing) with and without /s/ deletion.	47
1.2	Present tense conjugations of the verb <i>cantar</i> ‘to sing’ in the indicative, conditional, and subjunctive moods.	55
2.1	Population numbers from Latin American groups ($N = 11070$) in London, Ontario according reported ethnicity (Statistics Canada 2017).	73
2.2	Demographic information regarding the twenty first (1G) and second (2G) generation immigrant speakers included in the CoSLO corpus.	75
2.3	Subject pronoun forms considered in the current analysis.	87
2.4	Present tense conjugations of the verb <i>cantar</i> ‘to sing’ in the indicative, conditional, and subjunctive moods.	88
2.5	Indicative conjugations of the verb <i>comer</i> (to eat) in the present, imperfect, and preterite tense.	90
2.6	Questions used to address factors contributing to language maintenance for qualitative thematic analysis.	121
3.1	Linguistic, social, and extra-linguistic factors conditioning the overt expression of Spanish SPPs among two generations of Colombian Spanish speakers living in London, Ontario ($N=2366$).	128
3.2	Linguistic and social factors conditioning the overt expression of Spanish SPPs among first-generation Colombian Spanish speakers living in London, Ontario.	135
3.3	Linguistic and social factors conditioning the overt expression of Spanish SPPs among second-generation Colombian Spanish speakers living in London, Ontario.	143
3.4	Linguistic and social factors conditioning the overt expression of Spanish SPPs among first and second-generation Colombian Spanish speakers living in London, Ontario.	152
3.5	Individual speaker data reported for all speakers in the first and second generation groups including number of tokens extracted from each interview, frequency of use of overt SPPs determined for each speaker, and the distance from the respective generation mean for each value.	156
3.6	Questions used to address factors contributing to language maintenance for the qualitative thematic analysis.	161
3.7	Themes identified through thematic analysis of conversations with first and second generation speakers regarding their use of Spanish at home, in social situations, in professional contexts, and in public.	163

3.8	Themes identified through thematic analysis of conversations with first and second generation speakers regarding their attitudes towards the importance of heritage language maintenance.	179
3.9	Themes identified through thematic analysis of conversations with first and second generation speakers regarding their attitudes towards endogamous and exogamous relationships.	185
3.10	Themes identified through thematic analysis of conversations with first and second generation speakers regarding their knowledge of Spanish-learning resources in the community.	189
3.11	Themes identified through thematic analysis of conversations with second generation speakers regarding their cultural identity.	194
3.12	Rbrul results for individual speaker data ($N = 2366$). Data presented in descending order according to FW values for overall overt SPP usage.	203
3.13	Quantified language use scores according to frequency of use of Spanish across four different domains (home/family, social-local, social-international, and work) as reported in language questionnaire for first generation ($N = 10$) and second generation ($N = 9$) speakers, and factor weight values for their use of overt SPPs. Asterisks (*) denote speakers in high intensity of contact for each generation group (see section 2.3.3).	204
4.1	Present tense conjugations of the verb <i>cantar</i> ‘to sing’ in the indicative, conditional, and subjunctive moods.	226

List of Appendices

Appendix A Western Research Ethics Board Approval	292
Appendix B Interview Guides	298
Appendix C Language Questionnaires	305

Chapter 1

Introduction

Despite the fact that in Canada, Spanish represents the fifth most commonly used immigrant language (Statistics Canada 2017), the use of Spanish, as well as the social and linguistic effects of the English-Spanish contact situation on the use of this minority language by Hispanic immigrant communities in the country, remains understudied (Guardado 2002; Hoffman 2001). This thesis contributes to this gap in the literature by investigating the variable expression of subject pronouns in two generations of Colombian Spanish speakers living in the Canadian city of London, Ontario ¹. This population is in intense contact with English in their everyday lives, and as such it is an excellent place to look at the possible effects of language contact between English and Spanish on Canadian soil. More specifically, I investigate effects of the situation of language contact on speakers' variable expression of Spanish subject personal pronouns (SPPs). In addition, I also conduct a qualitative analysis in order to gain a better understanding of the language use patterns and language attitudes of speakers which can help in predicting a possible future outlook for the maintenance of Spanish in Canada.

Although the situation of language contact between Spanish and English and specifically the variable expression of Spanish SPPs has been well studied in the United States (Sankoff and Poplack 1981; Silva-Corvalán 1994b; Flores-Ferrán 2004; Montrul 2004; Orozco 2004;

¹For the purposes of this thesis the focus will be exclusively in an English-majority part of Canada, particularly in the province of Ontario. (For Spanish in Québec, see Pato 2013-2017)

Sorace 2011; Otheguy and Zentella 2012), the socio-historical contexts of the language contact situation in Canada and in the United States are drastically different as will be made clear in the current chapter. These differences can result in different outcomes for the situation of contact, both in terms of the language contact effects observable in the use of Spanish by Hispanic populations in each country and in the long-term outcomes for the maintenance of the minority language (Sankoff 2001). For this reason, this work contributes not only to the literature on Spanish in Canada and on the study of minority language communities, but also to the literature on the variable expression of Spanish SPPs within a different situation of language contact.

1.1 General overview and research objectives

According to the Extended Projection Principle, the use of subjects in sentence structures is universally mandatory (Camacho 2013). However, each language will vary according to the type of subject pronouns that it allows. As a whole, subject pronouns can either be thematic or non-thematic, and they can be null or overtly expressed. In the case of Spanish, null and overt subject pronouns are both allowed, as can be seen in examples 1 and 2 below.

- (1) Nosotros comemos pan.
 1PL eat.1PL bread
 ‘We eat bread.’
- (2) Ø comemos pan.
 (1PL) eat.1PL bread
 ‘(We) eat bread.’

As seen above, in examples (1) and (2), sentences in Spanish with an explicit and with an omitted subject pronoun are both grammatical. However, as is usually the case with language variation, this aspect is not random. Variation between the use and omission of subject pronouns in Spanish is conditioned by a number of syntactic as well as discourse-pragmatic factors (Montrul 2004).

This contrasts with languages such as English, where the use of subject pronouns must

be explicit (3), while sentences such as those in (4) which omit explicit subject pronouns are ungrammatical when the same first person plural indicative reading is intended.

(3) We eat bread.

(4) *Ø eat bread.

In the current study, the principal objective is to investigate whether transfer effects due to contact between Spanish and English can be observed in the Spanish of Colombian speakers in London, Ontario. In the context of the United States, some studies have suggested that the situation of contact has led to signs of convergence between the two languages in the speech of Spanish-English bilinguals (Montrul 2004; Otheguy et al. 2007), particularly in the increased use of overt subject pronouns by second and third generation heritage speakers (Montrul 2004), and speakers who have been residing in a situation of contact with English in the United States for over five years (Otheguy et al. 2007). In contrast to studies suggesting that the observed frequency of subject pronominal expression is due to contact with English, another view suggests that the increased use of overt subject pronouns may in fact be caused by dialectal levelling effects due to contact with Caribbean varieties of Spanish. Caribbean Spanish shows some of the highest rates of use of overt subject pronouns in the Spanish-speaking world with an average of about 30%, in comparison to 16%-20% rates seen in Mainland varieties of Latin American Spanish (Cuza 2017). Since Caribbean Hispanics, including speakers from Cuba and Puerto Rico, constitute a well-established group in the United States, newcomer Hispanics may be assimilating to their language variety. Increased overt subject pronoun rates in heritage speakers may be reflecting this dialectal assimilation (Camacho 2016; Cuza 2017), rather than a syntactic convergence with English. It is this hypothesis that I primarily target through this thesis by looking at these research questions in a Canadian context:

1. What are the social and linguistic factors that condition the variable expression of Spanish SPPs in the Spanish of Colombian speakers, from the region of Cundinamarca, living in London, Ontario?

2. Are there differences between the first and second generation immigrant speakers of Spanish in regards to their variable use of subject personal pronouns? If so, how do the two generations differ from each other?
3. How does the variable expression of SPPs in the Spanish of London, Ontario (within a Canadian context) compare to findings reported for Hispanic communities in the United States?

In addition to investigating the linguistic aspects of the Spanish spoken by Colombian Spanish speakers in Canada, a secondary goal of the current thesis is to learn more about the experiences and the communities of these first and second generation Canadian Hispanics as they relate to the likelihood of Spanish language maintenance and inter-generational transmission. The main question that I target in this respect is:

4. How do the attitudes and language use habits of Colombian Spanish speakers in London, Ontario, align with factors that are known to influence language maintenance success for minority or heritage languages?

This last qualitative question investigates in particular aspects such as language attitudes of first and second generation Spanish speakers towards their own home language, the cultural identity of the second generation speakers, and the resources available to the Hispanic community of London, Ontario to contribute to language maintenance efforts. Although these last qualitative considerations do not represent the main focus of the current thesis, they are important in the larger scope of the investigation. They help us gain a better understanding of the social context in which Hispanics in Canada live, and they allow us to be more accurate in our predictions regarding the possible outcomes of the language contact situation between English and Spanish in Canada.

Within the scope of the current thesis, a comparison with similar studies conducted in the United States is also addressed. This allows us to better understand how the English-Spanish contact situation in Canada aligns with or differs from the English-Spanish contact situation

in the United States. This, however, is addressed mainly as a discussion of the results of the current investigation in order to situate our findings alongside those concerning Hispanic populations in the United States.

The current thesis is organized as follows. The introductory chapter is divided into four main sections: 1) Language contact; 2) Spanish-English language contact in North America; 3) The Null Subject Parameter (NSP); and 4) Spanish subject pronouns. Following this introductory chapter, chapter two discusses my methodological considerations and approaches. This chapter describes: the main research questions; the data collected, including presentation of the CoSLO corpus and of the community of speakers; the methods used for data collection, including the study materials and protocols; the transcription and token extraction procedures; all factors (social and linguistic) considered and the coding procedures used for each factor; and the quantitative and qualitative data analysis methods used. Chapter three is then designed to present results of the quantitative and qualitative data analyses to address the research questions (1-3) above regarding the variable subject pronoun expression of Spanish speakers in London, Ontario, and question (4) above regarding speakers' attitudes and use of Spanish correspondingly. This is then followed by a discussion chapter which mirrors the structure of the results chapter describing the significant results of the investigation as they relate to the four research questions above. Finally, the last chapter highlights the most important conclusions which are reached through my quantitative and qualitative analyses and discusses possible future research directions.

1.2 Language Contact

To begin this chapter, I give a general overview of the subject of language contact while focusing on the way that language contact influences bilingual and heritage language speakers. In the second part of this chapter, I also discuss the socio-historic aspects relating to language contact between Spanish and English in North America. This is essential for the

current project since a proper investigation on the effects of language contact necessarily considers extra-linguistic factors which influence the outcomes of language contact. Further, the discussion helps in highlighting some of the differences, and similarities, between the situation of English-Spanish contact in the United States and Canada. These social, demographic, and historical differences outlined mark the need for an investigation which specifically focuses on the variable expression of Spanish subject pronouns in Canada. It is important to note that the discussion outlined here on language contact is not meant to be an exhaustive summary, but rather an overview of this expansive topic focusing on those aspects that are most relevant to the current research.

1.2.1 Types of Language Contact

Contact between languages is a phenomenon that has been present throughout history in practically every part of the world (Thomason 2001). While sometimes we see a situation of language contact arising from voluntary actions by individuals, as is the case of migrant populations in search of opportunities, there are numerous ways through which languages can come in contact with each other. In fact, many factors such as politics, natural disasters, climate change, wars, religion, culture, education, economic landscape and even technology can come into play to bring languages into a situation of contact (Wei 2000). In linguistics, the study of language contact, which formally gained prominence following the seminal work of Weinreich (1968) as well as that of other pioneers such as Haugen (1953), forms an expansive field and encompasses a large range of situations.

In the current section, I provide a summary of a few main types of language contact situations, as described in the existing scientific literature, and some of the effects that have been observed to arise from these situations. However, as described by Weinreich (1968), the many extralinguistic factors surrounding a language situation will influence the observable effects of the situation of contact, and therefore it is important to remember that each situation of language contact is in fact unique and outlining a specific number of language contact types will

never be a precise endeavour.

The effects that ensue from language contact situations for individuals, societies, and the languages themselves can also vary greatly and differ according to each specific situation, even when the same languages are involved (Muysken and Appel 1987; Penny 2004). This may depend on socio-demographic, geographic, or temporal factors. Montrul (2012) notes that when analysing situations of language contact, many different factors need to be considered to get a complete picture of the situation. Factors such as the status of the language within the community (official vs. non-official; majority vs. minority, etc.), the prestige associated with the language (both overt and covert prestige, at the level of the community and at a global scale), and even the social standing of the speakers of the language within the community. These factors all play important roles in determining the outcomes and effects that can be observed as a result of the specific language contact situation. In addition to the consideration of societal factors, some linguists consider that it is the individual speakers themselves that are the locus of contact in situations of language contact and therefore it is this dimension that should be more closely considered when looking at the effects of language contact situations (Sankoff 2001; Valdés 2005). Further, it is sometimes difficult to determine which language-related changes or effects resulted in fact from language contact, and which could be attributed to other internal and/or social factors. Thomason (2001) explains that while some changes in the speech of a language community, such as lexical borrowings, can be very clearly attributed to language contact, others can be more difficult to determine. Nevertheless, linguistic research on this topic has made attempts at predicting the likelihood that specific effects from language contact will ensue from different types of contact. The following sections will speak to two broad types of language contact as described above, language contact at the society level and language contact at the individual level.

1.2.2 Contact at the level of society

At the level of society, Muysken and Appel (1987) outline three types of contact situations. First, they consider situations where within a society, one subgroup of the population is bilingual and another is monolingual. This is the case of diglossic communities where two languages are regularly used in a single community, but where the two languages are not equal in regards to social positioning and status in the community. In these situations, one language is a super-ordinate language and is used for formal and official purposes such as government, business, and education, while the other language is a subordinate language and is reserved for informal and familiar exchanges such as at home or with friends. Despite the fact that the super-ordinate language may not necessarily be the language spoken by a majority of the population, it generally enjoys higher prestige status and is considered socially dominant. Unfortunately, the social status attributed to each of the languages also tends to be transferred to the individuals in the society that speak that language themselves. Therefore, speakers of the socially dominant language tend to be considered of a higher social standing and speakers of the subordinate language(s) are considered of a lower social standing and are sometimes stigmatized along with their language. Consequently, the speakers of the subordinate language must have proficiency in both languages spoken in the community to be able to participate in everyday activities and are often bilingual, while speakers of the super-ordinate language do not need to speak the second language of the community and are likely to be monolingual. These situations can be observed, for instance in postcolonial societies, or societies with a recent immigrant population. It is important to note that one factor that contributes to intergenerational transmission and the survival of the subordinate language in the community relates to the covert prestige associated with this language. For example, in the case of some Creole communities or of speakers of AAVE in the United States, the ability to speak using the subordinate language marks a sense of belonging to the speakers.

Diglossic situations can further be categorized into two subtypes of language contact situations according to the level of prestige of the subordinate language from a global perspective.

For instance, while Spanish in the United States is a minority language which enjoys lesser social prestige than English, it remains a super-ordinate dominant language throughout Latin America and in Spain, and shares a similar level of prestige to English at a global scale ² This is not the case, by contrast, of the languages spoken by indigenous communities in the United States or Canada. Similarly to Spanish in the United States, languages such as Navajo in the United States, and Anishinaabemowin in Canada (to name only a couple) are minority languages. However, these indigenous languages are not spoken by a monolingual socially dominant population elsewhere and thus, even at a global scale, they remain mostly unknown and enjoy little social prestige. When speakers from these latter communities contract their use of their traditional languages favouring instead the use of the English (or French) majority language in more contexts, and if the rate of intergenerational transmission of the languages decreases leading to a community language shift towards English, these languages are likely to become endangered and may eventually die globally. By contrast, when the communities of Hispanics in Canada and in the United States undergo the same language shift towards use of English (or French) in all contexts, this results in the loss of the heritage language in these communities, but the Spanish language will neither become endangered nor die since there are many other communities around the world where Spanish is used daily, monolingually, as a majority language and where it continues to be passed one from generation to generation.

A second case of language contact described by Muysken and Appel (1987) are situations where two groups within a population speak different languages and live in close contact, but where only a handful of speakers are bilinguals. The few bilingual speakers facilitate communication between the two communities. This was the case for instance during the time of initial contact between European colonizers and indigenous communities, where only a few First Nation members were sufficiently proficient in the language of the colonizers and they served as interpreters between their communities and the Europeans. This situation is often of very short duration, and unless it is followed by colonization of one group by the other group

²Albeit not exactly the same since English has nowadays become a lingua franca world-wide.

leading to a more intense situation of contact, tends to have very few long-term effects on either language.

Finally, the third type of language contact situation according to Muysken and Appel (1987) are situations where members of a population are all bilingual or multilingual speakers. Examples of these situations can be seen in populations of many countries and regions in Africa and Asia where it is not uncommon to find speakers proficient in two or more languages. For instance, in India many speakers are often proficient in two or more languages such as Bengali, Hindi, Urdu, and English among others. Similarly, in Morocco, speakers will often speak both Berber and Moroccan Arabic, while they are educated in Modern Arabic and attend religious services offered in Classical Arabic (Wei 2000). This kind of situation is actually very common in many regions around the world. In fact, in most regions of the world it is more uncommon to find speakers who are strictly monolingual. This type of language contact where a large proportion of individuals is bilingual or multilingual can result in a situation of sustained language contact with multiple generations of bilingual speakers who routinely use two or more languages in different aspects of daily life. These situations may lead to language mixing and transfer effects.

It is notable, however, that this sustained state of bilingual or plurilingualism, and more importantly the variety of languages that bilingual speakers learn and are exposed to can become threatened by the prevalence of a “global” language, as is the case of English nowadays. For example, according to Kuzelewska (2016), in Switzerland’s region of Zurich, where the official language is German, it was decided in 1997 that English should be the first foreign language taught in schools, while minimizing the curriculum dedicated to Switzerland’s other official languages (French, Italian, and Romansh). This was received nationally with some controversy at the time but was accepted due to English’s status as the international language.

In contrast to Muysken and Appel (1987), Sankoff (2001) considers different types of language contact according to the ways in which the situation of contact began. Throughout history, languages have come in contact through different means, one of the most prevalent be-

ing language contact following conquest of one population by another. In this case, a socially dominant population conquers another population and imposes their language on the conquered population. Examples of this type of contact are ubiquitous throughout human history and include, as previously mentioned, conquest of Indigenous communities by European colonizers, as well as imposition of standard languages over local and regional language varieties as a move to encourage national unity. The latter is the case of many populations, including France, where standard (or Parisian) French was imposed over all populations of the country through the educational system hoping to see language shift in future generations to enhance patriotic and national unity feelings in the citizens of the country. Contact by conquest, explains Sankoff (2001), is more likely to result in various generations of bilingual speakers within the colonized population, and longer maintenance of the language of those colonized.

This situation differs from a situation of language contact by immigration where individuals who are speakers of a given language, for a variety of reasons, travel and establish themselves in a territory where a different language is spoken. Immigrant speakers arrive to the new community and in their attempt to quickly assimilate to their new environment, they may prioritize learning the language of the area and pass it on to their children to ensure their overall future success, and assimilation within the new community. Sankoff (2001) explains that language shift as an effect of this language contact situation is likely to occur within a shorter time frame than in the case of language contact by conquest. Usually two or three generations of speakers that are bilingual in the local dominant language and heritage language can be seen, but sometimes language shift can occur in as quickly as one generation.

The rate at which a population will experience language shift may also depend on whether it is a colonized or immigrant language since this may also affect whether the language is a true minority language or a subordinate language. In the case of language contact by conquest, it is possible that the language remains numerically strong, meaning that a large proportion of the population are native speakers of the language. In this case, it is possible to have a situation of diglossia lasting several generations aided by covert prestige of the colonized culture and

language. In the past, however, this has been affected by the violent means through which a population was colonized, or through the transmission of unfamiliar illnesses, which reduced the population numbers of the colonized populations, leading to their populations becoming smaller minority communities. In these cases, language shift and death is more likely to occur more rapidly. In the case of immigrant populations, their language is almost always a minority language in their new community. Due to this status, and the efforts of the immigrant communities to integrate themselves to their new environment, a diglossic situation is unlikely and quicker language shift is likely to occur unless language maintenance becomes an important goal for the minority community.

1 Community language varieties

Given a situation of sufficiently intense and sustained language contact we can see the rise of novel language varieties (Thomason 2001), which are unintelligible to monolingual speakers of the individual languages in contact themselves. At one extreme this is evidenced by the formation of pidgin and creole languages, but this can also be seen within other contexts such as in immigrant and other bilingual communities. Bilingual speakers in a given society can come together to form a speech community and a “community language variety” (Valdés 2005). This is most commonly seen in the cases resulting from immigration, but can also be seen in situations where linguistic minorities formed through other means such as by political developments (Sankoff 2001). Community language varieties are used mostly by heritage speakers of the minority language, and are characterized by phenomena such as code-switching, lexical nonce borrowings, as well as aspects attributed to attrition such as syntactic convergence, semantic expansion, and simplification (Silva-Corvalán 1994b; Thomason 2001; Valdés 2005). These features, as a result of the speakers’ bi- or multilingualism, can emerge at the individual level as a part of cognitive processing or during production of speech, and when adopted by the community at large, become characteristics of the community language variety (Treffers-Daller and Mougeon 2005; Valdés 2005).

Additionally, in some cases, bilingual speakers who are members of a minority language population, as is usually the case of heritage language speakers, can intentionally introduce certain distinctive language elements into their speech (Muysken 2013). In the case of heritage language speakers in a community where a different language functions as a majority language, they may insert lexical elements and discourse markers in utterances that are otherwise all in the majority language. This strategy, which Muysken denotes as backflagging, is mostly used by speakers from a second generation immigrants who are fluent in the majority language, and is done in an effort to highlight their ethnic identity.

Community language varieties can differ significantly from monolingual varieties of the languages in contact, and tend to be stigmatized by monolingual speakers of both languages (Valdés 2005; Otheguy and Stern 2011). Otheguy and Stern (2011) for instance note that in the United States, the overall use of the term “Spanglish” to refer to the way in which Hispanic individuals speak Spanish, whereby they use elements of English when speaking primarily in Spanish, or where their use of Spanish reflects community varieties’ use which differ from monolingual Spanish use, in fact can be hurtful. By referring to Hispanics in the U.S. as speakers of Spanglish, they are failing to recognize these individuals as speakers of Spanish, and their language use is thus stigmatized.

We reject the use of the term Spanglish because there is no objective justification for the term, and because it expresses an ideology of exceptionalism and scorn that actually deprives the North American Latino community of a major resource in this globalized world: mastery of a world language. (Otheguy and Stern 2011)

At one extreme of this situation, we can look at the emergence of creole languages. In many Caribbean islands, for instance, during the 16th and 17th centuries, European colonizers enslaved populations of First Nation groups, such as the Arawak and Karib nations, and populations which they captured and brought from West African nations. The result of these actions linguistically was an intense and sustained language contact situation between the European language of the colonizers (often English, French, Spanish, or Portuguese), First Nation

Caribbean languages, and multiple West African Languages. Enslaved individuals from different backgrounds were unable to understand each other or their colonizers. As a result, Pidgin languages, which contained elements from different languages in contact, emerged to facilitate communication between the community members. In the following generations, children who grew up in the midst of this situation eventually developed Creole languages which include features from European, Caribbean, and West African languages. Examples of these languages include Haitian Creole, Palenquero, Crucian, Jamaican Creole, among many others. While this situation is at an extreme, it is an example of how language contact can be at the root of new language varieties. However, it is important to note that in this situation of contact by conquest and by enslavement of the population, individuals were rarely able to develop full proficiency in the languages that surrounded them, with some exceptions in cases of Spanish colonies where some enslaved populations were in fact “allowed” to learn the colonizers’ language (McWhorter 1995), and at the level of the society a community language was needed to maintain day to day communication. This differs from true bilingual societies or populations, where we can see the formation of language varieties which differ from normative monolingual varieties and which may not be intelligible to monolingual speakers of either language variety as a result of transfer and attrition effects, but where speakers tend to have advanced proficiency in both of the languages in the community.

2 Heritage speakers and society

Heritage speaker is a term used to describe persons who grew up in a household where a minority language (which in this case we will describe simply as a language not used as an official language in the society in which they live) is used. More specifically, according to Montrul and Ionin (2012), heritage speakers are individuals who grew up in a household where at least one parent was a speaker of a minority language. From a social standpoint, heritage speakers grew up in families that not only used this minority language at home but which were part of the minority community. Each minority community shares not only language, but also other cul-

tural aspects. As a result, heritage speakers, even when the minority language proficiency is set aside, acquired throughout their childhood a knowledge of the corresponding culture, including music, traditions, food, religious beliefs, etc. This is important to note because these aspects contribute greatly to these individual's personal identity and group membership, which in turn, may in fact affect their use of language. One example of how identity can affect language is the intentional introduction of linguistic markers that signal these feelings and memberships. Intentional introduction of language variants among bilingual heritage speakers can be done in different ways but often takes shape of introduction of more frequent code-switching (to signal proficiency in both languages), borrowed lexical items, and tag words or expressions typical of one language, such as the use of Spanish *¡oye!* or *¡ay!* (Muysken 2013). This may be used as part of identity formation, according to constructivist identity theories, to indicate membership to a particular heritage group, or to the bilingual community.

Potowski (2014a) notes that in the case of U.S. bilinguals, individuals find ways of expressing their membership in the Latino or Hispanic groups to which they belong through the use of music and most notably language itself. The use of these linguistic strategies has been shown to be done only by these bilingual speakers when communicating with other bilingual speakers belonging to the same linguistic community and does not indicate a competence problem in either language system. For example, in Chicago, heritage Spanish speakers from Mexican and Puerto-Rican backgrounds co-exist and come together to form a linguistic community (Potowski 2014a). Students from these communities at the high school level use language features such as “reverse-Spanglish”, the use of English pronunciation while speaking in Spanish, as expression of their overall Latin@³ identity, or “Latinidad”, and bilingual identity. Additionally, speakers from each Hispanic group can also use differing Spanish lexicon in specific cases, (ie., Mexican word *popote*, or Puerto Rican word *sorbete* ‘straw’) to indicate their membership to the specific heritage group. Speakers in these groups often also take pride in having

³This notation is used to employ a more gender-inclusive (including both males and females) term than the more traditional term “Latino” which implies a primarily male term. It is important to note that other variations are also used such as “Latino and Latina”, “Latina/o”, and “Latinx” (Torres 2018).

full proficiency, (without audible accent), in both English and Spanish. In other words, these heritage speakers use three different language systems (English, Spanish, and heritage bilingualism) to indicate their full membership to each of these groups independently indicating multiple identities.

3 Language shift and language death

In the case of language contact at the community level, we have explored the factors that tend to be associated with the maintenance of a minority language in a community. However, should efforts to maintain the minority language fail, other consequences of language contact at the society level are language shift or language death.

Language shift refers to the cases where speakers of a minority language abandon the use of this language in favour of using a more dominant majority language. This is related to the notion of language attrition, but differs in that language shift refers to the situation where the use of the minority language in question disappears from the community as a whole (Pauwels 2016). It is also important to note that this may or may not also result in language death.

Language death accompanies language shift only if there are no other communities in the world who natively speak the language being abandoned by a given community in favour of another language. As described above, language death is more common in the case of situations where language contact came about as a result of an invasion or conquest, such as with First Nation languages in the United States and Canada. In the United States and in Canada, the shift of First Nation communities towards the use of English, while abandoning the use of traditional First Nation languages, will result in language death. However, in these same countries, shift of Hispanic heritage speakers from the use of Spanish towards English or French, while abandoning all use of Spanish, will not result in language death since many other countries and regions around the world use Spanish as a majority language.

Language shift in itself, whether it leads to language death or not, can occur suddenly or gradually depending on the situation of contact and the extralinguistic factors in play.

In the United States, Silva-Corvalán (1994a) has described how Hispanic speakers have begun to undergo language loss, whether it is due to first language attrition or incomplete language acquisition. This is evidenced by register contraction, and use of a simplified grammatical system. For instance, Spanish speakers in Los Angeles were observed to make less use of the subjunctive mood, favouring instead indicative forms, specially in cases where the standard norm allows for optionality between subjunctive and indicative forms (Silva-Corvalán 1994a). She explains that this change, at the level of society, has been occurring over the course of several generations, however, by the third generation of immigration, children from Hispanic ancestry may have very little exposure to Spanish at home (perhaps a little more if they live or are close with grandparents who acquired fluency in the language) and at best may develop an understanding proficiency. This gradual abandonment of the heritage language by Hispanic families may be indicated (at the level of individuals) over generations in the form of small evidence of transfer effects, syntactic convergence, and simplification of the verbal system.

In contrast, language shift can occur in the course of a single generation depending, again, on extralinguistic factors. Guardado (2010) noted that while investigating a Hispanic community from Chile in Vancouver, it was reported by parents of students that the use of Spanish in the larger community tended to be discouraged. However, he explains that among the most important factors in cultivating maintenance of the heritage language is the synchronous adoption of Hispanic cultural identity, which is in turn associated with language use. The combination of these extralinguistic factors may discourage even second generation Hispanic community members from using and practicing Spanish language outside of the home, and may in turn lead to a faster language shift to English in these populations. In fact, Guardado (2010) does report that parents of children in these communities were already struggling in helping their children to maintain their Spanish language use. This then may indicate that in Canada we will be seeing minority language loss and shift to English in as few as two generations.

1.2.3 Contact at the level of the individual

In the case of language contact at the individual level we can also speak of different kinds of contact such as language contact through preference and language contact due to necessity. In the first case, we refer to the case where a speaker chooses to learn a second language out of personal preference, for instance for employment or travel purposes. In a different situation, immigrant speakers for instance, are in a situation of language contact out of necessity. They must learn a second language after arriving to a new community to be able to participate in daily life. It is in the brain of bilingual individuals, during cognitive processing of language that the main point of contact occurs between two (or multiple) languages (Sankoff 2001; Valdés 2005). As these language systems interact, influence of one language over another may introduce linguistic variation which deviates from normative monolingual use in either of the two separate languages into the speech of an individual (Potowski 2014b). In the end, it is this individual variation, which through day-to-day communication can spread in use throughout the speech community leading to language variation at a larger scale and possible change if the introduced variables become more stable (Montrul and Ionin 2012).

At the level of the individual, the outcomes of language contact can vary greatly and what effects are observed depend on each individual situation (Thomason 2001, Treffers-Daller and Mougeon 2005). For the purposes of the current dissertation, I focus only on effects of language contact, at the level of the individual, which are common in a situation of immigration or of heritage speakers.

1 Language transfer effects

Transfer effects are generally thought to be the consequence of a situation of language loss where the speaker's second language becomes more cognitively dominant and some elements from this L2 begin to influence the use of the L1 (Montrul 2004).

Which domains of language are more vulnerable, if at all, to influence from transfer effects is a somewhat controversial topic. While Thomason and Kaufman (2001) argue that given

a sufficiently intense language contact situation, all domains of language become susceptible to influence due to transfer, other linguists consider that syntax proper is not susceptible to influence from transfer (Silva-Corvalán 1994b; Montrul 2004). Furthermore, it is important to note that while transfer is possible, simply having two languages in contact does not guarantee a language transfer effect will follow. According to Weinreich (1968), transfer effects are likely to occur only if there are parallel, competing structures between the two languages in contact. The author explains that within the bilingual brain, transfer occurs as a result of similar surface structures with underlying different analyses, which results in two competing forms. The bilingual individual then chooses between the two competing structures, but due to the ambiguity, it is often the case that the individual will opt for the form which allows for the least restrictions - that is, the form that is least cognitively demanding. The important element here, however, is that there needs to be two parallel structures which are in competition for a transfer effect to be possible (Weinreich 1968; Torres Cacoullos and Travis 2015).

One example is the case of lexical transfer evident through the misidentification of cognates. A bilingual individual associates different lexical items to representations of the given concept in the real world. That is, a Spanish-English bilingual may associate both 'dog' and *perro* with the concept of the four-legged domestic animal in the real world. However, depending on which of the two languages is more dominant for the specific individual, one of the two lexical items may be more strongly associated with the concepts and may be consequently more accessible. This situation can in some cases lead as well to lexical borrowing and code-switching as the person introduces within their speech the word from the other language when this is more easily accessible (Poplack 1988). In terms of misidentification of cognates, in the case of bilingual individuals, this same cognitive process may lead instead to the use of "false friends" when a lexical item in one language resembles the lexical item associated with the concept at hand in the other language (Jarvis 2009). To illustrate this more clearly, take for instance the English word 'embarrassed', which relates to a feeling of shame (more or less). When a Spanish-English bilingual is showing evidence of Spanish language attrition, they may

choose the word *embarazada* ‘pregnant’, which shares a close surface structure with the English word ‘embarrassed’ but which semantically differs greatly, in an inappropriate context in Spanish. This is due to the two similar, or parallel, surface structures entering in competition within the bilingual’s brain in relation. However, due to their differing underlying semantic content, this leads to a transfer error in speech (possibly a very confusing, or embarrassing one).

Importantly, although similar to these cases of lexical transfer, explains Silva-Corvalán (2008), it is more common to see lexico-semantic calques where the lexical items share some semantic content as well. This could then, in the case of heritage speakers and bilinguals in an immigrant setting, lead to semantic extension of certain items, which could in turn become crystallized in the community as part of the community language variety if their use spreads between speakers. It is mostly agreed, however, that not all domains of language are equally susceptible to transfer effects, and evidence from empirical studies does show patterns which suggest that the lexical and phonological domains of language are the first to be influenced by transfer effects in a situation of contact (Sankoff 2001; Thomason 2001), while morphology and syntax are less (if at all) vulnerable (Poplack 1988; Montrul 2011).

According to Silva-Corvalán (1994b) transfer from one language to another at the level of the individual cannot occur at the level of abstract syntactic structure. However, she explains, it is important to characterize how we define syntax itself first. While abstract syntax is defined as the level of syntax which includes basic syntactic structures with empty category slots, and thus not including any lexico-semantic, phonological, discourse or pragmatic elements, concrete syntax represents structures where lexical items have now been included and where semantic and discourse pragmatic considerations are at play. Therefore, due to the vulnerability of the other language domains to influence, transfer effects can be observed at the level of concrete syntax.

Similarly, Hulk and Müller (2000) suggest that it is at the syntax-pragmatics interface level, where language becomes susceptible to transfer effects, and not syntax proper. It is suggested

again, that while syntax on its own is not susceptible to transfer, during language production, discourse and pragmatic factors can then influence choices and create changes in regards to syntactic structure used by an individual. This is also seen in a study by Montrul (2004), where she studied the use of subject pronouns by heritage speakers of Spanish living in the United States. In this study, she found that while heritage speakers had a strong knowledge of the syntax of Spanish subjects, as evidenced by their acceptance and use of both null and overt subject pronouns, their speech deviated from monolinguals and violated to some extent discourse pragmatic conditions of null and overt subject use. This, she suggests, supports the notion that it is interface domains that are susceptible to transfer effects due to erosion or attrition of the first language. In addition, it is this erosion of pragmatic and semantic features that leads to the reduction of morphosyntactic complexity of the Spanish grammar, and which accompanies convergence with the grammar of English where pragmatic and semantic features do not constrain subject expression.

According to Silva-Corvalán (2008), a bilingual speaker continuously experiences heavy cognitive demands when switching between languages. This continuous use of the two languages, as described above, can lead to competition between parallel structures in the two languages leading in turn to transfer effects. As a result of this situation then, the two grammars in the bilingual individual may begin to show evidence of syntactic or grammatical convergence. This is hypothesized to be due to reasons of economy as it may lessen the cognitive demands on the individual when switching between languages.

Additionally, simplification can also be observed as a consequence of language contact and loss (Silva-Corvalán 1994b). In the case of simplification, we see the increased frequency of use of a form X in a given context, while gradually abandoning the use of a second form, Y, in a given context, when previously both forms had been actively used in the language and used according to differing constraints (Gutierrez 2003). This then results in a simplification of the language by the elimination of competing options that were previously used. Gutierrez (2003) illustrates the idea of simplification in the case of the use of the verbs *ser* and *estar* (to be)

in Spanish-English bilingual speakers in the United States. Traditionally, while both Spanish verbs *ser* and *estar* are translated as ‘to be’, they are used in different contexts by Spanish speakers. In this population of Mexican-Americans, however, speakers have expanded the use of *estar* to contexts where previously the verb *ser* was traditionally used (and is used in standard monolingual Spanish). Namely, according to Gutierrez (2003), bilingual speakers in this community have extended the meaning of *estar* to include that of “quality referring to a class”, where *ser* was previously used. This, he argues, can be due to a simplification of the system to reduce cognitive load on the bilingual brain, strengthened as well by convergence with the English system where only one verb is used for ‘to be’. However, it is still disputed in these cases whether the changes observed are due to the situation of contact or simply accelerated by it and would be part of the gradual evolution of the language on its own.

2 Heritage speakers and bilingualism

Heritage speakers are individuals who grew up in homes where a minority language, which differed from the official majority language of the community, was spoken. While these speakers grew up receiving aural input from the minority language at home and, at times, from members of their minority language community, most of the aural input they received by the time their language skills were fully developed (around the age of 12 years of age approximately) is likely to have come from the majority language.

Heritage speakers can vary greatly in regards to level of proficiency in the minority language. While some speakers may have native-like proficiency in both the minority and the majority language of the community, for some others their dominant language may be the majority language while having at best limited listening and understanding abilities in the minority language. Further, heritage speakers may or may not have been exposed to written materials in the minority language and may or may not be able to read and write in their heritage language.

Many studies in the United States have compared heritage speakers to second language (L2) learners. For instance, Montrul and Ionin (2012) studied one group of Spanish heritage

speakers and compared them against a group of L2 Spanish learners in regards to their interpretation and use of articles. The authors concluded that both groups showed similar transfer effects. However, although heritage speakers and second language speakers of a language may show similarities in the errors produced in their use of language (by normative standards), many differences between these two groups remain. Heritage speakers differ from second language speakers in their age of onset of acquisition, their mode of acquisition, and their direct experience with the language.

To begin with, as mentioned above, heritage speakers begin acquisition of the minority language at home. As a result, there is debate as to whether the transfer effects that are observed in their use of Spanish are due to attrition or to incomplete acquisition effects. Regardless of which of these two processes led to the observed transfer effects, these speakers are known to have begun acquisition at a much earlier age, usually before puberty and therefore before the age of critical period.⁴ This has been argued to have repercussions on the heritage speakers' language proficiency and their intuitions in this language (Montrul 2012).

In a study by Montrul (2012), the performance of heritage speakers, second language learners, and monolingual speakers of the same language when completing certain linguistic tasks was compared. Montrul (2012) explains that heritage speakers aligned more closely with monolingual speakers when given tasks that evaluated their implicit knowledge of the language even when their oral language proficiency more closely resembled the second language speakers' proficiency. This may indicate that even when heritage speakers do not show a high level of oral proficiency in their heritage language, their earlier exposure to the language can have lasting repercussions in their language abilities and may even have repercussions in their abilities to acquire native-like proficiency in the heritage language in the future.

One additional difference that distinguishes heritage speakers as a unique linguistic group lies in the mode of acquisition of the language. For heritage speakers, the initial acquisition of the language began at home, and most of their learning took place via oral communication

⁴The critical period hypothesis as it relates to the current study is discussed further in the methodology chapter of this work (see section 2.5.2.1).

with their parents, friends, and other close members of the community in informal contexts. This distinguishes them significantly from second language speakers who would have mostly learned the language explicitly via formal education. This may result in heritage speakers being less successful than L2 learners in tasks that evaluate explicit knowledge since most of their learning has been implicit in their minority language and they are unlikely to have received formal education in regards to their heritage language (Montrul 2012). The mode of acquisition experienced by heritage speakers also distinguishes them, however, from monolingual speakers. Since heritage speakers, even when they are advanced in their oral proficiency of the minority language, may have never or rarely had to use the language in formal contexts, this leads to register contraction, with heritage speakers having minimal knowledge of language uses (specialized terms, expressions, or forms of address) that are reserved for formal situations (Montrul and Ionin 2012). In addition, heritage speakers are often only exposed to community language varieties which may differ significantly from monolingual or standard varieties of the language.

In terms of education, or when assessing a speaker's proficiency, these characteristics of heritage speakers' language use (namely register contraction, and community language variety use) can pose challenges since what is perceived as "errors" in the speech of these speakers, when judged in terms of monolingual or standard norms, may in fact be characteristic of the unique community language varieties or simply a result of register contraction. In discussing education programs for heritage learners, it is important to keep in mind the unique characteristics of heritage speakers that set them apart from both second language and monolingual speakers. The mode of acquisition for heritage speakers may result in a lower level of meta-linguistic comprehension of the language, which may in turn result in them performing lower in meta-linguistic tasks, but they may also have a higher level of listening or speaking proficiency, as well as a more developed intuitive knowledge of the language and the culture associated with the language than second language learners. These are some reasons why it is often advocated to offer separate language programs for heritage speakers who want to increase their level of

language proficiency. These programs would then address the needs of heritage speakers such as expansion of register, focus on written language, and explicit instruction of standard language norms to allow them to better communicate with monolingual speakers and on a wider scale.

3 Language attrition and language loss

The outcomes of language contact at the level of the individual can include effects related to language loss, whether due to language attrition or incomplete acquisition.

Language attrition is defined as the process through which a bilingual individual begins to lose linguistic proficiency in their first language. It differs from incomplete acquisition in that it occurs in speakers who had previously completely acquired their first language. This situation can occur when an individual is immersed in a situation where they use their second language with more frequency and their use of the first language becomes more and more limited until the second language begins to become more dominant within the bilingual's brain. This is the case often seen in situations of immigration where individuals begin to have more and more restricted contexts in which they use their first language. In this context, it can be accelerated by exogamous marriages or relationships, which further lead to fewer situations where the minority language can be used and where the minority language loses prominence in the home domain (Pauwels 2016). Language loss via attrition can also be characterized by a number of language transfer effects including syntactic convergence, register contraction, and/or simplification of the verbal paradigm (Silva-Corvalán 1994b).

One example of simplification of the verbal paradigm is observed when looking at the population of Spanish speakers in Los Angeles, where second and third generation heritage speakers make less use of the subjunctive mood, favouring instead indicative forms, especially in cases where the standard norm allows for optionality between subjunctive and indicative forms (Silva-Corvalán 1994a). The author explains that this change has been occurring over the course of several generations. She also explains that interestingly the simplification of the

verbal paradigm is a mirror image of the process observed in creole language complexification. That is, while in creole formation it is conditional and subjunctive mood forms that develop last (and that are less common among creole languages), these are the first items to be lost in speakers undergoing language loss (Silva-Corvalán 1994a).

1.2.4 Minority language maintenance

Minority languages are by definition languages that are used by a community that is a minority within a larger population. There are, however, different kinds of minority languages. The term of minority languages and minority communities is frequently associated with immigration and the term may be considered to be synonymous with heritage languages, such as is the case of Spanish in Canada or of German and Dutch in the United States. As described previously, these are languages that are brought into a nation as populations immigrate (for a variety of social, economic, environmental, political, or historic reasons) and that are used for the most part at home. These minority immigrant languages are rarely considered official languages of the new nation, and the use of these languages remains for the most part restricted to the home domain and only within the immigrant community.

These immigrant languages differ from cases where minority languages are territorial languages. In this case, the minority languages are native or indigenous to the territory where they exist, but are still only spoken by a minority of the population. In these cases, it is due to social, political or historical reasons that they are now the language of a minority. Territorial minority languages may have an official status in the country. This is the case of languages such as Catalan in Spain, and of French in Canada which are minority languages at a national level (although they may not be minority languages within their specific territories). These in turn, in my opinion, further differ from the case of non-standardized (often non-official) language varieties such as African American Vernacular English (AAVE) in the United States and Creole languages throughout the world, which are minority languages with no standardized writing system. I make this distinction because in the case of these non-standard languages,

it is sometimes the case that even the speakers themselves do not recognize the languages as distinct systems separate from the standard varieties of the dominant standard variety of the language (in the case of AAVE)⁵, or their lexifier languages (in the case of Creole languages) and therefore, the speakers themselves may not recognize themselves as speakers of a minority language variety.

Despite the different kinds of minority languages, in minority language studies, one main point of interest remains determining the reasons or the factors that can help in predicting whether a minority population will be successful at maintaining the use of their language (Linton 2004; Pauwels 2016). While some studies have focused on investigating internal factors such as the quantity and quality of input received by speakers while learning the minority language (Potowski 2014a), others have focused on external factors such as the community's attitude towards the minority language, and even the government and educational policies which may support or hinder the minority language use and transmission (Linton 2004). In this section, I discuss first the factors at the macro-level, that is at the level of the community, and then the factors at the micro-level, that is at the level of the individual or of the family, that influence the success of a given minority group in maintaining the use of the minority language in the community.

It is important to note, however, that an all-encompassing description of factors that predicts language maintenance does not yet exist, and given the immense variation between different situations of language contact, findings in one community don't always apply to another community. As Pauwels (2016) describes, the development of a predictive model is still ongoing, and therefore, in this section I merely note some of the factors most commonly associated, and most frequently associated with successful language maintenance.

⁵According to Rickford (1998), it has been suggested that AAVE may have a Creole origin, but this is a debated topic and its discussion lies outside of the scope of the current thesis.

1 Macro-level factors

The attitudes that the majority and the minority language communities themselves hold towards the use of the minority language in question can greatly impact the likelihood that a given language is maintained.

Attitudes of speakers towards the use of their minority language may be linked to their cultural identity and in these cases, it is more likely that an effort will be made to use and maintain the minority language as young members of the minority community build their identity. If the link between cultural identity and language is weak, however, the pressure to learn and use the minority language is reduced and therefore speakers are less likely to emphasize this as a priority. This is explained by Pauwels (2016) as the notion of “core value”. A core value for a given community refers to any specified trait or characteristic that is intimately connected to the group’s sense of identity. If a group is at risk of losing any of the core values, the identity of the group can be considered to be threatened. Therefore, in relation to language use, if the use of a minority language is considered a core value for the community, then the rejection or inability to use it could lead to an individual being excluded from the group.

In the case of Hispanic communities in the United States, language isn’t always considered a core value for cultural identity and this may be a risk factor contributing to language shift, as observed, over two or three generations (Potowski 2014a). For instance, in New York, for the Puerto Rican community, the ability to speak Spanish is not connected to the ability to be Puerto Rican. According to Pauwels (2016), in fact, many children of Puerto Rican descent in the community are monolingual speakers of English and still actively participate in the Puerto Rican cultural traditions and maintain a Puerto Rican identity. This, combined with the English majority environment by which they are surrounded contributes to the lesser intergenerational transmission, as you can be Puerto Rican and maintain your cultural and ethnic heritage ties without having advanced Spanish proficiency.

In addition, a minority language will struggle to be maintained, even in cases where parents, family, and even educational resources have been committed to language transmission, if the

local teenage and social networks of young speakers of the minority culture have negative views regarding the use of the minority language. This is due to the fact that after the age of about 10, most speakers' influences regarding their language use shift from the home and educational domain to their social network (Potowski 2014a).

In yet more extreme cases, it is possible that the minority language is not just discouraged or used infrequently, but that a true stigmatization of its use and its speakers is present in the larger community. For example, returning to the case of Creole languages, these languages are associated with their historical roots which lie in slavery. Since these languages emerged from a situation of slavery, and the populations enslaved, who were speakers of these languages, were strongly stigmatized even after the end of slavery, these languages were (and still are in many cases) referred to as “bad/broken English” or “Français cassé” (Bartens 2001; Siegel 2007; Wigglesworth et al. 2013). For this reason, despite the fact that these languages have become a mark of cultural heritage for their speakers, and the fact that some speakers of these languages today consider them “the heart of their home” (Plyley and Hernandez 2019), older populations in some creole communities insist that their children avoid the use of these languages (Bartens 2001; Wigglesworth et al. 2013), in hope that this will give them better hopes of economic, social success, and upward social mobility. This leads to a shift towards the majority language in the region (usually towards English or French), or decreolization in these particular cases (Valdman 1989). One consideration, however, that keeps these languages alive and in use is the cultural association to the languages and the covert prestige that they hold within the community. In the island of St. Croix in the United States Virgin Islands, speaking Crucian creole, for instance, marks you as a member of the inner community and separates you from those from “babylon”, a name used to refer locally to the contiguous U.S. Plyley and Hernandez (2019).

Another important factor in determining the probability that a minority language will be maintained is the number of individuals in the given region who speak and use it. According to Linton (2004), the number or local concentration and distribution of minority language

speakers in a given area also can be connected to a critical mass model. These critical mass models mainly describe the notion that there is a tipping point situation where once there is a certain number of minority speakers together in a given area, there is a subsequent increase in the number of individuals who are motivated to maintain their bilingualism. Linton (2004), whose study focuses mainly on Hispanic immigrant populations in the U.S., describes that for individuals arriving to a new country where their language is a minority language, there are a number of contextual factors that will influence the decision on which languages to use in different contexts. This decision, he explains, is not unlike a cost-benefit analysis where the use and maintenance of the language is looked at in terms of how beneficial and desirable it is vs. how much effort and what costs are associated with it. Therefore, when an immigrant individual arrives in an area where the minority language is widely used and where resources are available in that language, the effort to maintain the language (and to pass it on to the next generation) may be low, especially if educational resources have been developed by the existing community, and the benefits can be high as it allows for social interactions with other members of the community who likely also value the use of the minority language and by extension bilingualism. On the contrary, if an individual arrives in an area where the minority language is not frequently used (and may even be discouraged or stigmatized) and where resources are not easily available, the cost of maintaining the minority language may be high enough to outweigh the costs. This would in turn be even harder if the community itself encourages or expects a quick assimilation by the incoming immigrant populations and if bilingualism in general is discouraged by education and other government policies.

One example of this contrast is described by Potowski (2014a), who mentions that a higher success rate of Spanish maintenance has been observed in the region of Miami Dade, where nearly half of the population is reported to “use Spanish at home” according to the 2013 census, than in New York where the proportion of the population who is Hispanic isn’t quite as high. While in Miami it is not uncommon to find third and sometimes fourth generation speakers of Spanish, in New York it is more likely to see a shift towards English within two generations.

The size of a minority language population isn't the only factor that can influence language maintenance in a community. Presence of linguistic enclaves, where minority language speakers may not be as numerous but where they remain closely concentrated geographically and where contact with other groups is minimal, can also contribute to the maintenance of the language. This is the case, for instance, of Amish Dutch populations, as well as some Yiddish speakers in the U.S. who have maintained their languages over multiple generations due in part to the fact that they have remained relatively isolated from the dominant language community.

It is also important to highlight that larger or more closed off communities of minority language speakers may also correlate with a higher likelihood that the community may work to create further opportunities for the future generations to practice and learn (even become literate) in the minority language - that is, as long as there are positive attitudes and a strong desire from the community to maintain the language. According to Pauwels (2016), this also contributes strongly to the language maintenance efforts. For instance, communities may develop schools or other educational initiatives aimed at helping the future generations to learn and use the language. These programs, however, do even more than that. They provide a center where youth from the community can come together, share their negative and positive experiences with the minority language and culture, and even begin to negotiate their cultural and heritage identity. This is a major issue among later generations of heritage speakers since they are often faced with dual identities which are conflicting. In fact, according to Pauwels (2016) there is inconclusive evidence that these schools are effective in maintaining the minority language, but the evidence does suggest that these community-based spaces "provide a safe space for students to forge identities linked to the heritage or minority language and to experiment with various linguistic practices" (Pauwels 2016: 133) which may reflect a bilingual or dual identity. That being said, the availability or possibility to have these spaces is in many cases dependent on local and national policies and support.

Language policies can vary widely from one jurisdiction to another and the particular positions adopted in regard to minority language maintenance can have strong repercussions. For

instance, Pauwels (2016) explains that following both World Wars, there was strong resistance in Australia against allowing the maintenance of German in the country. This represents one of many situations when a nation has taken a strong stance forbidding the use of other non-official languages, often hoping instead to encourage assimilation of the minority communities into the dominant culture. In these cases, any efforts by the minority community members to maintain their language or to instruct the next generation on its use would have to be done in secret, which would limit the access to these programs. In the case of French in Ontario, similar situations arose in the 1800's where efforts were made to impose English as the sole language of instruction. At the time, apparently tolerant policies which allowed for the inclusion of French in the classroom were seen as necessary to avoid damaging relationships with the Francophone population, but they were expected to be "ephemeral" strategies since government members such as Egerton Ryerson expected a natural shift over generations and a "voluntary assimilation" of Francophones towards the use of English language and "British" traditions (Gaffield 1987).

Another example of policies adopted by countries that may negatively impact the maintenance of a minority language is found in the United States where the educational policies implemented in several states encouraged an English Only approach to education as early as 1919, following the first World War. The English Only educational policy was born initially of a fear of national loyalty issues and was established with the hope that one language throughout the country might unify the population under a single national identity (Linton 2004), and seeks to establish American English as the official language of the United States. Linton (2004) explains, however, that there are other models that have been proposed and supported in the United States. For instance, the use of an English-Plus approach where the goal would be to encourage bilingualism with most American citizens speaking English and one or more other languages. This was considered to be in the best national interest to improve international relationships and to help maintain a culturally rich and diverse population. This approach, however, has been less successful in practice. For instance, in 1968 a Bilingual Education Pro-

gram was established in the U.S., but these programs remain “transitional”, maintaining the main goal of helping children to quickly transition to English use. In addition, these programs have been critiqued as they are described as positioning bilingualism as a disadvantage rather than an advantage, and to promote isolation of bilingual children (Linton 2004).

In other cases, some minority languages may be given an official status and the government may provide some financial help to provide these minority communities with resources in their own languages. This is the case of Canada, for example, where an approach of official bilingualism was adopted. Today, following the Official Languages Act (1969), Canada has both English and French as official languages and according to this act, speakers of either official language have the right to access services and information in their own language. This has been enforced through different acts and regulations such as the French Language Services Act (1986), which was instituted in Ontario to protect the language rights of Franco-Ontarians.

2 Micro-level factors

In order to maintain and foster intergenerational transmission of a minority language, the macro-level factors discussed in the previous section greatly contribute to a minority population’s chances of success in maintaining their language. However, there are also significant factors at the level of the individual, and even with the community-level (or macro-level) factors being favourable, if the levels at the level of the individual (or micro-levels) are unfavourable, the success of a given individual or family in this respect might be affected. Within the context of the individual I include factors based on personal cognitive and motivational aspects of language acquisition, but also social factors as they relate to the family and the home. More specifically, in this section, I discuss the importance of the quantity and quality of the input received by minority language speakers during language acquisition (and beyond), and the motivation and sources of motivation for maintaining the heritage language.

One of the greatest challenges in raising a bilingual child in a minority context is ensuring that the child receives a sufficient amount of language input from the minority language. The

quantity of input during L1 acquisition, correlates with the rate of acquisition. For instance, children that receive a higher quantity of input will learn more vocabulary, and will more quickly develop certain aspects of the language's grammatical system such as tense, aspect, and noun gender knowledge (Ramirez 2013). However, ensuring children receive enough input in minority contexts becomes even more difficult once the child enters the formal education system at around the age of four. In fact, in minority contexts, for instance where the language is an immigrant or heritage language, it is common to see incomplete acquisition of the minority home language in both simultaneous and sequential bilingual children (Pérez-Leroux et al. 2011). This, although attributable to a combination of social, cognitive, and linguistic factors, could be in large part due to the reduced amount of input and the fact that the child then spends a considerable amount of time receiving input only from the majority language as they begin to develop social relationships (e.g., with teachers and schoolmates) where the minority language is not used and perhaps even discouraged.

Another situation that can influence the amount of input that children receive from a minority language is the nature of the family or home context. For instance younger siblings in a family tend to receive less input in their minority L1 as they interact with their older siblings who often use the majority language during play time. Another challenge to language maintenance can be exogamy, or marriage between members of a minority language community and members of the majority language community (Pauwels 2016). In this situations the majority language enters the home domain and further reduces the amount of input that the child receives from the minority language. In addition, it may also influence the cultural ties that a child feels towards the minority culture, which in turn may affect their motivation for maintaining the use of the language. That being said, however, it is important to note that exogamous relationships do not on their own doom the probability of intergenerational transmission and can in fact be beneficial if an environment encouraging bilingualism and multiculturalism can be created in the home (Pauwels 2016). In fact, the attitudes of parents towards the minority language, the minority community, and bilingualism itself are essential predictors of intergenerational

transmission.

Guardado (2008) studied the experiences and attitudes of Hispanic families in Vancouver, Canada, and determined that two factors were essential in determining the success of intergenerational transmission: cultural awareness, and *familismo*. He explains that within the home domain, parents with positive and encouraging attitudes towards bilingualism and the minority language culture in general were more successful in maintaining the use of Spanish. In addition, he determined that the concept of *familismo*, or loyalty and strong affiliation with the family can also serve as strong motivation to maintain the use of Spanish. For instance, children who are internally motivated to maintain ties with their Hispanic family members in their country of origin are more likely to attribute more value to the use of Spanish and thus to aim to maintain it. This factor was more significant in determining success of intergenerational transmission than even perceived importance of the use of the minority language for practical purposes such as job opportunities. In addition, according to Pérez-Leroux et al. (2011), the attitudes of Hispanic parents in Toronto are also important predictors in the maintenance of Spanish, especially when considering the perceived ethnolinguistic vitality of the minority language. Families that considered that the minority language was more alive in Toronto, with more opportunities to use it for social interactions, were more likely to consider language maintenance a priority.

1.3 Spanish-English Language Contact in North America

Contact between English and Spanish in North America can be traced back to the 1800's following arrival of English and Spanish colonizers to the Americas. Notably, intense contact initiated at the border between Mexico and the United States followed by the treaty of Guadalupe Hidalgo in 1848 which resulted in Mexico being forced to concede over half of its territory to the United States (Lozano 2018). The land conceded, nowadays compose all or significant parts of the states of Arizona, California, Colorado, Nevada, New Mexico, Texas,

Utah, and Wyoming. As part of the treaty, any Mexican citizens who lived in these lands previous to the treaty was then given U.S. citizenship. This led to a long intense contact situation between English and Spanish in the United States that persists to this day.

1.3.1 Spanish-English contact in the United States

In addition to the 1848 acquisition of Mexican territory (Lozano 2018), in 1898 following the Spanish-American war, the United States acquired the island of Puerto Rico, an island that was and remains home to a population of Caribbean Spanish speakers. Residents of the island, similarly to all inhabitants of the previously Mexican territories, were given U.S. citizenship. The inclusion of Puerto Rico further increased the presence of Spanish in the United States, thus intensifying the situation of contact between Spanish and English in the United States, especially considering that many citizens from Puerto Rico immigrate into other areas of the United States, such as New York, Florida, and Chicago, in search of financial and professional opportunities. It is important to note as well that this not only increased language contact between Spanish and English, but since Mexican and Puerto Rican Spanish constitute two different varieties of Spanish, this also led to dialect contact.

Since these initial points of contact, many generations of Spanish-English bilingual speakers have remained in the United States and the current population of Hispanics in the country constitutes an important portion of the overall population (Silva-Corvalán 2008). In fact, according to data from the 2016 census, there were at the time of the census 57.7 million speakers who reported speaking Spanish as their native language. This makes Spanish the second most used language in the United States after English. Further, to this day, language contact in the United States between Spanish and English, and dialectal contact between different varieties of Spanish, continues and intensifies as Hispanic groups from a variety of countries including most notably (but not exclusively) Cuba, El Salvador, Dominican Republic, and Colombia, continue to immigrate. This constant immigration from monolingual Spanish speaking countries helps in the revitalization of Spanish language use, which assists in maintaining the

language in communities where otherwise we would be likely to see language shift towards English (Pauwels 2016; Silva-Corvalán 2008).

Throughout history, attempts have been made by government and education officials of the country to unify all of its citizens under one language, English. This has brought about complex social relationships between English and Spanish populations, specially since Spanish in the United States does not enjoy an equal level of overt prestige as English. In fact, according to Lozano (2018), during the early part of the 20th century, when considering whether or not to give statehood to the territories of New Mexico and Puerto Rico, the one condition presented was to incorporate English-only education, starting from primary school, in the regions. While New Mexico conceded to the terms and gained its statehood, Puerto Rico declined, opting instead to maintain local education in Spanish, and was therefore denied statehood. The main argument for this decision was that it was considered that persons who were not proficient in English were unable to fully participate in government processes in the United States due to lack of understanding.

The complex social and political situation surrounding Spanish-English contact in the United States has also given rise to stigmatization of Spanish populations, and covert pride of Latin@ culture among Spanish speakers. This in turn comes accompanied by the use of linguistic elements to mark Latin@ identity and pride. For example, when speaking with other bilingual speakers, individuals will make frequent use of code-switching, and/or insert Spanish discourse markers in an otherwise monolingual English conversation. This is considered to be done as a marker of identity since speakers who engage in these linguistic practices can avoid using them when speaking to monolingual speakers of either Spanish or English. In fact, some linguists have expressed that Spanish-English bilinguals are in fact speakers of three linguistic systems, one being Spanglish (Otheguy and Stern 2011). The frequency of use of this term in itself, both in linguistic literature and in popular culture, highlights the prevalence of this community language variety due to the intensity of language contact, and the overall strength of Hispanic culture in the United States.

Given the intense situation of Spanish-English contact in the United States, it is no surprise that there is great academic interest in studying it from linguistic, social, and historical standpoints. Linguistically, some of the more commonly studied phenomena include code-switching, simplification of the verbal system, syntactic convergence, and borrowing (Poplack 1988; Silva-Corvalán 1994b, 2008; Otheguy et al. 2007).

1.3.2 Spanish-English contact in Canada

In Canada, the contact between Spanish and English is a much more recent phenomenon. By contrast with the United States, where examples of both contact by conquest and contact by immigration have been recorded since the mid 19th century, in Canada contact between Spanish and English only occurs through contact by immigration as several waves of Spanish-speaking citizens have arrived since the mid-20th century (Bonnici and Bayley 2010; Pato 2013-2017). According to Bonnici and Bayley (2010), the first wave of Spanish speakers in Canada immigrated during the mid-20th century as a result of situations of political unrest in Spain and Latin American countries such as Chile, Argentina, and El Salvador. Following this initial wave, other immigrants arrived from countries like Colombia during the early part of the 21st century. Nowadays, in Canada, according to the 2016 national census, 495 090 individuals reported Spanish as their native language which is still used at home, making Spanish the fifth most widely spoken immigrant language after Mandarin, Cantonese, Punjabi, and Tagalog (Statistics Canada 2017).

Despite this, however, and partly due to the somewhat recent arrival of Hispanic populations to Canada, most of the linguistic literature that exists regarding situations of Spanish-English contact in North America focuses on the United States, and literature in a Canadian context remains scarce (Guardado 2002; Hoffman 2001; Bonnici and Bayley 2010; Pérez-Leroux et al. 2011; Pato 2013-2017).

The prevalence of Spanish as an immigrant language in Canada has been increasing over the past few decades. Nevertheless, Hispanic immigrants who come to Canada find themselves

in a social situation that is very different from the one in the United States. In contrast with the United States where Hispanic communities have been established for centuries, and where official and commercial services can be accessed through the use of Spanish, (which in my opinion constitutes a veritable second language - unofficially so - in the United States), in Canada these same services will only be accessible in English or French according to the region of the country in which Spanish immigrants find themselves. This results in conditions that are much less favourable for passing Spanish on to future generations and decreases the likelihood of maintaining its use. This also means that for adult Hispanic immigrants the pressure to gain a high proficiency level in English and/or French is greater than in the United States since without it they will be mostly unable to access necessary services and participate in everyday life.

Guardado (2008) conducted a qualitative study to investigate the attitudes of Hispanic parents in Vancouver towards the use of Spanish at home. He notes that parents have an overall great interest in their children maintaining the use of Spanish, and consider that growing up with this bilingualism grants their children a sense of pan-ethnicity, or a broader global identity. Therefore, from his study he concluded that Hispanic parents in Canada do strive to cultivate the use of Spanish as a heritage language at home. He also found, however, that the Hispanic families reported experiencing great pressure to use English both directly and indirectly in order to succeed in Canadian society. Some families even reported pressure from school officials for the abandonment of Spanish, even when at home. It was believed that it was in the children's best interest to gain full English proficiency as quickly as possible to succeed in school and to be integrated into the school's social environment. This is supported by Linton (2004) who explains that while living in Canada, unlike in the United States, speaking either French or English is an absolute necessity for navigating everyday life, accessing resources, and advancing professionally. This kind of heightened pressure can negatively impact acquisition of the heritage language since children receive less, and sometimes insufficient, input in Spanish, and have very few opportunities to use it on a daily basis (Valdés 2005).

Among the studies conducted regarding the language contact situation between English and Spanish in Canada, we find quantitative and qualitative studies focusing on both social and linguistic aspects. Hoffman (2001), for instance, studied a population of Spanish speakers from El Salvador residing in Toronto, and focused on investigating the frequency of aspiration and deletion of /s/ in coda position by second generation speakers of Spanish from El Salvador vs. monolingual speakers from the same country. Although a common phenomenon in El Salvador, /s/ aspiration and deletion is a feature with low overall prestige in monolingual varieties of Spanish. In Toronto, however, Hoffman (2001) found that speakers from El Salvador were making use of this feature in their speech as a marker of identity. This was found despite the prevalence of Hispanic speakers from other regions of Latin America, such as Colombians, where the use of /s/ aspiration and deletion is considered a low prestige variant.

It is also important to note that as Canada has two official languages, French and English, the immigration of Hispanic populations to Canada has led not only to English-Spanish contact, but also to French-Spanish contact. Pato (2013-2017) studied French-Spanish contact in Québec by collecting the corpus of Spanish in Montreal (*Corpus Oral de la Lengua Española en Montréal*; COLEM). Subsequently, through the use of this corpus, Cruz Enríquez (2014) examined the morphosyntactic effects of language contact in Spanish-French bilinguals, but this discussion lies beyond the scope of this investigation.

Given the prevalence of Hispanic communities in Canada and the fact that Spanish represents, as of the year 2019, the fifth most commonly spoken immigrant language in the country, the research that exists on the use of Spanish in Canada remains insufficient (Hoffman 2001; Guardado 2008; Pato 2013-2017; Pérez-Leroux et al. 2011). Through this research project, I seek to contribute to the growing body of literature on this subject, by looking into the effects of the English-Spanish contact situation on the use of variable expression of Spanish subject pronouns.

In the next section, I transition into a discussion of the relevant literature on the Null Subject Parameter (NSP), which is in turn followed by a section which more specifically focuses on

Spanish subject pronouns.

1.4 The Null-Subject Parameter

According to Chomsky's (1982) government/binding theory, each person is born with a set of language features, or principles, which are common to all natural languages. Each child will then set certain parameters within each principle according to the input they receive from the language(s) they are exposed to from birth. Therefore, while principles are universal and invariant, parameters vary according to each language (Camacho 2016). In regards to subject pronouns, it is hypothesized that all natural languages abide by the extended projection principle (EPP), which states that all tensed clauses, in all natural languages, require the use of a subject (Chomsky 1982; Camacho 2013). However, the EPP contains as well the Null Subject Parameter (NSP), which can be set for each language to either allow or disallow for the use of null subjects (Camacho 2013).

1.4.1 Null and non-null subject languages

Some natural languages allow for the insertion of a null *pro*, or phonetically empty, subject into a syntactic structure when forming a sentence. These languages, known as pro-drop or null-subject languages (NSLs), according to Government and Binding Theory are thought to be languages for which the NSP has been set to [+ null subject] due to the language input that an individual received during language acquisition of their native language (L1) (Chomsky 1982).

The possibility of having null subjects in a language has been argued to be due to the rich morphology of these languages. For example, languages such as Spanish or Italian, both of which are languages that allow for the use of null subjects, have rich morphology within their verbal paradigms. It is argued that the rich verbal morphology in these languages allows for the recovery of the semantic content of the subject of the clause even when the subject itself

is phonetically absent. For instance, if we consider the example in (5) below, we can see that although the null *pro* subject has been inserted into the phrase, the verb *habla-mos* (Speak-1PL) contains sufficient information in its form that we can easily determine the subject of the clause is ‘we’, the first person plural.

- (5) \emptyset hablamos en la tarde.
pro speak.1PL in the afternoon
 “(We) speak in the afternoon.”

The same, in contrast, is argued to not be possible for non-null subject languages which have a more limited verbal morphology from which the subject cannot be recovered. For instance, if we consider the example below in (6) from English, we see that the verb form *speak*, in contrast to the Spanish form above, does not allow us to determine the subject of the clause and inclusion of the null subject renders the phrase ungrammatical with the intended interpretation.

- (6) * \emptyset speak in the afternoon.
pro speak.1PL in the afternoon

It is important to note that in general that the use of a null subject in null subject languages is optional, meaning that most clauses could include an overt or a null subject pronoun without affecting the meaning or grammaticality of the clause. Languages that are considered to be null-subject languages include Italian, Spanish, Portuguese, Russian, and Greek among others (Nagy 2015; Camacho 2013).

The activation of the NSP is also thought to activate other characteristics, which are therefore expected to be shared by all [+null subject] languages but not by [-null subject] languages. These characteristics include the ability to have subject inversion, the ability to extract *wh*-subjects from an embedded clause (D’Alessandro 2015), the inability to have overt expletive pronouns, and the unavailability of overt pronouns with arbitrary reading (Camacho 2008).

NSL’s allow for the inversion of subject placement within the syntactic structure. For instance, a subject in a clause from a [+ null subject] language may be found in pre-verbal subject position in a phrase (7), or in post-verbal position (8).

- (7) Sofia come pan.
Sofia eat.3SG bread
“Sophia eats bread.”
- (8) Come pan Sofia.
eat.3SG bread Sofia
“Sofia eats bread.”

This flexibility is not displayed by [-null subject] languages, however, where a subject position is more rigid. For instance in English, which is not a null-subject language, there is a preference for subjects to be in pre-verbal position.

A second characteristic of the NSP includes the ability to extract a *wh*-subject from an embedded clause. This is shown by the case of the THAT-trace condition violation. According to the THAT-trace condition, the subject of a subordinate clause cannot be extracted to form a *wh*-subject to a position outside of the clause if there is a subordinating conjunction such as ‘that’. This condition is evidenced by the examples below (9-10). Taking the phrase in (9) as a starting point, we can see that extracting the subject ‘Paul’ to a position above the conjunction to act as a *wh*-subject as in (10) gives us an ungrammatical construction in English.

- (9) Lorena said that Paul ate bread.
- (10) *Who_(t) did Lorena say that (t) ate bread?

This condition is respected by languages such as English and French which are [- null subject] languages. However, [+ null subject] languages such as Spanish can violate it, as shown in (11, 12).

- (11) Lorena dijo que Paul comió pan.
Lorena say.3.SG.PST THAT Paul eat.3.SG.PST bread
“Lorena said that Paul ate bread.”
- (12) ¿Quién_(t) dijo Lorena que (t) comió pan?
Who say.3.SG.PST Lorena THAT (t) eat.3.SG.PST bread
“Who did Lorena say that (t) ate bread?”

We can see in (11) and (12) that both phrases remain grammatical in Spanish, despite the *wh*-extraction of the subject in the embedded sentence in (12). These two conditions (subject inversion and THAT-trace violation) help to exemplify the NSP properties that are active in Spanish, a [+ null subject] language, but not in English, a [-null subject] language.

1.4.2 Non-compliance with the NSP

The theory of the NSP activation as an explanatory tool for the differences between languages in regards to their use of subjects is widely known and used in the literature, but not unanimously accepted as initially proposed and many different possible additions and deviations have been proposed.

Some critiques of the NSP highlight that this theory assumes a binary distinction of languages that is too restrictive and does not accurately represent the actual variation observed in languages (Heap 1990; Huang 1994; Holmberg 2005; Camacho 2008; Sessarego and Gutierrez-Rexach 2017). Language varieties around the world show a lot more variability in respect to their actual use of subjects than the NSP theory could accommodate. For instance, Heap (1990) explained that we could have a binary differentiation between [+ null subject] and [- null subject] languages where, for example, Italian would be [+ null subject] and standard French would be [- null subject], only if we restrict the analysis to Standard versions of the languages, and this would then ignore the evidence from non-standard varieties in regions such as those ranging from Venice and Florence to Nice and Grenoble.

1 Caribbean Spanish

This is also the case, for instance, of Dominican and Puerto Rican Spanish (Camacho 2008; Cuza 2017). Although as we have seen, Spanish is considered a prototypical [+ null subject] language, these varieties of Spanish have been shown to violate some of the characteristics associated with NSLs. For instance, Dominican and Puerto Rican Spanish show a more rigid subject position with reduced frequency of *wh*-inversion (13), and use of overt expletive subject

pronouns.

- (13) ¿Qué tú quieres?
 What 2SG want.2SG.PRS
 “What do you want?”

As can be seen in (13) above, the subject pronoun *tú* in the question is produced explicitly and remains in the pre-verbal position. This phrase structure, however, would be considered ungrammatical in most other varieties of Spanish. Instead, what is usually observed in Spanish is that the subject would undergo inversion with the verb, thus moving to post-verbal position or be omitted as shown below in (14):

- (14) ¿Qué quieres (tú)?
 What want.2SG (2SG)
 “What do you want?”

The observations from Dominican and Puerto Rican varieties of Spanish provide evidence of systems that although they allow for the use of null subjects, violate the NSP by showing a more rigid word order where the subject does not undergo inversion (Camacho 2008; Cuza 2017).

Another point of evidence from these varieties is the use of overt expletive pronouns. Caribbean varieties of Spanish have been observed to use overt expletive pronouns as shown in (15). However, this is often a characteristic of [- null subject] languages, and thus violates the NSP characteristics for [+ null subject] languages such as Spanish.

- (15) Ello llueve
 EXPL rain.3SG.PRS
 “It rains.”

(Hinzelin and Kaiser 2007)

In the example above, in the case of the weather verb *llover*, which in Spanish is usually expressed with a null subject, there is the neutral pronoun *ello* acting in to fill the subject slot in the syntactic structure. This overt expletive pronoun is rare in the Hispanic world, but its use

has been recorded in cases of Caribbean Spanish, such as the Spanish of Cuba and Dominican Republic (Hinzelin and Kaiser 2007; Cuza 2017; Camacho 2013).

Some theories have been proposed to account for these observations in these Caribbean varieties of Spanish. Some explanations include: 1) that these language varieties are partial NSLs, 2) that the speakers of these varieties are in fact managing two grammars, one [+ null subject] and one [- null subject] grammar, 3) and that these varieties are in transition to becoming [- null subject] varieties (Heap 1990; Camacho 2013; Sessarego and Gutierrez-Rexach 2017).

Partial Null Subject Languages were proposed in response to overwhelming evidence from different languages and non-standard language varieties that don't fit the absolute binary [+/- null subject] categories. For instance, Finnish allows null subjects but only in the case of certain grammatical persons. Heap (1990) highlights that this pattern, where within a language subjects are obligatorily null with some grammatical persons but optionally null with others, does not seem to support the notion of a binary system where a parameter has either been activated or not. Rather, at best it would indicate that the NSP could be turned on for each grammatical person within the language independently.

The idea that these speakers are in fact managing two grammars, one which is [- null subject] and one which is [+ null subject], has also been proposed (Toribio 2000). However, this theory seems to need some refining. In this case, it is thought that speakers of Caribbean varieties of Spanish are bilingual speakers within the same language and that the only difference between these two languages is the ability or inability to use null vs. overt subjects in certain contexts. Following an assumption that, evolutionarily, all systems tend to simplify and remove redundancies, it would seem to me that having two separate grammars within our language center would be ultimately an expensive strategy over time. Further, most bilingual and bi-dialectal speakers are able to separate and use their language systems individually according to context. Many Caribbean speakers, however, never show this switch between languages and dialects and thus it seems like an unlikely explanation.

In addition, even English, which I just showed above in Section 1.4 exhibits the characteristics of a [- null subject] language, in some situations does allow for the omission of phonetically realized subjects. This will be discussed in more detail in the following subsection.

Finally, one proposal is that these varieties of Spanish in the Caribbean are transitioning from [+ null subject] languages to [- null subject] languages (Camacho 2013) in a similar manner as Old French did (Adams 1987). Similarly to Old French, there is the suggestion that an erosion of phonetic markers of the morphological endings of verbal conjugations could be an explanation. That is, in Caribbean Spanish, there is the well-documented phenomenon of aspiration and deletion of /s/ in coda position (Heap 1990; Tennant et al. 2015). This results in some cases in added ambiguity in the morphological paradigm as shown below in Table 1.1.

Subject Pronoun	Indicative Present Tense (No /s/ deletion)	Indicative Present tense (/s/ deletion)
Yo (1SG)	/kanto/	/kanto/
Tú (2SG.informal)	/kantas/	/kanta/
Ústed (2SG.formal)	/kanta/	/kanta/
Él/Ella (3SG)	/kanta/	/kanta/
Nosotros (1PL)	/kantamos/	/kantamo/
Ustedes (2PL)	/kantan/	/kantan/
Ellos/Ellas (3PL)	/kantan/	/kantan/

Table 1.1: Phonetic realizations of present indicative conjugations of the verb *cantar* (to sing) with and without /s/ deletion.

Table 1.1 shows that when the /s/ in coda position is deleted, the distinctiveness of the morphological forms becomes less evident. This means that the subject of the construction will then be harder to recover by relying only on the morphology of verbal forms. The functional compensation hypothesis suggests that this increase in ambiguity pushes the use of overt subject pronouns (Poplack 1980; Hochberg 1986). This is then similar to what was observed as part of the change for Old French as the morphological verbal endings became less phonetically salient. However, Adams (1987) notes that an increase in the use of overt subject pronouns in Old French was notable before the loss of phonetic distinctiveness of morphological verbal endings. Nevertheless, it is interesting to see a similar pattern being followed by

Caribbean varieties of Spanish and some linguists suggest this may indicate that these Spanish varieties are transitioning to [- null subject] language varieties as French did (Camacho 2008; Cuza 2017).

2 English subject pronoun omission

English allows for the omission of thematic subject pronouns in some cases (Haegeman 1999; Torres Cacoullos and Travis 2015). For instance, subject omission in English is possible in certain registers such as informal spoken forms and in some written registers, namely in cases where the medium favours economy such as e-mails, notes, and letters, among others (Haegeman 1999). Torres Cacoullos and Travis (2015), note as well that omission of subject pronouns is possible in English in *and*-coordinated clauses (Torres Cacoullos and Travis 2015) as in (16) below.

(16) She_i went to the store and Ø_i bought candy.

In this example above, we see an *and*-coordinated clause where the subject of the second clause has been omitted. This is an allowable subject pronoun omission in English as long as the subject of the second clause is co-referential with the subject of the first clause.

(17) - Where are you?
- Ø just left work now.

In example (17), in response to the question “where are you?”, a speaker may omit the subject pronoun that would normally be required in the English utterance. These constructions, where a subject pronoun is omitted, are common in English in contexts such as diary entries, or informal exchanges (Hochberg 1986; Harvie 1998). Today, with the rise in prevalence of written informal communication thanks to social media outlets such as Facebook, Instagram, Snapchat, etc., it is not difficult to think of similar examples where speakers of English might omit a thematic subject pronoun as in (18) whether in response to a question or not.

(18) “Took this pic yesterday...”

(Example taken from public Instagram post)

Note that in these cases (17, 18), the omission of the thematic subject pronoun leads to the understanding that the referent is a first person subject. Any other grammatical person subject would lead to these constructions being ungrammatical.

Note as well that in English, subject pronouns can also be omitted depending on the specific contexts. For instance, if we reconsider the question above, and change it so the subject is now a third person as in (19), the omission of the third person subject pronoun is still acceptable.

(19) - Where is she?

- Ø just left work now.

In this case, the omission of the third person singular pronoun is acceptable since there is only one person the statement could be relating to. However, in absence of the context of the question, the phrase would be ungrammatical for the third person singular resulting in an assumed first person singular reading as in (20).

(20) Ø just left work now.

The examples above (17, 19 and 20) then show that in English, subject pronoun omission leads to a first person singular reading as default in the absence of additional context.

In considering the use of null-subjects in informal registers of English, we recognize that English does not restrict itself to a total [- null subject] character. By extension this variation demonstrates that a binary distinction cannot be made in natural languages in the case of subject use.

In light of the evidence from partial NSLs, (such as Dominican Spanish and Puerto Rican Spanish) and variation even within [- null subject] languages (such as those just outlined for English), it has been suggested that this is possible if we include certain adjustments to the NSP. One suggestion is to propose that the NSP includes multiple micro-parameters, and speakers

from different language varieties in fact activate a selection of these micro-parameters, (not just one large null subject parameter), which leads to increased variation.

That being said, even if we don't adhere to a completely binary or parametric categorization of languages according to the NSP, we can still state that languages vary according to the degree or frequency, and the specific grammatical contexts in which overt and null subjects are allowed. In using this latter statement, then we can state that English remains more restrictive than Spanish in regards to the use of null subjects. For instance in the cases above, as mentioned, subject omission in English will by default give the idea of a grammatical first person referent in absence of appropriate context, and English allows for null-subject usage only in specific contexts such as written media (such as diary entries) and in media that favour economy (Haegeman 1999). In contrast, Spanish allows for optionality in regards to the use of overt and null subjects with any grammatical person, and in any communication medium or context. Therefore, the question as it concerns the current study is not whether contact between a categorical [-null subject] language and a [+ null subject] language can have an effect on the frequency of use of overt subject pronouns. Rather, the question is whether a language which shows a higher frequency of overt subject pronoun usage can influence and raise the frequency for a language which shows a lower frequency of use of overt subject pronouns in bilingual speakers. I therefore focus on investigating the likelihood of convergence between the two languages in contact based on frequency cues from input rather than on a binary, parametric view of bilingualism.

1.5 Spanish Subject Pronouns

Spanish can be categorized in general terms as an SVO language with subjects that, when produced overtly, usually appear in pre-verbal position as shown in (21).

- (21) Ellos comen pan.
 3PL eat.PRS.3PL bread
 "They eat bread."

There are some exceptions to this general word order. These include the case of presentational phrases (22), and cases including subject-verb inversion such as questions (23).

- (22) Llegó María.
 arrive.PST.3SG María
 “María arrived.”
- (23) ¿Qué comen ellos?
 What eat.PRS.3PL 3PL
 “What do they eat?”

Note that in the examples (23) and (22), the subject, in each case, is in postverbal position showing a VS structure rather than an SVO structure.

Spanish is also considered to be in general a null subject language and in many cases, (although admittedly with some exceptions), it displays the characteristics of null subject languages, including the grammatical use of either a null or an overt subject in some clauses. There are, however, some cases in Spanish where the subject expression is non-variable and the subject of the construction is either necessarily null or overt. In this section, I will briefly discuss the non-variable contexts in which subjects or subject pronouns are categorically null or overt before focusing on variable subject pronoun expression (SPE) in Spanish. In addition, in the case of variable SPE, I will also discuss the nature of the variation in both Spanish monolingual and bilingual communities.

1.5.1 Non-variable subject expression

This section focuses on cases in Spanish where subject expression is either obligatorily null or obligatorily overt.

1 Categorically Null Subjects

Some constructions in Spanish such as weather verbs (24), impersonal verbs including existentials (25), and modal verbs (26), do not take a thematic subject and deviate from the

more typical two argument structure (Melis and Flores 2005).

- (24) Ayer llovió.
yesterday rain.3SG.PST
“It rained yesterday.”
- (25) Hay tres perro-s.
there-be.3SG.PRS three dog-PL
“There are three dogs.”
- (26) Parece que ella tiene hambre.
seem.PRS.3SG RLV 3.SG.F have.3SG.PRS hunger
“It seems that she is hungry.”

In the case of the examples (25-26) above, none of the main clause constructions refer to a thematic subject, or an element (often an entity performing an action or demonstrating a certain state) which fulfills the requirement by the verb for a thematic argument. Therefore, any subject included in these sentences would be an expletive subject. Expletive subjects are semantically empty lexical insertions whose main role is to fill the space of subject within the syntactic structure (Camacho 2013).

In English, expletive subjects such as ‘it’ or ‘there’ need to be overtly expressed in these types of sentences since English does not allow for the use of null subjects. However, in Spanish, expletive subjects are categorically null in almost all varieties with the exception of Dominican and other Caribbean varieties of Spanish (Henríquez Ureña 1939; Toribio 2000).

Toribio (2000) points to the use of an expletive overt pronoun *ello* in sentences of Dominican Spanish such as in example (27) below, where we see a weather verb being introduced by an overt expletive *ello*.

- (27) Ello no está lloviendo aquí pero allá sí.
EXP.3SG NEG be.3SG.PRS rain.PRS but there yes
“It is not raining here but over there it is.”

(Bullock and Toribio 2009, :11)

However, constructions such as in (27) above are considered ungrammatical or questionable in most varieties of Spanish. These constructions are mostly observed in Caribbean varieties, which as described earlier in this chapter differ significantly from other varieties of Spanish and have even been argued to be moving towards becoming non-null subject language varieties (Cuza 2017). These varieties are thus an exception, with most varieties of Spanish categorically disallowing the expression of expletive subjects.

2 Categorically Overt Subjects

Some cases where subject pronouns are categorically overt due to the nature of the construction include instances where the subject pronoun was used emphatically, as in the example below (28a).

- (28) a. Fue *ella* quien comió.
 be.PST.3.SG 3.SG.F who eat.PST.3.SG
- b. #Fue \emptyset quien comió.
 be.PST.3.SG \emptyset who eat.PST.3.SG
 “It was she who ate.”

In the case of the above construction, we can imagine a situation perhaps where someone is identified as the person who ate in contrast to someone who did not. In this case, in the example (28a) above, omitting the overt pronoun *ella*, would result in the construction in (28b) which would be ungrammatical since it is an emphatic sentence meant to highlight the agent of the action and would thus lack necessary information.

Similarly, another example of a situation where a pronoun would be obligatorily overt would be in response to the question “who?”. For instance, if we consider example (28) above, we can imagine a situation where someone may have asked *¿Quién comió?* ‘Who ate?’. In this case, the answer to the question requires a subject or subject pronoun to identify the agent of the action. It may be given in a phrase similar to the one above or it may be answered more simply with a *Ella* ‘She’, or *Ella comió* ‘She ate’, but again, including a null subject in this phrase would lead to \emptyset *comió* ‘#ate’ which results in a pragmatically inappropriate response.

1.5.2 Variable Subject Pronoun Expression

In order for a language phenomenon to be variable, we need to have two equivalent expressions where they both provide the same semantic content (Lavandera 1978). This is easier to note in phonetic analyses, for instance in the case of allophonic variation where the pronunciation of a given sound as one of two allophones in a language does not in any way alter the meaning conveyed by the word of which it is part. We argue this is also possible, however, in the morphosyntactic domain in cases such as variable subject expression in Spanish. While in some cases the use of overt subjects is associated with a more emphatic or assertive affirmation (Lipski 2002), situations where subject pronouns are used with an emphatic purpose in this way require categorically overt subjects and therefore these situations would not be part of the variation described when discussing variable subject expression. In the case of variable subject expression, there is true optionality and the use of a null or overt subject in a given clause renders semantically and grammatically equivalent constructions. The use of a null or overt subject pronoun in a given clause is conditioned by a number of factors, making a null pronoun more likely than an overt pronoun in a given context and vice-versa, but this is typical as well of other variation phenomena, including phonetic variation, since linguistic variation is rarely, if ever, completely free variation.

1 Spanish Variable Subject Expression in monolingual contexts

In a large number of cases where null subjects are possible there is optionality: the speaker can choose to use an overt subject instead while maintaining the grammaticality and semantic equivalency of the phrase (Lipski 2002).

On average, Spanish subject pronouns are overtly expressed between 19% and 41% of the time in variable contexts (Otheguy et al. 2007). This rate varies between different varieties of Spanish. While Hispanic regions of continental Latin America, such as Mexico, Colombia (Bogotá), and Ecuador, use overt subject personal pronouns between 19% and 27% of the time, speakers in Caribbean regions such as Cuba, Puerto Rico, Dominican Republic, and the

Atlantic coast of Colombia use overt SPPs between 33% and 41% of the time (Otheguy et al. 2007; Orozco 2015).

Variability between regions of the Hispanic world, however, lies mainly at the level of frequency of SPP expression. The factors that significantly condition the use of overt vs. null SPPs seem to remain consistent among many varieties. In variable contexts research supports the contention that the variability of SPP expression is constrained by a number of grammatical and pragmatic factors (Flores-Ferrán 2004; Montrul 2004), including clause type, switch reference, grammatical person and number, verbal Tense, Aspect, Mood (TAM), and negation (Otheguy et al. 2007; Orozco 2015). The most consistently significant factors among different varieties of Spanish are switch reference, and grammatical person and number (Silva-Corvalán 1994b; Flores-Ferrán 2004; Otheguy et al. 2007; Otheguy and Zentella 2012; Orozco 2015).

One factor that seems to consistently condition the variable use of subject pronouns in Spanish, even among different varieties of Spanish, is the grammatical person and number of the construction in which the pronoun (null or overt) is found. This may be due to the fact that not all verbal forms in Spanish are equally unambiguous. That is, some forms within the verbal paradigm allow for the recovery of the subject of the clause more easily than others. For instance, let us consider the conjugation of the verb *cantar* ‘to sing’ as an example (Table 1.2).

Subject Pronoun	Indicative	Conditional	Subjunctive
Yo (1SG)	<i>canto</i>	<i>cantaría</i>	<i>cante</i>
Tú (2SG)	<i>cantas</i>	<i>cantarías</i>	<i>cantes</i>
Ústed (2SG.POL)	<i>canta</i>	<i>cantaría</i>	<i>cante</i>
Él/Ella (3SG)	<i>canta</i>	<i>cantaría</i>	<i>cante</i>
Nosotros (1PL)	<i>cantamos</i>	<i>cantaríamos</i>	<i>cantemos</i>
Ustedes (2PL)	<i>cantan</i>	<i>cantarían</i>	<i>canten</i>
Ellos/Ellas (3PL)	<i>cantan</i>	<i>cantarían</i>	<i>canten</i>

Table 1.2: Present tense conjugations of the verb *cantar* ‘to sing’ in the indicative, conditional, and subjunctive moods.

In table (1.2) we can see conjugations for the verb *cantar* ‘to sing’ for all persons and numbers, in the present tense, for the indicative, conditional and subjunctive moods. We can notice that in the indicative mood, the third person singular conjugations of this verb are equal to that

of the formal second person singular. Also, the forms for the third person plural and for the second person plural are identical⁶. In the case of the conditional and subjunctive conjugations we observe the same pattern, but in these cases we now have a three-way equivalency since the first person singular forms are equal to the second person formal singular *usted* and third person singular forms.

These similarities between conjugated forms of the verb *cantar* introduce a certain level of ambiguity, and if a null subject is used they can lead to confusion between the speaker and the listener in a conversation (29).

- (29) *Canta una canción.*
 sing.PRS.IND.(3SG/2SG.POL) DET.F song
 “(he/she/you) sing(s) a song.”

In example (29) the sentence *Canta una canción* ‘(he/she/you) Sing a song’ is ambiguous. Without having any additional context, it is impossible to know who sings the song. The best that we can infer is that it is either a second person (formal) singular or third person singular subject. In addition, the mood of the sentence is ambiguous since the sentence could be in the indicative or the verb form could represent the second person singular imperative such as *¡Canta una canción!* ‘Sing a song!’. This ambiguity may be resolved by the surrounding context or preceding continuing topic (for instance if we knew that the person was speaking about a favourite singer). If the surrounding context is insufficient, however, it is likely that an overt subject pronoun would be introduced to disambiguate (D’Alessandro 2015).

Verbal ambiguity can be magnified in some varieties, namely in Caribbean varieties of Spanish where /s/ aspiration and deletion are commonly observed phenomena. In these varieties, an /s/ in coda position has three possible allophones which are /s/, /h/ or /ø/, such that a word like *casas* (houses) can be pronounced as /ka.sas/, /ka.sah/, or /ka.sa/ depending on a number of linguistic and social factors (Poplack 1980; Tennant et al. 2015).

This variety therefore introduces further ambiguity in regards to verbal forms when the

⁶In this study, I do not consider the peninsular second person plural form *vosotros* since it is not a form used in the variety Spanish of Cundinamarca, Colombia.

inflectional morphology involves an /s/ in final position as is the case in the second person singular forms of regular Spanish verbs in the present tense (Table 1.2). For instance, let's return to the case of the verb *cantar*. In this case, the informal second person singular form of the verb in the present indicative is *cantas* where the final /s/ pronunciation differentiates it from *canta* which is the form for the formal second person singular, and the third person singular. If the final /s/ of this form is omitted, we now end up with three identical phonetic forms of the verb. According to the functional compensation hypothesis (Poplack 1980; Hochberg 1986; Cameron 1993), in response to this additional ambiguity, speakers will use subject pronouns more frequently.

The *functional compensation hypothesis* seems to be supported by empirical evidence from Caribbean Spanish, since in fact it is these varieties that show some of the highest average frequencies of use of overt subject pronouns. This is then thought to be due to the increased ambiguity introduced by the phonetic phenomena associated with speech from these varieties (Poplack 1980; Hochberg 1986; Cameron 1993; D'Alessandro 2015). It is important to note, however, that this hypothesis is not universally accepted (Poplack 1980).

Some studies seem to provide evidence against the functional hypothesis. For instance, the basis of the functional compensation hypothesis, that is, that overt subject pronouns are more frequently used in cases where compensation for a higher level of ambiguity is needed, does not seem to hold when considering other null-subject languages such as Cantonese and Japanese which have no subject verb agreement (Cameron 1993; Turan 1996). In addition, according to Ranson (1991), the functional hypothesis as is does not hold when considering the Spanish of Andalusia where the overt expression of subject pronouns does not increase with more ambiguous word endings which result from /s/ aspiration.

Another factor that seems to be consistent in significantly conditioning the overt use of SPP's across different varieties of Spanish is change of referent, also referred to as switch reference. Many studies, focusing on different varieties of Spanish, have found that Spanish speakers are more likely to use an overt SPP in a given clause when they are introducing a new

referent into the discourse (Silva-Corvalán 1994b; Flores-Ferrán 2004; Montrul 2004; Orozco 2004; Otheguy et al. 2007; Otheguy and Zentella 2012). In addition, it seems to be the case that the further a SPP is from its referent, the more likely it is to be expressed overtly (Flores-Ferrán 2004). This factor seems to be among the most significant and consistent factors conditioning the variable expression of SPP's in Spanish and it may in fact tie in as well to the idea of reducing ambiguity. When a new referent is being introduced, or when the referent is changing within the discourse, using a null SPP may lead to confusion.

- (30) Ella estaba en entrenamiento y después \emptyset llegó tarde
 3SG.fem sing.IMP.3SG DET song and later *pro* arrive.PST.3SG late
 y \emptyset no pudo ver el juego.
 and *pro* NEG can.PST.3SG see.INF
 “She was in training and then (he/she/it/you(formal)) arrived late and (he/she/it/you
 (formal)) couldn't watch the game.”

Example (30) above shows one context in which some confusion might arise from an over use of null subject pronouns, specifically in contexts in which an overt subject is pragmatically necessary to ensure sentence clarity. In this case, it is unclear whether it was the same person who was the subject of all the verbal constructions (i.e., She finished training and then arrived late and was not able to watch the game), or whether there are two or more people who are subjects to the other verbal constructions (i.e., As she trained someone else arrived late and was unable to watch the game). The statement could be made clear by overtly expressing the subject whenever there is a new referent or a change of referent. This factor, therefore, seems to also tie the use of overt SPPs in Spanish to the idea of clarifying information that is not readily recoverable from the immediate physical or linguistic context surrounding the expression. That being said, there are other factors that have been found to significantly condition the variable expression of SPPs that are not as clearly tied to disambiguation such as clause type, the semantic content of the verb, and the presence of negative elements in the clause (also known as polarity).

The type of clause seems to be a significant factor in conditioning the use of overt subject

pronouns in various studies. For instance, Orozco (2015) found through his analysis of the Spanish of Barranquilla, Colombia, that clause type is a significant factor in conditioning variable expression of SPPs. Specifically, he found that subordinate clauses favoured the use of overt SPPs, coordinate clauses favoured null SPPs, and independent clauses remained neutral. That being said, his results, although significant, were marginal. However, the author notes that his results differ somewhat from those found in similar studies which did not find clause type to be a significant conditioning factor for variable expression of SPPs (Peskova 2013; Torres Cacoullos and Travis 2015). This difference in significance, however, may point to a difference between monolingual and bilingual speakers (Abreu 2009).

In a study by Abreu (2009), an analysis of monolingual, bilingual and second language speakers of Spanish showed that clause type was a significant factor for English-Spanish bilinguals and Spanish second language speakers, but not for monolingual Spanish speakers. In the case of bilingual and second language speakers, overt subject pronouns were favoured in main clauses while in dependent and relative clauses null subject pronouns were favoured. The author mentions that this could be pointing to an effect from contact with English, but it remains unclear how English-contact or English language dominance may be influencing subject pronoun expression according to clause type. Further, she suggests that this may be due to main clauses being more likely to be used by speakers to introduce a new referent, while dependent or relative clauses are more likely to be used to continue with a pre-existing referent. This, however, would not explain the lack of significance of this factor for monolingual speakers.

The semantic content of the verb has also been noted to influence the use of overt vs. null subject pronouns. Peskova (2013) studied the variable expression of SPPs in relation in the Spanish of Buenos Aires. In relation to semantic verb type, Peskova (2013) looked at epistemic vs. perceptive verbs, where the former category included verbs such as *creer* 'to believe' and the latter included verbs such as *mirar* 'to look'. They found that this factor was significant with epistemic verbs favouring the use of overt SPPs more than perception verbs.

Semantic verb content was also considered by Orozco (2015) in his analysis of Barranquilla

Spanish. He categorized verbs in his data according to whether they were denoting mental or emotional activity, an external activity (denoting physical actions), or a passive state (stative verb). He found that verbs related to mental or emotional actions, such as ‘to feel’, ‘to love’, ‘to think’, were more likely to be accompanied by overt subject pronouns, while verbs denoting a physical action were more likely to be accompanied by a null subject. It seems then that in general, in Spanish verbs denoting mental or stative activities, such as verbs denoting beliefs, thoughts, or passive states, are more likely to be accompanied by overt SPPs while verbs denoting physical actions or other less passive actions correlate with null SPPs. This may in turn connect to an idea that in the case of mental or stative verbs, speakers seek to emphasize that the thought or idea being expressed is the speaker’s own since these verbs tend to indicate subjective ideas such as beliefs or thoughts (Harrington and Pérez-Leroux 2016). In contrast, physical action verbs may be more transparent in their context regarding the subject of the clause and may require less information to be overtly expressed to convey the same meaning.

Studies have also shown that the variable expression of subject pronouns can be influenced by the negative or affirmative polarity of the sentence in which the subject is found (Gridstead 1998; Nagy 2015; Lastra and Butragueño 2015). For instance, Lastra and Butragueño (2015) who considered the Spanish of Mexico city, determined from their data that negative clauses favoured the use of null subject pronouns, while affirmative clauses favoured the use of overt subject pronouns. The authors explained that instances of negative clauses tended to also be serial in nature, and in fact isolated negative clauses were rare in their data (31).

- (31) “[Ø] empecé a ter-/bueno/ [Ø] empecé otra vez la carrera// pues sí/sí [Ø] la terminé /ya/ ahí en la Sep/ [Ø] **ya no pude aguantar** o sea/ ya era así de que/ [Ø] **ya no quería ir**// [Ø] **ya no quería**/ o sea nomás [Ø] me acord-/ “¡chin!/ [Ø] ya me tengo que ir a trabajar” // **no**/ [Ø] **no quería ir**/ a veces pues <-pus> (risa) [Ø] **ya ni iba**/ ¿no?/ mejor [Ø] me iba a mi casa.” (ME-257-32H-05, turn 40)

(Lastra and Butragueño 2015, 16)

“I began to fin-/well/ I began again the program// well yes/yes I finished it /already/

there in the Sep/ **I couldn't stand it anymore** I mean/ it was already like as if/ **I no longer wanted to go// I no longer wanted to/** I mean I just rememb-/ “shoot!/ I now have to go to work” // **no/ I did not want to go/** sometimes well <-well> (laughter) **I no longer even went/** no?/ I rathered go to my house.”

This kind of repetition, explain Lastra and Butragueño (2015), is significant since the seriality of the construction leads to the referent for all the clauses being nearby which can in turn contribute to favouring the production of a null SPP over an overt SPP. Similarly to switch reference, negation may be favouring the use of null pronouns due to the referent being more easily recovered from the linguistic context thanks to repetition.

A similar finding was reported by Harrington and Pérez-Leroux (2016) who conducted an analysis of negated and affirmative epistemic clauses and how the mood of the dependent clause in these constructions influenced overt expression of SPPs in monolingual Spanish. Namely, they conducted their analysis using monolingual Spanish data from the *Habla Culta* section of the *Corpus del Español* (Davies 2002), which included oral interview transcripts from different varieties of Spanish including the Spanish of Bogotá, San Juan, La Paz, and Buenos Aires among others. Harrington and Pérez-Leroux (2016) concluded that while mood of the dependent clause didn't seem to influence subject realization in affirmative clauses, there was a significant effect in the case of negated clauses. In negated clauses, overt subjects were favoured if the clause was a dependent clause in the indicative mood, while null subjects were favoured for dependent clauses in the subjunctive mood. The authors propose that these results may show that when a dependent clause of a negated epistemic construction is in the indicative the speaker may have intended a more contrastive tone emphasizing that the statement is the speaker's *own* opinion rather than a general view which would be more likely to be expressed using the subjunctive. This would then be related to the higher likelihood of incorporating an overt subject pronoun in order to increase emphasis and to highlight the subjectivity of the statement, and therefore the authors conclude that there is a link between the use of overt subject pronouns and contrastive focus. In other words, they propose that the use of overt

subject pronouns seems to be related to increased subjectivity and intended contrastiveness, while the use of null subjects creates a more objective reading.

The idea of overt subject pronouns being used for additional contrast is also supported by others (Ordóñez and Treviño 1999; Mayo and Clark 2010). For instance, Ordóñez and Treviño (1999) argue that preverbal overt subjects are not occupying [SPEC,IP] position as it is traditionally thought, but are rather in the left periphery of the structure, patterning closer with left-dislocated direct and indirect objects. They argue that in Spanish, it is the verbal agreement morphology that, in acting as a clitic, satisfies the agreement requirements in the structure. According to their analysis, then, movement of the subject to preverbal position would only occur if there are other discursive reasons for doing so, such as emphasis or contrastive focus, in a manner similar to other left-dislocated structures. Therefore, while the reason why negative clauses seem to favour null subject realization is not uniformly agreed upon, according to Turan (1996), negation can have an effect on an NP making it less definite since negation indicates that either the entity itself or the action relating to the entity is negated. It is then possible that this effect of negation on NPs could render them as less emphatic. Since emphatic clauses in Spanish tend to be correlated with a categorical use of overt subject pronouns, then the presence of negation could be favouring null subjects by approximating them to clauses that are less emphatic and which, therefore, are then less likely to appear with overt subjects. We may be able to explain this by envisioning a “scale of emphasis” where clauses emphasizing a subject, for instance clauses answering to the question “who?”, or epistemic clauses selecting an indicative dependent complement, are most emphatic and favour an overtly expressed subject, and clauses where there is no thematic subject to be emphasized, such as weather or impersonal verb constructions, are least emphatic. Within this scale then negative clauses may lie lower than affirmative clauses and would therefore be more likely to favour null over overt subject pronouns.

In addition to linguistic factors, some studies have found that social factors such as level of education may influence the expression of subject pronouns. For instance, a study focusing

on the populations of Bogotá, Colombia, and Mexico City, Mexico, found that speakers with a higher level of education used overt first person singular pronoun *yo* 'I' with a higher frequency than speakers with a lower level of formal education (Lipski 1989). The author concluded that speakers with a higher level of education might be paying closer attention to possible ambiguities and overextend the use of overt subject pronouns to contexts where they would otherwise be unnecessary.

2 Spanish Variable Subject Expression in bilingual contexts

The use of SPP's by bilingual speakers, and how this use can be influenced by contact between languages has been of the focus of much linguistic research (Poplack 1988; Hulk and Müller 2000; Flores-Ferrán 2004; Montrul 2004; Hurtado 2005; Abreu 2009; Nagy 2015; Otheguy et al. 2007). Particularly, several studies have investigated the possible influence of English on Spanish in regards to the use of SPPs in English-Spanish bilingual speakers in the United States (Silva-Corvalán 1994a; Montrul 2004; Otheguy et al. 2007; Abreu 2009; Orozco 2004).

According to Müller and Hulk (2001) in order for there to be the potential for cross-linguistic influence, the two languages in contact must have characteristics that overlap at the surface level but that are in fact distinct from each other at the underlying syntactic level. This surface overlap creates a competition between the two structures and can lead to cross-linguistic influence of one language system over the other which will then be reflected in the speech production of the bilingual speaker. This is in fact the case of Spanish and English in respect to the use of SPPs. Both English and Spanish are SVO languages and therefore resemble each other in regards to the distribution and placement of subject pronouns in most phrases, but the underlying characteristics that condition their use differ in important ways as we explored in the previous section. Recall, for instance, that while Spanish use of overt SPPs is conditioned by discourse-pragmatic factors such as the novelty of the referent in the discourse, English lacks these pragmatic constraints and the use of overt SPP's is almost always mandatory. This

type of overlap then satisfies the condition as stipulated by Müller and Hulk (2001) and makes it theoretically possible for there to be cross-linguistic influence between the two systems in the speech of Spanish-English bilinguals (Paradis and Navarro 2003).

In order to explore the possible influence between these two linguistic systems in regards to the use of SPPs, Otheguy et al. (2007) studied a population of Hispanic speakers in New York City. They investigated the speech of 1) Hispanic speakers newly arrived in the United States, 2) Hispanic speakers who were considered to have been living in the community for an extended length of time (about five years or longer), and 3) heritage Spanish speakers who were native or near native to the city of New York. Following their analysis, they found that the length of residence in New York, (and thus the duration of the situation of language contact), was positively correlated with an increase in frequency of use of overt Spanish subject pronouns in relation to their Spanish monolingual counterparts from their country of origin. This led to the conclusion by the researchers that the situation of intense language contact with English was leading to a convergence of the Spanish and English linguistic systems in the bilingual's use. It is important to note that this differed from the argument by ? who believed that bilinguals do not converge in their grammars, but rather that the two grammars of bilinguals can influence each other given specific conditions as described above.

Several similar studies focusing on different communities of Hispanic speakers in the United States, such as Chicago and Miami, corroborate the findings in the study by Otheguy et al. (2007) showing that Spanish-English bilingual speakers use overt SPPs with a higher frequency (Montrul 2004; Orozco 2004). These authors propose that due to the situation of intense language contact between English and Spanish, bilingual speakers show evidence of convergence in their grammars coming to a less cognitively demanding strategy where instead of manoeuvring between one grammar where overt subject expression varies (in many cases optionally) according to a number of pragmatic and grammatical factors and one where the expression of subjects is constant and obligatory, speakers overgeneralize by converging towards the more consistent and less complex grammar. However, some researchers consider that this observa-

tion may be due to reasons other than the situation of language contact itself.

For instance, Paradis and Navarro (2003) compared the use of overt and null subject pronouns between two monolingual children and one Spanish-English bilingual child. They found that in fact the bilingual child tended to use overt subject pronouns with more frequency than the monolingual children and argue that while cross-linguistic influence is apparent in the bilingual child's language production, other explanations should be considered. The authors suggest that the increase in use of overt SPPs in Spanish bilinguals may be due to the variety of Spanish from which they receive their linguistic input during language acquisition. In their study, Paradis and Navarro (2003), an analysis of the parents' use of subject pronouns was conducted and they found that the Spanish input that the bilingual child was exposed to showed higher use of overt SPPs than that which the children from monolingual families were exposed to. This provides evidence to support the assertion that the bilingual child's language use is showing a reflection of the language input they are receiving rather than the result of an internal mechanism where their two linguistic systems are converging. However, no definite conclusion could be reached in this study and more research to tease apart these two aspects is needed.

Another alternate explanation to the observed increase in frequency in bilingual speakers is that it is a general consequence of language contact in transitional bilinguals. Sorace (2011) found that an increased frequency of overt subject pronominal expression is apparent even in cases where the two languages are both [+ null subject] and therefore this increase in frequency is in fact rather a consequence of transitional bilingualism and not of contact with a specific language. In her study, she studied Italian-Spanish bilingual speakers and found that these speakers, who were heritage speakers of Spanish, showed an increased use of overt subject pronouns similar to that seen in bilingual contexts with English in the United States. She concluded then that it is the bilingual cognitive demands that lead bilingual speakers to try and disambiguate more frequently, and consequently, to use overt subject pronouns with more frequency.

Other studies have then focused on looking at how social and extra-linguistic factors may

condition the variable subject expression in bilingual Spanish. Shin and Orozco (2013) found that the use of variable subject pronouns in New York varied as a function of speakers' gender and socioeconomic level. Specifically, their results suggested that Hispanic women in New York, as well as Hispanic speakers of higher socioeconomic status, lead the increase in the use of overt SPPs. The authors point out that women leading a linguistic change is not a surprising fact and actually aligns well with the pattern observed in studies of linguistic change in monolingual communities (Trudgill 1974). Women are often the members of the community that have closer contact with younger speakers, including their own children, and therefore are more likely to be the individuals to pass on certain variations in linguistic forms to these new generations. It is however contrary to expectations to see that it was speakers of higher socioeconomic status in New York City that seemed to be leading the change towards an increase in the use of overt pronouns since this feature is often associated with lower prestige in bilingual communities.

Higher rate of use of overt SPPs can be seen as a symbol of lower prestige due to the fact that its use is associated with Caribbean varieties of Spanish which in general enjoy less prestige in the Spanish-speaking world. This is mentioned in Garrido (2007) where she notes that when Colombians were interviewed regarding their attitudes towards different varieties of Spanish, it was the Spanish of Bogotá that was considered the most prestigious variant, and the Spanish of the Caribbean and Pacific coasts which were the least prestigious variants. Therefore, Shin and Orozco (2013) note that the higher use of overt SPPs among Spanish speakers from a higher socioeconomic status in NYC is counter-intuitive, but may be due to the fact that more affluent members of the Hispanic community have smaller social networks and are more susceptible to influence from English, showing a more "anglicized" Spanish, and thus revealing an opposition between prestige monolingual and prestige bilingual features. It is also possible, in my opinion, that in the case of the bilingual context of New York City, features of Caribbean Spanish may enjoy a higher prestige than in monolingual contexts since these varieties have been present in these regions of the United States for longer and therefore

may act as a model for later arriving groups of Hispanics.

Another suggestion that has been put forth to explain the results from previous studies focusing on Spanish-English bilingual speakers, which highlight a higher rate of use of overt SPPs, is cross-linguistic priming effects (Torres Cacoullós and Travis 2010; Abreu 2012). ? studied the variable use of SPPs in bilingual Puerto Rican speakers living in Florida. She determined that priming effects were among the most significantly relevant factors conditioning the use of overt SPPs. Mainly, her results from a multivariate analysis showed that speakers were more likely to produce overt SPPs following an immediately preceding overt SPP regardless of whether it was produced by themselves or the interviewer. More interestingly, she found that overt SPP production in Spanish clauses was significantly more likely if there was an English clause including an SPP preceding it. In this case Abreu concludes that it is the result of priming during online discourse production rather than an effect of syntactic convergence between the Spanish and English grammar of bilinguals. That is, bilingual speakers tend to code-switch between the two languages, in this case English and Spanish, and since English produces subject pronouns categorically overtly most of the time and since priming is a universal non language-specific effect, it is the presence of the overt pronoun that primes speakers to also use overt SPPs in the following clause (even when in Spanish) leading necessarily to higher rates of overt subject pronoun production in bilingual than monolingual speakers. These results corroborated the conclusions reached previously by Torres Cacoullós and Travis (2010) who stated that cross-linguistic priming, rather than language convergence leading to changes of grammatical patterns, was a primary factor contributing to the slight increase in the use of overt SPPs in the Spanish of Spanish-English bilinguals from New Mexico.

It is also important to note that not all studies of Spanish-English bilinguals have observed an increase in the frequency of use of overt SPPs. In fact, in her study of the Spanish of Los Angeles, Silva-Corvalán (1994b) found that second and third generation speakers of Spanish were instead showing signs of a decrease in their use of overt subject pronouns in variable contexts. She also noted a more rigid SVO order and less inversion in contexts where new

referents are introduced. In Spanish, as described in the beginning of the current section, although the general word order is SVO, inversion to a VS order is common in monolingual Spanish when introducing a new referent into the discourse. This inversion, along with the decrease in use of overt SPPs in Spanish heritage speakers in LA, may indicate an erosion of unspecified pragmatic features rather than effects of syntactic convergence with English. That is, speakers tend to overgeneralize and use null subjects more frequently even when the context, due to pragmatic factors such as referent continuity, would disfavour it, and speakers fail to undergo SVO to VS inversion when the pragmatic factors would favour it. This then, according to Silva-Corvalán (1994b), is not an example of direct language transfer but rather an indirect effect of language contact in general, emanating instead from language attrition of these bilingual and heritage Spanish speakers.

In contrast, other studies have found that contact with English may in fact influence the use of SPPs across generations. For instance, the presence of negation in the same clause as the pronoun in question has been found to be a relevant factor in conditioning the variable expression of subject pronouns in second generation heritage speakers of null subject languages (Nagy 2015). In her study of heritage speakers, they found that second generation speakers of Russian in Toronto, but not first generation speakers, used significantly more null subjects with negative rather than with affirmative clauses. The author suggests that this may indicate influence from contact with English.

English, although canonically a non-null subject language, has been noted to use null subjects in specific contexts such as diary entries (Haegeman 1999), and according to Harvie (1998) English has increased its use of null subjects alongside negative expressions such as “can’t do it” or “don’t know”. Therefore, the increase in null subjects in negative clauses among second generation speakers of null-subject languages in contact with English could be a contact effect due to the higher incidence of null subject use in English within these contexts.

Chapter 2

Methodology

2.1 Overview and research questions

In Canada, little research exists regarding the use of Spanish, despite the fact that this language now represents the fifth most frequently used non-official language in the country (Statistics Canada 2017). Through this study, my aim is to contribute to this growing body of knowledge by considering the variable expression of Spanish subject personal pronouns (SPPs) in two generations of Colombian immigrants living in London, Ontario. The analysis presented in this dissertation aims to answer the following research questions:

1. What are the social and linguistic factors that condition the use of overt and null subject personal pronouns in Colombian Spanish in London, Ontario?
2. Are there differences between the first and second generation immigrant speakers of Spanish in regards to their variable use of subject personal pronouns? If so, how do the two generations differ from each other?
3. How does the variable expression of SPPs in the Spanish of London, Ontario (within a Canadian context) compare to findings reported for Hispanic communities in the United States?

Finding answers to these questions will not only shed light on the use of Spanish within a Canadian context, but also contributes to our understanding of the effects of English-Spanish language contact and the possible transfer effects resulting from these situations.

In the context of the United States, some studies have suggested that English-Spanish contact has led to signs of convergence or cross-linguistic influence between the two languages in the speech of Spanish-English bilinguals (Montrul 2004; Otheguy et al. 2007). Particularly, some studies have found an increased use of overt subject pronouns by second and third generation heritage speakers (Otheguy and Zentella 2012; Montrul 2004). However, this remains controversial and other studies have not found an increase when considering other Spanish-English bilingual populations in the United States (Flores-Ferrán 2004; Silva-Corvalán 1994b). In addition, some studies that have found an increase in frequency of use of overt SPPs suggest that the increased use of overt subject pronouns may be caused by dialect accommodation effects resulting from contact with Caribbean varieties of Spanish, which show some of the highest rates of use of overt subject pronouns in the Spanish-speaking world with an average of about 30%, in comparison to rates between 16% and 20% seen in Mainland varieties of Latin American Spanish (Cuza 2017). Since Caribbean Hispanics, including speakers from Cuba and Puerto Rico, constitute a well-established group in the United States, newcomer Hispanics may be assimilating to their language variety and increased overt subject pronoun rates in heritage speakers may be reflecting this dialectal assimilation (Camacho 2016; Cuza 2017), rather than a syntactic convergence with English. It is this hypothesis that I primarily target through this thesis by looking at the above research questions in a Canadian context.

The questions relating to the variable expression of Spanish SPPs in Colombians living in London, Ontario, are addressed in this work primarily through a quantitative analysis which is described in detail later in this chapter (See sections 2.4, and 2.7.1). In addition to the quantitative data, however, I incorporate an explanatory mixed-methods analysis to further consider individual speaker variation incorporating qualitative data. This allows for a more thorough understanding of the variation observed. I also use data from the sociolinguistic interviews con-

ducted to investigate how the attitudes and daily language habits of the speakers within the sample align with factors that are known to contribute to the maintenance of a minority or heritage languages, such as: use of Spanish within different social contexts; attitudes from speakers and larger community towards the language itself; attitudes towards exogamous/endogamous relationships; and sense of identity and association to the heritage culture (See section 2.6 for details on these factors). These questions are addressed through a thematic qualitative analysis that aims to ultimately determine whether Spanish in Canada is likely to be maintained or undergo language shift, and therefore qualifies the current investigation as an embedded mixed methods study which is, however, primarily quantitative in nature.

This investigation contributes to the understanding of Spanish variation by exploring the social and linguistic factors that condition the variable expression of Spanish SPPs in a Canadian context, while also indicating that non-linguistic factors, such as cultural identity, familism, and frequency of language use, can have an important effect on this variable among heritage Spanish speakers. This study also contributes to the field of language contact by providing evidence that cross-linguistic interference during bilingual first language acquisition can have an impact on the hierarchy of factors that condition the variable expression of Spanish SPPs among second generation speakers, while contact with English in itself seems to have little influence on this variable among first generation speakers who learned English as a second language later in life.

2.2 The data

2.2.1 London, Ontario : Londombia

According to Statistics Canada (2017) the majority of immigrant groups who arrive to Canada settle in some of the more populous cities in the country such as Toronto, Vancouver, and Montréal. By contrast, smaller cities receive less immigration. One such city is London, Ontario, a medium-sized city, with 521,756 inhabitants, located in South Western Ontario

(Statistics Canada 2017). In this thesis, I focus on the population of immigrants who reside in London, Ontario. As mentioned, London receives less immigration in comparison to larger Canadian cities. The immigrant populations who reside here are also less diverse in regards to the place of origin. In fact, London, Ontario has been nicknamed *Londombia* thanks to the high concentration of Colombians who have made of this city their home (Oakland 2003).

The high concentration of individuals from a single country of origin, combined with the lower overall rate of immigration, allows us to study the influence that English may have on the use of one variety of Spanish by immigrant speakers, while minimizing the influence that different varieties of Spanish exert on each other. Since most members of the Spanish-speaking immigrant community are all from a single nationality, we can consider that influence from other varieties of Spanish will be a minimal factor in any changes in language use that are observed. Nevertheless, influence from different varieties of Spanish can not be eliminated.

In London, Ontario, in addition to Colombian Spanish speakers there are Spanish speakers from different regions of Latin America including El Salvador, Mexico, Peru, and Argentina, as well as some from Caribbean countries such as Cuba and Dominican Republic, among others (Statistics Canada 2017). The population distribution for the different ethnicities, based on the data provided by Statistics Canada (2017), is shown in the table below.

While these varieties may also influence the Spanish used in the city, Colombian Spanish speakers compose 39.2% of the Latin American Hispanic population, and most of the other varieties are also mainland varieties of Spanish which are characterized by rates of overt SPP usage similar to the rates found in the Spanish of Cundinamarca, Colombia (Orozco 2015). Nevertheless, in London, Ontario some of the population does identify ethnically as coming from a Caribbean Hispanic nation or territory (Cuba, Dominican Republic, Puerto Rico), but according to the data, they represent about just a little over 5% of the population. Therefore, given the low prevalence of Caribbean varieties of Spanish we don't expect that contact with these varieties will influence the frequency of use of overt SPPs by Colombian speakers in a significant manner. Even after considering the contact between different varieties of Spanish

Country	Population
Colombia	4340
El Salvador	1995
Mexico	1620
Nicaragua	555
Cuba	380
Chile	365
Venezuela	345
Guatemala	335
Peru	260
Honduras	245
Argentina	190
Dominican Republic	180
Ecuador	70
Costa Rica	60
Panama	35
Puerto Rico	35
Uruguay	25
Bolivia	20
Paraguay	15

Table 2.1: Population numbers from Latin American groups ($N = 11070$) in London, Ontario according reported ethnicity (Statistics Canada 2017).

in London, Ontario, the Canadian contact situation remains drastically different from that in the United States, where it is more common to find large populations of Spanish speakers from Cuba, Puerto Rico, and other Caribbean countries. These varieties show some of the highest rates of overt subject pronoun usage in Spanish-speaking America, and contact with these varieties could be a factor in any changes in SPPs expression among speakers of other varieties of Spanish in the United States (Camacho 2013). In Canada, and more specifically in London, Ontario, Caribbean Spanish speakers, although present, are much less prevalent than in the United States. Therefore, by studying the speech of Colombian speakers in London, Ontario, we can isolate and investigate the effects of contact between English and Spanish more directly, while minimizing the influence from other significantly different varieties of Spanish.

One important consideration to highlight is that within the *Londombia* of London, Ontario, there are speakers from different regions of Colombia, and the Spanish varieties of Colombia do differ from one region to another (Flórez 1961). The different dialectal areas of Colombian

Spanish have been discussed in several works (Flórez 1961; Orozco 2004). Florez (1961) divides the Spanish of Colombia into seven different dialectal regions: 1) Coastal (including both the Caribbean and Pacific coasts); 2) Greater Antioquia; 3) Cundinamarca/Boyacá; 4) Greater Tolima; 5) Greater Cauca; 6) Greater Santander; 7) Llanera region (including the South East of Colombia incorporating the Amazonian region). However, Orozco (2004) explains that some other dialectal studies, such as that by Montes Giraldo (1982) differ. For instance, Montes Giraldo (1982) further subdivides the coastal region of Colombia into two macro-regions, separating the speech of the Caribbean coast from the speech of the Pacific coast of the country. According to the 2016 Census data provided by Statistics Canada (2017), a total of 4340 inhabitants in London consider themselves to be Colombian. Unfortunately, the census data does not report specifically on the regions of Colombia where immigrants come from. Nevertheless, it is the general impression of the community that a large proportion of individuals come from either Bogotá (and the surrounding cities), Cali (Greater Cauca region), or Medellín (Greater Antioquia region). For the purposes of this study, I focus on the Spanish of the region of Cundinamarca, a region of Colombia located on the Eastern mountain range of the Andes, which includes cities such as Bogotá, Fusagasugá, Chía, and Tenjo among others, and which represents a mainland Latin American Spanish variety. Speakers from other areas of Colombia such as Cali and Medellín were excluded from the study in order to control for the dialect origin variable.

2.2.2 The CoSLO corpus

A total of 40 interviews were conducted following participant recruitment and of these, a total of twenty interviews, averaging approximately one hour each in duration, were ultimately included within the corpus of Colombian Spanish in London, Ontario (CoSLO). This represents approximately 1200 minutes of recorded speech data in total.

The corpus includes speakers from two groups, representing two generations of Colombian immigrants in London. Although speakers from any city within the *departamento* (region) of

Cundinamarca, Colombia were eligible to participate in the study, all participants from this region were specifically from the city of Bogotá. The demographic data for these participants including age, gender, and age at time of arrival in Canada are listed below in Table (2.2).

	1G Male	1G Female	2G Male	2G Female
Age				
Young (18-34)	0	0	5	5
Medium (35-51)	3	3	0	0
Older (51+)	2	2	0	0
Education Level				
High (Postsecondary)	5	5	5	5
Low (Highschool)	0	0	0	0
Age of Arrival				
Infancy/Born in Canada (0-3)	0	0	1	2
Childhood (3-8)	0	0	4	3
Early Adulthood (18-35)	3	4	0	0
Late Adulthood (36+)	2	1	0	0

Table 2.2: Demographic information regarding the twenty first (1G) and second (2G) generation immigrant speakers included in the CoSLO corpus.

Each of the two generation groups included in the CoSLO corpus has a total of ten speakers. Both generation groups were also balanced for gender, and therefore include five female and five male speakers.

Although efforts were made to ensure a well-balanced corpus in relation to the socioeconomic status and level of education of speakers, all participants were considered to be from high socioeconomic status as determined by their level of education in Spanish, or in the case of second-generation speakers, the level of education of their parents. That is, all first generation speakers, and at least one parent of all second generation speakers, had completed at least some post-secondary education in Colombia. In fact, several participants in the corpus sample had advanced degrees in engineering, finance, speech-pathology, dentistry, and accounting, among others (more details on how socioeconomic status was determined are included in section 2.5.2.2). Although information regarding their education in English was not specifically collected, many of the first-generation speakers had also completed post-secondary certificates or degrees in Canada after arriving in the country, and most second-generation speakers were

currently enrolled in university programs or planning to enroll in the near future.

2.3 Data collection

1 Participant recruitment

The recruitment of participants was carried out during the months of January 2019 through December 2020 following approval by Western's research ethics board in December 2018 (Appendix A). During this time, posters, and online posts in social media were used to advertise the need for volunteer participants. Also, being a part of the community myself, I was able to personally contact and speak with friends and acquaintances from the Colombian community. Within our protocol, mediated snowball sampling was also included. Therefore, additional participants were contacted from the initial contacts made. Many participants were willing to help with this and spread the word with their own contacts. I am very grateful to these individuals in the community, and to others within my department who were instrumental in sharing information about the study and need for participants.

Recruitment posters were posted in a number of local stores and establishments whose target clientele includes the Hispanic population of the city, including the *Latino Market*, and the *Hernandez Variety*¹. Recruitment posters were also posted at different locations at Western University to appeal to local students, and at the local indoor soccer center (BMO center) which is frequented by several men and women from the Hispanic population of London for weekly games. The success of physical posters throughout the city was at best limited, however, and online advertisements proved to be much more effective.

Online advertisements expressing the need for volunteer participants for the study were posted on online social media sites such as Facebook and Twitter. Following initial postings, individuals from the community further shared the post on sites such as Instagram and LinkedIn. Further, in addition to posting the advertisement on my personal social media accounts, I also

¹No relation to the author.

contacted local agencies and groups that are dedicated to serving the local Colombian and Hispanic communities and asked them whether they would advertise the need for participants on their sites. Their collaboration allowed me to reach the wider Colombian community in London.

2 Participant eligibility

To be an eligible participant, individuals needed to be from the region of Cundinamarca, Colombia, or a direct descendant of at least one speakers from this region.

Participants also needed to have lived exclusively in London, Ontario for the majority of their time of residence in Canada. Individuals who lived in other cities of Canada, or who lived in the United States for longer than two years prior to coming to Canada after emigrating from Colombia were excluded from the study. This was done to avoid including participants who had a significant influence from other contact varieties of Spanish such as those spoken in the United States or in other regions of Canada. For instance, one participant was excluded when they revealed during their interview that they had previously lived in Montréal for over ten years prior to arriving to London.

Participants from two generations of immigration were recruited. The first group, which represents the first generation of immigration, includes speakers who arrived in Canada, and more specifically in London Ontario, from Colombia as adults, and who had been living in London, for at least five years at the time of the interview. Several participants who had been living in London, Ontario for shorter periods of time were excluded. This was done to ensure this study's procedures aligned with previous investigations and thus ensure that the results are in fact comparable to these similar investigations. For instance, although seemingly arbitrary, the five year residence cut-off point was selected to follow the methodology used by Otheguy et al. (2007). In their study, speakers who had been living in New York for less than five years were considered newcomers, and only after the five year point were speakers considered local New York residents.

The second group, which represents the second generation of immigration, includes speakers who were born in Canada or who arrived in Canada before the age of eight, and who had at least one Colombian parent who fits the criteria for the first generation group.

Participants who had lived in London, Ontario during the majority of their time in Canada were not disqualified from participating if they currently were living in a city other than London to attend university, or if they had moved away less than two years ago. This was an important consideration, particularly in the case of some second generation speakers who grew up in London, but who recently had moved away to attend university in Toronto or other cities in Ontario.

Every effort was made to ensure that each group was balanced for gender. Therefore, both male and female speakers were eligible to participate and to be included in the study. Note that speakers were given a chance to self-report on their gender identity, and their eligibility to participate would not have been affected due to reported gender identity. Nevertheless, all participants in our study identified binarily as either male or female.

3 Data collection materials and protocol

During the time period from February 2019 to December 2020, a series of interviews were conducted with members of the Colombian community of London, Ontario following receipt of approval for the study from Western's Ethics Review Board (Appendix A).

For every interview, I was the interviewer, and as a native Spanish speaker from Bogotá, Colombia, I was able to minimize dialect accommodation effects. Furthermore, as a member of the community, I was able to enter the community more easily and I was able to relate to the experiences and background (both linguistic and cultural) of the participants in this study, which facilitated communication with participants during interviews.

Following recruitment, individuals interested in participating who met the study's criteria, were invited to meet with the interviewer for a single one hour session. Participants were given the choice to participate in person, or at distance via phone or video-calls. Including this

optionality within the study design proved to be very useful during the COVID-19 pandemic as study procedures and participant recruitment did not need to stop to accommodate for social distancing measures.

At the beginning of each meeting, an in-depth explanation of the study motivations, procedures, and requirements was given to participants, and they in turn were given a chance to ask questions and to confirm their willingness to participate. Participant consent to participate and to be recorded was collected using a physical consent form during in-person interviews, and orally collected and recorded for each participant during distance interviews.

Once participants had consented to participate, the interview session initiated. Each study session consisted of two components: a language questionnaire (Appendix C), and a semi-guided sociolinguistic interview (Appendix B).

During the first part of the session a language questionnaire was provided. The language questionnaire administered was based on a questionnaire initially created and used to investigate language use patterns and language proficiency of French-English bilingual speakers in Ontario (Mougeon et al. 1982). In the current study, the questionnaire created was administered to collect information regarding each participant's language proficiency and everyday use of Spanish and English. This data was further used as an indirect measure of the intensity of language contact as is described later in this chapter (section 2.5.2.3). The intensity of language contact is an important factor to determine in any situation where language contact effects are being considered (Bylund 2009; Kaltsa et al. 2015; Montrul 2004).

In cases where the meeting with the participant was done in person, participants received a paper copy of the questionnaire. Alternatively, in cases where the meeting with the participant was done over the phone or online, participants were sent a link to their e-mail account which allowed them to access the questionnaire in an online format. In both cases, in person and online, the questions asked and the answer options were the same, and only the modality of delivery was changed.

During the second part of the session, the semi-guided sociolinguistic interview was con-

ducted and audio-recorded using a Zoom H4N recorder, either directly in face-to-face interviews, or via phone or video-call in the case of distance interviews ².

The questions asked during the semi-directed interviews covered topics regarding the participants' occupation, hobbies, family and relationships, and their life in Canada as immigrants (Appendix B). The primary goal of the interview was to elicit as many uses as possible of constructions where overt subject pronouns could be optionally inserted. A secondary goal, however, was to gather data about participants' attitudes, opinions, and habits regarding their use of Spanish and English in order to investigate how these correlate with factors known to impact language maintenance in heritage language communities, and to gain a better understanding of their language use patterns.

4 Other methodological considerations

During the course of this study, participants were given the opportunity to participate in person or to participate in a distance interview via phone or via a video-call³. In the case of in person interviews, the Zoom H4N recorder was used in the stereo mode and the internal microphones of the recorder were used. Alternatively, in the case of interviews conducted over the phone or online, interviews were recorded using two separate tracks. The H4N recorder was connected to the audio source (either the computer or the phone audio jack) using a 1/8 inch to 1/4 inch cable, and separately to an external lavalier microphone and a pair of headphones used by the interviewer. Following these recordings, Audacity software was used to merge the two tracks into a single .wav file, and in some cases to reduce ambient noise and normalize the sound. Although the use of distance methods may reduce the audio quality of the recordings to a certain degree, this was minimized through the use of the Zoom H4N recorder, and it was ultimately not considered to be an issue in this study since only morphosyntactic and qualitative data (ie., no acoustic measurements) were considered. Further, this distance methodology

²Specific details on how distance interviews were recorded can be found in section 2.3.4

³This required an amendment to be submitted to Western's Ethic Review Board (Appendix A), which was approved in April 2019.

proved to be helpful when interviewing participants who were at the moment away from London temporarily due to school or work commitments, or who had a particularly busy schedule which did not allow them to schedule an in person meeting. The distance method was also helpful during the COVID-19 pandemic of 2020 to continue participant recruitment and interviews while maintaining social distancing measures.

Although some audio quality was lost due to the sound coming through a device such as a computer or phone, these recordings were often of better quality in regards to surrounding ambient noise since participants were able to communicate from their home or office. Some participants who met in person preferred to meet at public venues such as local cafés or shopping malls which unavoidably had higher levels of background noise and thus reduced quality of the recording. In all cases, quality of the recordings was sufficiently high to conduct the necessary analyses with very few instances of inaudible interview fragments.

2.4 Transcription and coding

2.4.1 Overview

Following the collection of interviews, each interview was transcribed using Praat and/or Phon open-access software (Hernandez and Rose 2019; Rose et al. 2006). Once all twenty interviews were transcribed, the interviews were coded according to the linguistic factors considered (See section 2.5.1).

According to Orozco (2015:20) a data sample of approximately 1000 tokens is sufficient in studies regarding variable use of overt vs. null subject pronouns to get a representative and accurate view of a population's use of subject pronouns and establish a reliable pronominal rate. Using over 3000 tokens in his data, explained Orozco (2015), yielded non-significantly different results from studies using half as many tokens following statistical analysis.

For this project our aim was to at least match this proposed 1000 token threshold for each of our two sub-populations (first and second generation speakers) in order to be able to establish

reliable pronominal rates for each. With this goal in mind, between 100 and 150 tokens were extracted for each interview for a total of 2366 tokens ($NG1 = 1239$; $NG2 = 1127$) from about 317 minutes of audio recordings. The specific time coded for each interview varied between participants according to differences in speech. That is, while some speakers used longer sentences and engaged in longer story-telling narratives earlier on, others used shorter sentences in their responses. With the speakers from the former group fewer minutes of interview were coded, while the latter group required longer portions of the interview to be coded. In the end, no less than 100 tokens from no less than 15 minutes of interview time were extracted for each speaker. Subsequently, the tokens extracted were coded to test for the linguistic and social factors of interest using Excel and R studio.

2.4.2 The envelope of variation

Following the principle of accountability (Labov 1972), I identified and extracted all instances where variation between an overt and a null subject personal pronoun (SPP) could be present. That is, all instances where there was true optionality between the use of a null or an overt subject pronoun were included. This includes all instances of finite verbal constructions where a thematic subject was required by the accompanying verb, where this subject was clearly ascertainable, and where this role was not already filled by a lexical subject (Otheguy and Zentella 2012).

Most cases where the SPP was overt were included in the analysis with the exception of emphatic clauses, or other similar cases where the use of a null subject pronoun is unlikely (See section 1.5.1.2). For instance, cases where the subject pronoun used in response to a question which requires an overt subject (32).

- (32) - *¿Quién quería bailar?*
 - *Nosotros queríamos bailar.*
 “Who wanted to dance?”
 “We wanted to dance.”

As explained in section (1.5.1.2), in the situation of the question asked in example (32), it would be ungrammatical to produce a null pronoun as part of the response. Therefore, the pronoun *nosotros* (we), is in this case obligatory, and no variation would be possible.

Another type of construction that was excluded from the analysis even though they may have included an overtly pronounced subject pronoun, was the case of relative clauses where the head of the clause was co-referential with the subject of the relative clause. For instance, in phrases such as *Julian es un niño que es muy amable* ‘Julian is a boy who is very kind’, it is unlikely to have an overt SPP in the position following the relative pronoun *que* ‘who’. For this reason, these are cases where it is considered to be a categorically null context and which therefore lie outside the envelope of variation Otheguy and Zentella (2012). However, in practice, in regards to these relative phrases, there are cases where speakers will introduce a resumptive pronoun in this position such that the phrase, *Julian es un niño que él es muy amable* ‘Julian is a boy who **he** is very kind’, can result. These constructions were also excluded, and thus are some of the cases where an overt SPP might have been produced but excluded from the data. That being said, these were extremely rare in the data analysed.

Other cases where an overt SPP was excluded from the analysis included instances where the pronoun *uno* (one) was produced. Although some variability has been documented for this pronoun (Hurtado 2005; Orozco 2015), it tends to be strongly associated with overt production of the subject personal pronoun, with reported instances where it is produced overtly in as much as 86% of the time. For this reason, and considering that this pronoun was relatively rare within the analyzed sample, instances of this pronoun were excluded.

On the other hand, in cases where the SPP was null, tokens were included if there was an available subject slot where a SPP could have been inserted. In contrast, cases where the structure did not have an available subject slot, as is for instance the case of weather-related and existential expressions, were excluded from the analysis. In these cases, the subjects of the constructions would be considered expletive. Since in Spanish, expletive subjects are rarely

overtly expressed (with the exception being some Caribbean varieties of Spanish, as discussed in Section 1.3.1.2), this represents situations where there is no optionality, the SPP is categorically null, and thus lies outside the envelope of variation (See section 1.3.1.1). Another case where a verb was found bare, and where an available subject slot might have been available but which were constructions excluded from the analysis were the cases of imperatives such as *digamos* ‘let’s say’, or *mira* ‘look’, which were used in the data as an introduction to a main statement such as in the examples below (33-34).

(33) *Digamos que recién llegué acá nos hablábamos mucho más seguido.*

“Let’s say that right when I arrived here, we spoke much more frequently.”

(34) *Mira, yo te voy a decir una cosa.*

“Look, I’m going to tell you something.”

Imperative expressions such as these above are used categorically with null subjects and therefore were excluded from the analysis.

Finally, in the current investigation we focus only on the variable use of subject pronouns and therefore, verb constructions which included a nominal subject were also excluded from the analysis (Otheguy and Zentella 2012).

1 Inanimate subjects

It is important to note that similarly to studies such as Otheguy and Zentella’s (2012) analysis of Spanish in New York City, in this analysis I chose to exclude instances where the SPP’s referent was an inanimate object (35-36) despite the fact that some varieties of Spanish do show some variability in regards to overt vs. null production of SPPs referring to inanimate subjects.

(35) *...me traducía las canciones que yo quería saber qué **decían**.*

“[He] used to translate for me the songs that I wanted to know what **[they] said**.”

(Otheguy and Zentella 2012, :53)

(36) *...venía un tren y el tren lo que hace es que él espera que esté en el túnel para tocar la bocina.*

“..a train was coming, and what the train does is that is that **it waits** to be in the tunnel to sound the horn.”

(Otheguy and Zentella 2012, :54)

In Spanish, usually null subject pronouns are used to refer to inanimate subjects as in (35). However, cases where an overt subject pronoun is used, as in (36) are observed (Cuza 2017; Otheguy and Zentella 2012). According to Otheguy and Zentella (2012) these instances, as in (36), where an overt subject pronoun is used to refer to an inanimate subject, were rare enough in their data to be excluded without affecting their results and conclusions significantly.

In the current investigation, I reached a similar conclusion following initial consideration of my data. From initial token extraction observations, I determined that when referring to inanimate subjects speakers in the sample used null subjects nearly categorically, and therefore, constructions referring to inanimate entities were excluded from the final analysis following the protocol used by Otheguy and Zentella (2012). As Cuza (2017) explains, these constructions, where a third person subject pronoun is used with an inanimate referent, are rare and mostly used in Caribbean varieties of Spanish, and therefore it is not surprising that these constructions were mostly absent within our data.

2.5 Factors considered: Variable expression of Spanish SPPs

2.5.1 Linguistic Factors

Once tokens were extracted from the 15 minute sample for each interview, each token was coded for each of the linguistic factors of interest. The linguistic factors considered were: the verb’s tense and mood, the verb’s semantic content (active, mental, stative), the type of subject pronoun used (grammatical person and number), switch of referent in discourse (same, differ-

ent, utterance initial), clause type (main, coordinated, subordinate), and polarity of the clause (negative, affirmative). The factors being considered were included as factors in the current analysis since they were found to significantly condition variation between the use of null/overt subject pronouns in Spanish in previous similar studies (Cameron 1993; Silva-Corvalán 1994b; Flores-Ferrán 2004; Otheguy and Zentella 2012; Orozco 2015). One additional linguistic factor that was considered was the position of subject pronouns in relation to the verb. This factor, although it can only be considered for instances where the subject pronoun is produced overtly, is a related relevant factor since investigations of language change focused on subject personal pronouns have noted that changes in the rate of expression of Spanish SPPs tend to co-occur with an increased preference for pre-verbal subjects (Barrera-Tobón and Raña-Risso 2016). Further, these two factors are related since the ability to have null subjects, and the ability to have overt subjects pre-verbally and post-verbally are both considered to be part of the NSP (section 1.4). A [-null subject] language such as English both disfavours the omission of subject pronouns and favours a more rigid SVO order. Convergence towards a more English-like system could therefore be observed through changes in both of these characteristics and is therefore an important factor to consider alongside the rate of expression of Spanish SPPs.

1 Subject pronoun type (grammatical person and number)

Each token is coded to indicate the person and number of the subject pronoun as shown below in Table 2.3. These aspects are important when considering the variation between null and overt subject pronouns.

For instance, in a study by Flores-Ferrán (2004), the author identified that Spanish SPPs were more likely to be expressed overtly when accompanied by verbs in their first, second, and third person singular forms, with first person plural forms being more commonly associated with null SPP usage. These results are supported by other studies as well (Cameron 1993; Avila-Jiménez 1995; Flores-Ferrán 2002), and the pattern observed ties in to the idea of “rich agreement”, which proposes that null-subject languages, such as Spanish, allow for the omis-

Person	Number	Gender	Pronoun form
1	SG	-	yo
2 (informal)	SG	-	tú
2 (formal)	SG	-	usted
3	SG	fem	ella
3	SG	masc	él
1	PL	fem	nosotras
1	PL	masc	nosotros
2	PL	-	ustedes
3	PL	fem	ellas
3	PL	masc	ellos

Table 2.3: Subject pronoun forms considered in the current analysis.

sion of overt subjects in specific contexts thanks to the rich morphology of the verb forms in these languages (Cole 2009, D’Alessandro 2015). In other words, in cases where the inflectional morphology of the verb itself contains enough information (along with the additional context available) to recover the subject of the clause, the SPP provides redundant information and can therefore be omitted (Cole 2009)⁴.

It is important to reiterate, as was discussed previously (section 1.5.2.1), that within the inflectional paradigms of Spanish verbs, not all forms are equally unambiguous and thus the subject is not equally easy to recover in all cases. For instance, in the indicative mood, third person singular verb forms which are equivalent in form to the second person singular formal (*usted*) forms, and third person plural forms which are equivalent in form with second person plural forms, are more ambiguous, while other forms in the paradigm, such as first person singular and first person plural forms, are distinctive in their morphology leading to less or no ambiguity. Take for example the verb *comer* ‘to eat’, where the third person singular form in the indicative is (*él/ella*) *come* ‘he/she eats’ and is identical to the second person singular formal form (*usted*) *come* ‘you eat’. However, when considering the same verb in the indicative, other forms such as the first person plural (*nosotros*) *comemos* ‘we eat’, or the first person singular (*yo*) *como* ‘I eat’ are unique in their morphology leading to no ambiguity with other forms.

⁴Note, however, that this idea of richness of inflectional morphology as a factor in licensing the use of null subjects remains problematic due to the fact that other null subject languages such as Chinese and Japanese do not show rich inflectional morphology (Cameron 1993).

This is further highlighted below in Table (2.4) using the example of the verb *cantar* ‘to sing’, which is recreated from Table (1.2) in section (1.5.2.1) of the introductory chapter.

Subject Pronoun	Indicative	Conditional	Subjunctive
Yo (1SG)	<i>canto</i>	<i>cantaría</i>	<i>cante</i>
Tú (2SG)	<i>cantas</i>	<i>cantarías</i>	<i>cantes</i>
Ústed (2SG)	<i>canta</i>	<i>cantaría</i>	<i>cante</i>
Él/Ella (3SG)	<i>canta</i>	<i>cantaría</i>	<i>cante</i>
Nosotros (1PL)	<i>cantamos</i>	<i>cantaríamos</i>	<i>cantemos</i>
Ustedes (2PL)	<i>cantan</i>	<i>cantarían</i>	<i>canten</i>
Ellos/Ellas (3PL)	<i>cantan</i>	<i>cantarían</i>	<i>canten</i>

Table 2.4: Present tense conjugations of the verb *cantar* ‘to sing’ in the indicative, conditional, and subjunctive moods.

Further, in the case of second person singular informal pronouns, pronouns are also coded for specificity (Cameron 1993). In his study which looked at the variable use of subject pronouns in the Spanish of Madrid, Spain, and San Juan, Puerto Rico, Cameron (1993), showed that the rate of use of overt vs. null forms of second person singular subject pronouns varied according to the function (or specificity) of the pronoun. That is, whether the second person singular (informal) pronoun *tú* ‘you’ was being used to address the person to whom they were speaking directly, or whether it was being used in non-specific statement. He found that while the two cities, Madrid and San Juan, did not differ in relation to the rate at which they used this second person singular pronoun, there were differences between the cities in relation to the rate with which they used the pronoun according to the general or specific function. While in Madrid the overt form of the second person singular pronoun was favoured when being used with a specific referent, the opposite was true in Puerto Rico. In San Juan, Cameron (1993) found that the overt form of the second person singular pronoun was favoured when being used in a non-specific context. In the current analysis, therefore, I code for person and number (Table 2.3), as well as for the specificity (general vs. specific) of use of second person singular pronouns.

Further, I want to highlight that since the sample for this study only has data from speakers from Cundinamarca, Colombia, I don’t expect to find the additional ambiguity introduced by

aspiration and deletion of coda /s/ which is a Caribbean phenomenon since this is not a phonetic characteristic of this variety of Spanish (Flórez 1961). In the Spanish of Cundinamarca, coda /s/, is consistently retained phonetically. For this reason, /s/ aspiration and deletion, and the possible additional ambiguity introduced through these variable productions are not taken into consideration in this study.

It is also important to note that constructions including the subject personal pronoun *uno* ‘one’ are not included in the current analysis. This decision was made based on a few considerations. Firstly, most studies considering the use of Spanish in the United States do not incorporate instances of this pronoun (Cameron 1993; Flores-Ferrán 2004; Otheguy et al. 2007), and therefore excluding it allows us to have a better frame of comparison with these similar studies conducted in North America. Secondly, the production of the subject personal pronoun *uno* ‘one’, although not categorically produced as overt, is strongly associated with overt SPPs usage. In fact, Lastra and Butrageño (2015: 12) state that the pronoun *uno* ‘one’ “is not really a variable form” when considering another mainland Latin American Spanish variety from Mexico, and they reported that the pronoun *uno* ‘one’ was expressed overtly in 85.1% of variable contexts. Similarly, in a study by Hurtado (2005), the pronoun *uno* ‘one’ was produced overtly 86% of the time in possible variable contexts. In addition, other studies do not include this pronoun which may impact comparability by influencing the average for third person singular constructions. Also, note that within the data collected and coded, there were limited instances of this pronoun, with only 82 instances within 2366 tokens for both generations. Therefore, in this study I exclude this pronoun, which not only might influence averages due to its limited variability, but which is also sufficiently rare in the sample.

2 Verb form (tense and mood)

In regards to verb form, in addition to grammatical person and number, the tense and mood of the extracted token verbs are coded. As discussed above in relation to grammatical person and number, it is expected that verbal forms that introduce more ambiguity correlate with

higher use of overt SPPs since the subject of the construction is less easily recoverable through morphological cues alone. This varies not just as a factor of grammatical person and number, as discussed previously, but also as a factor of verb tense. For example, in the case of verbs in the indicative, in the present, preterite, and future tenses, forms in first person singular, second person singular informal, and first person plural are more unique morphologically, while other forms are more ambiguous, such as third person singular forms which are equivalent to second person singular formal forms, and second person plural forms which are equivalent to third person plural forms. However, in the imperfect tense, an additional source of ambiguity is introduced as the first person singular forms are identical to the second person formal and third person singular forms. This can be seen below in Table 2.5.

Subject Pronoun	Indicative present	Indicative imperfect	Indicative preterite	Indicative future
Yo (1SG)	<i>como</i>	<i>comía</i>	<i>comí</i>	<i>comeré</i>
Tú (2SG- informal)	<i>comes</i>	<i>comías</i>	<i>comiste</i>	<i>comerás</i>
Ústed (2SG-formal)	<i>come</i>	<i>comía</i>	<i>comió</i>	<i>comerá</i>
Él/Ella (3SG)	<i>come</i>	<i>comía</i>	<i>comió</i>	<i>comerá</i>
Nosotros (1PL)	<i>comemos</i>	<i>comíamos</i>	<i>comimos</i>	<i>comeremos</i>
Ustedes (2PL)	<i>comen</i>	<i>comían</i>	<i>comieron</i>	<i>comerán</i>
Ellos/Ellas (3PL)	<i>comen</i>	<i>comían</i>	<i>comieron</i>	<i>comerán</i>

Table 2.5: Indicative conjugations of the verb *comer* (to eat) in the present, imperfect, and preterite tense.

Therefore, as shown in Table (2.5), the level of ambiguity varies according to not just grammatical person, but also according to verbal tense. In considering the indicative I have shown how the imperfect verbal forms introduce more ambiguity to the verbal paradigm than tenses such as the present, preterite, or simple future. It is also important to note that although not included in Table (2.5), compound tenses follow the same pattern in that they introduce variable ambiguity as a function of the verbal paradigm of the auxiliary verb, either *estar* ‘to be’ in the case of the progressive, *haber* ‘to have’ (auxiliary) in the case of perfect constructions, or

ir ‘to go’ in the case of the periphrastic future. In the current investigation, the data was initially coded considering all tense and aspect distinctions. However, due to low token counts when considering them separately, the code was ultimately restructured to include broader categories of tense (present, preterite, imperfect, and future), and eliminate aspect distinctions.

It is also important to note the tenses that were included under these four broad categories. Specifically, the “present” category included all present and present progressive forms; the “past” category included preterite perfect and pluperfect forms; the “imperfect” category included all imperfect as well as imperfect progressive forms; and the “future” category included both simple and periphrastic future instances, as well as any instances of the future perfect.

The extent of possible ambiguity also varies according to verbal mood. For example, when considering the mood of the verbs alone, it is also expected that speakers will use more overt SPPs with subjunctive and conditional mood constructions, which in comparison to the indicative have a higher level of possible ambiguity, since in these cases the first person singular is equivalent in form and ambiguous with third person singular and second person singular formal forms (Table 2.4), creating a three-way ambiguity of forms not present in the indicative present forms. Therefore, the data was coded including consideration of verbal mood distinctions for indicative, conditional and subjunctive moods. It is expected that due to the increased ambiguity in conditional and subjunctive verbal paradigms as compared to the indicative mood paradigm, overt SPPs will be more common with conditional and subjunctive constructions than with indicative constructions. Following the same assumption, is also expected that overt SPPs will be more common with imperfect indicative forms than with present, preterite, and future indicative forms.

3 Semantic content of the verb

The semantic content of the verb in each construction within the envelope of variation is also taken into consideration. Several studies, including Otheguy and Zentella (2012) and Orozco (2015), have found the use of overt vs. null subject pronouns to be conditioned by the

semantic content of the verb. In Orozco's (2015) investigation, he categorizes each verb token as a stative verb, a verb denoting mental activity, or a verb denoting physical activity. Results of his investigation showed that while stative and mental activity verbs favoured the use of overt subject pronouns, physical activity verbs favoured the use of null subject pronouns. In this investigation I follow Orozco's (2015) protocol and code each verb as either a stative, a mental activity, or a physical activity verb.

4 Switch reference

Among the most consistent findings in relation to the variable expression of Spanish SPPs, we find the switch reference factor. Studies have shown that in Spanish, overt SPPs are favoured in cases where a new subject referent is introduced into discourse (Montrul 2004; Otheguy and Zentella 2012; Silva-Corvalán 1994b). By extension, whenever the referent changes within discourse, speakers are more likely to use an overt SPP. This factor has been found to be consistently significant in studies spanning several varieties of Spanish.

Overt SPPs may be favoured when a new or changed referent is introduced into discourse to reduce ambiguity during conversation. Recall that some verb forms in Spanish, despite their rich morphology, can introduce ambiguous information into discourse due to the fact that verb conjugations are similar in certain cases depending on grammatical person and number, as well as tense and mood of the verb. These ambiguities can play a significant role when introducing a new referent subject into the conversation. In some cases, due to the fact that some conjugated forms are identical for different person and number SPPs, it may not be obvious that a new referent was introduced unless an overt subject pronoun is used.

In this investigation, this factor is taken into consideration, and I follow the protocol described in Otheguy and Zentella's (2012) work for this purpose. For each token construction, the preceding subject was considered, and when coding it was specified whether the subject in the current token is in utterance initial position (and therefore has no preceding referent), whether the referent is different from the preceding referent, or whether there is continuity

of referent. However, following initial analysis, it was determined that there were no significant differences in expression of Spanish SPPs between pronouns in utterance initial position and those where there had been a change in referent. For this reason, these two groups were collapsed into one group denoting no continuity of the referent.

5 Clause polarity (Negation)

An additional factor that is taken into account in this investigation is whether the clause containing the verb of interest is negative or affirmative. In a study by Lastra and Butragueño (2015) on monolingual speakers of Mexican Spanish, they found that negative constructions favoured the use of null subject pronouns while affirmative constructions favoured the use of overt subject pronouns. The authors note that the use of overt subject pronouns in negative constructions was rare within their data.

Similarly, in a study of language contact between English and three heritage null-subject languages spoken in the area of Toronto by Nagy (2015), the author found that in second generation speakers of Russian, negated constructions favoured a higher use of null subject pronouns. This, however, contrasted with the speech of first generation Russian speakers in the city since this group showed no differences between negative and affirmative sentences in relation to the use of overt vs. null SPPs.

Evidence from previous studies suggests that the polarity of the clause can influence the use of overt vs. null SPPs in null subject languages. Therefore, in the current study, each token is coded to indicate the polarity of the clause. Each clause will be coded as either negative or affirmative.

6 Clause type

All tokens extracted were coded according to whether they were found within a main, a subordinate, or a coordinate clause. The three main clause types considered were selected in keeping with the methodology conducted in similar studies (Orozco 2015).

In his study on the variable use of SPP's, Orozco found that clause type is a significant but weak conditioning factor in regards to the variable use of SPP's in the Spanish of the coast of Colombia, with coordinate clauses favouring null subjects and subordinate and main clauses favouring overt subjects (Orozco 2015). The author states, however, that his results are only marginal in regards to clause type, unlike results of other studies such as that by Otheguy et al. (2007) which found clause type to be among the four most significant factors to condition SPP expression in their data. Orozco explains that unlike other factors, clause type is not consistently uniform among varieties of Spanish in conditioning the use of overt vs. null SPP's. The current project then provides a valuable contribution to the understanding of the importance of this factor in regards to the variable expression of subjects in Spanish by providing data from Spanish in the bilingual context of Canada.

7 Position of pronoun in respect to verb

According to initial proposals within the generative framework, null subject languages such as Spanish include a set of clustered properties which are together known as the Null Subject Parameter (NSP) (Chomsky 1981). The Null Subject Parameter includes properties that are displayed by null subject languages. These properties include the use of null subjects, the ability to have subject-verb inversion in simple sentences, the availability of subject "long wh-movement", the use of resumptive pronouns in embedded clauses, and presence of overt complementizers in t-trace contexts (Camacho 2016). Nevertheless, of these five properties of null-subject languages, it has been suggested that the ability to have subject-verb inversion is the primary one:

Chomsky (1981) and Rizzi (1982) propose that the basic property of null-subject languages is subject-verb inversion, all the others follow from the availability of the postverbal subject position in NSLs, which licenses several grammatical structures that were not available in languages without that position (Camacho 2016, 3).

This property is observable in Spanish where, although the word order is primarily SVO

(37), we can also find sentences with VS word order (38).

(37) La chica llegó ayer.
 DET.F girl arrive.PST.3.SG yesterday
 “The girl arrived yesterday.”

(38) Llegó la chica.
 arrive.PST.3.SG DET.F girl
 “The girl arrived.”

(39) ¿Qué quieres tú?
 what want.2.SG 2.SG
 “What (do) you want?”

As we can see above, certain Spanish constructions such as interrogative (39), and sentences introducing a new referent (38), can have a subject-verb inversion resulting in a post-verbal subject and VS word order (Heap 1990).

Recall now that in cases where there is incomplete acquisition or attrition of Spanish by bilingual speakers in the United States, erosion of uninterpretable features at the discourse-pragmatics level which results in the increased use of overt subject pronouns in Spanish (in contexts where a null subject pronoun would be more appropriate), has been observed (Montrul 2004). Along with the increased use of overt subject pronouns, in some of these same cases relating to heritage speakers displaying signs of incomplete acquisition or attrition, it has also been noted that a more rigid word order may develop (Barrera-Tobón and Raña-Risso 2016; Silva-Corvalán 1994b).

In the study by Silva-Corvalán (1994a) on the population of Mexican-Americans living in Los Angeles, she finds that second and third generation speakers display a more rigid SVO order than first generation speakers. This, according to Silva-Corvalán (1994a), may indicate that Spanish-English bilinguals are undergoing erosion in regards to the pragmatic factors that condition the inversion of SVO order to VS order which are typical in monolingual Spanish in cases such as the introduction of a new referent, resulting in the observed more rigid SVO order.

Alternatively, Camacho (2016) also highlights that in Caribbean varieties of Spanish, such as the variety spoken in Dominican Republic, we can observe an increase in frequency of use of overt subject pronouns, and a more rigid SVO order. He suggests that these varieties of the Caribbean may be displaying an evolution by which they are becoming a non-null-subject language system.

In this thesis, I consider this factor by investigating the use of pre-verbal and post-verbal subject pronouns in each generation and examining whether a significant intergenerational difference exists in the population of Colombian Spanish speakers living in London, Ontario. I code each instance where an overt SPP in a variable context is used as pre-verbal or post-verbal according to its position in relation to the verb. If the results indicate that second generation Colombian-Canadian Spanish speakers use pre-verbal subjects with higher frequency than first generation speakers, this may reveal a more rigid SVO order in this generation, which could then serve as evidence of language contact effects. Since there is very low incidence of contact with Caribbean Spanish varieties in London, Ontario, this pattern would suggest effects of language contact with English.

2.5.2 Social factors and extra-linguistic factors

Along with coding for linguistic factors which may condition the use of variable subject pronouns in Spanish of Colombian-Canadians, I also include an analysis of related social factors that may have a role in conditioning this variable. This is done since language in itself is a social, communicative phenomenon and studying language use in isolation of the related social context in which it occurs would make little sense. Further, since the focus of the current investigation is to look at the possible effects that emanate from a situation of intense language contact, consideration of the relevant sociodemographic factors is an essential component of the investigation. As stated by Gillian Sankoff (2001:3), “Linguistic outcomes of language contact are determined in large part by the history of social relations among populations, including economic, political, and demographic factors”. Therefore these factors can’t be excluded from

the investigation when we consider the variable in a novel socio-historical context. However, it is important to note that I consider the social factors through a combination of quantitative and qualitative approaches. In the sections that follow, I describe exclusively the social factors that were considered within the quantitative analysis. For description of other social considerations take into account in this study, through a qualitative approach, please refer to section 2.6 of this chapter. For the purposes of this investigation, the social factors that were included in the quantitative analysis were speakers' gender, immigration generation, and self-reported bilingual proficiency as a measure of language contact intensity. In addition, interview modality was considered as a potentially important extra-linguistic factor.

Although age is a factor usually considered in investigations of language variation, this factor is not considered directly in this investigation. In relation to language change, an increase in the use of an innovative language feature by younger speakers may be an indication that the feature in question represents a linguistic change (Trudgill 1974; Labov 1994). This refers to the apparent time construct which states that once an individual acquires certain language features during their youth, these features remain stable throughout their lifetime and thus, their speech presents a mirror into the language use at the time at which that individual acquired the language (Bailey et al. 1991). In the case of the speaker population being considered, if younger speakers are shown to use novel features consistently, which differ from those of their parents' generation, this may indicate that these features represents a change in the use of language of the new generations and of the language into the future. However, although age effects may be relevant in studying the variable use of Spanish SPPs, within the specific population being studied, generation is a better factor to use in measuring these effects. Therefore, within this study, an analysis of generation, which correlates closely with speaker age since second generation speakers in my sample are all younger than first generation speakers (and often the children of the first generation speakers in the sample), is included in place of an analysis specific to speaker's age. In addition, previous studies have not found a relationship between age of participants and their variable use of SPPs in bilingual settings, but some studies have

found an effect of generation on speakers' use of subject personal pronouns (Silva-Corvalán 1994b; Montrul 2004). Therefore, generation seems to be a more relevant related factor to consider. It is important to note that all participants in this study were adults at the time of the interview, ranging from 18 to 67 years of age.

1 Immigration generation

In the current study, two groups of participants were studied, where each group was composed of speakers from one of two generations of immigrant Colombian Spanish speakers.

The first group of speakers included in the study represented the first generation of immigration. This included individuals who arrived to Canada from Colombia as adults. For all of the speakers in this group, English represents a second language acquired later in life. A total of ten speakers ($N_M = 5$; $N_F = 5$) were included in this group.

The second group, a group with also ten speakers ($N_M = 5$; $N_F = 5$), represented the second generation of immigration. This included individuals who arrived in Canada from Colombia at the age of eight or younger, or who were born and raised in Canada (more specifically in London, Ontario) to at least one Colombian parent who would qualify as a first generation immigrant speaker.

The age of eight was selected as the threshold point in our division of generations to ensure that speakers in this group were introduced to English within the Canadian context prior to achieving full acquisition of the first language (Bylund 2009; Silva-Corvalán 1994b). According to Bylund (2009), the age of twelve represents a critical age in the acquisition of a native language variety. After the age of twelve, the extent and type of transfer and attrition effects that immigrant speakers will experience when exposed to intense contact with a foreign language diminishes (Bylund 2009).

Further, in the study conducted by Silva-Corvalán where she investigates the Spanish use of Los Angeles based Mexican Americans, she categorized second generation speakers as those individuals who arrived in the United States before age eleven (Silva-Corvalán 1994b). She

explains that prior to this age the full competence of the first language has not solidified. A similar argument is made by Montrul (2012), who notes that we see a solidification of the first language system in children between the ages of eight and ten. Therefore, following the protocols in these studies, for the current analysis the age of eight was selected as the cut-off point in selecting participants for the second-generation group.

2 Gender and socioeconomic status

In this study, to get a representative sample, an effort was made to have a balanced group of participants in relation to gender. This means that within each generational group, an equal number of men ($N_{TOTAL} = 10$; $N_{1G} = 5$, and $N_{2G} = 5$), and women ($N_{TOTAL} = 10$; $N_{1G} = 5$, and $N_{2G} = 5$) was included.

It is also important to note that participants were given a chance to self-identify in regards to their gender at the time of the interview. Nevertheless, in the case of this study, participants identified binarily as either male or female, and therefore these are the only two gender identities included in this study. Further, having this binary division allowed us to have a more direct comparison with similar previous studies which found gender to be significant in relation to the use of variable use of Spanish SPPs. In a study by Shin and Orozco (2013), they found that the variable expression of Spanish SPPs in New York City varied as a function of speakers' gender and socioeconomic level. Specifically, their results suggested that Hispanic women in New York, as well as Hispanic speakers of higher socioeconomic status, led the increase in the use of overt SPPs. The authors point out that women leading a linguistic change is not surprising, and in fact aligns well with the pattern observed in studies of linguistic change in monolingual communities (Trudgill 1974). Women are often the members of the community who have closer contact with younger speakers, including their own children, and therefore are more likely to be the ones to pass on certain variations in linguistic forms to these new generations.

Shin and Orozco (2013) also considered the influence of speakers' socioeconomic status in

regards to the variable expression of Spanish SPPs. They found through their study that speakers in New York with higher socioeconomic status were more likely to have higher frequencies of overt SPP. This, however, was contrary to expectations since this feature is often associated with lower prestige in bilingual communities.

The higher rate of use of overt SPPs can be seen as a symbol of lower prestige in the Spanish-speaking world. This is mentioned in Garrido (2007) where she notes that when Colombians were interviewed regarding their attitudes towards different varieties of Spanish, it was the Spanish of Bogotá that was considered the most prestigious variety, while the Spanish of the Caribbean and Pacific coasts were considered among the least prestigious varieties. Note that the Spanish of Bogotá, since it is a Latin American mainland variety, is characterized by a lower frequency of use of overt SPPs as opposed to Caribbean varieties, and by extension higher frequencies of use of SPPs are also considered less prestigious. Shin and Orozco (2013) note that the contradictory results in the New York study, which suggest that speakers of higher socioeconomic status leading the change in increase of overt SPPs may be due to the fact that more affluent members of the Hispanic community in New York have smaller social networks and are more susceptible to influence from English, using a more “anglicized” Spanish, and thus revealing an opposition between prestige monolingual and prestige bilingual features. It is also possible, in my opinion, that in the case of the bilingual context of New York City, features of Caribbean Spanish may enjoy a higher prestige than in monolingual contexts since these varieties have been present in these regions of the United States for longer and therefore may act as a model for later-arriving groups. This may be further reflected nowadays through the current popularity and prevalence of Caribbean Hispanic music such as *reggaeton*, a genre which originated in the Hispanic Caribbean and which often features singers and linguistic features of Puerto Rican and other Caribbean Spanish varieties.

In the current analysis, the socioeconomic level of participants is not included. Although this is a factor that has been found to be significant in previous research, our sample is composed of high socioeconomic level speakers exclusively, as determined through their level of

education in Spanish. The socioeconomic level of speakers in the sample was determined through consideration of the level of education attained by speakers in Colombia as it relates to their use of Spanish, since it is their Spanish use that is the focus of the investigation. For second generation speakers, their socioeconomic level was determined according to the education level in Spanish of their Hispanic parents since this represents their main linguistic input during their acquisition of Spanish. Following participant recruitment, it was found that, coincidentally, all of our speakers were classified as being from a high socioeconomic background. All first generation speakers had completed at least some post-secondary education in Colombia, with many of them in fact holding advanced degrees in fields such as engineering, finance, dentistry, and speech-pathology among others. In addition, the parents of second-generation speakers similarly had completed at least some post-secondary education in Colombia.

It is of note that current occupation of participants in the study would not be an appropriate measure of socio-economic level in relation to their use of Spanish since most immigrants, specially those of the first generation, do not have jobs which align with their formal education or training (Guardado 2002). Upon arrival in Canada and the United States, the academic degrees and professional training that immigrants received in their country of origin are often not recognized. This means that in order to work in their new community, immigrants end up finding jobs which have nothing to do with their level or field of education. For instance, within my own sample, one speaker from the first generation group (1GM3) is currently working as a construction worker while their professional training in Colombia was in financial services and they worked as a bank's general manager. This is by no means a singular occurrence and is in fact the case for many of the first generation speakers in my sample. Therefore, in this study it is the speakers' level of education in Spanish which is considered since this is more likely to reflect their usage of Spanish than their level of education in English or their current occupation.

3 Language contact intensity (Length of residence, bilingual proficiency, and language use habits)

In section 1.3, it was shown that the English-Spanish situation of language contact in Canada is different from that in the United States due to the socio-historical differences surrounding the language contact situation in each country. However, in this study, I consider as well the ways in which different levels of language contact intensity with English experienced by Spanish speakers influence their use of Spanish SPPs. In order to do this I consider a number of different factors as measures of contact intensity. Namely, I consider the length of residence, which applies mostly to first-generation speakers, the degree of bilingual proficiency, and the extent of daily use of English and Spanish by speakers. These three considerations all represent different ways of assessing the intensity of language contact which speakers in the sample experience. In this investigation, not all of these measurements are equally applicable to both generations, as will be explained in further detail in the sections to follow, and for this reason, I consider that it is important to use various measures of language contact intensity to determine the influence that contact intensity can have on speakers' use of Spanish SPPs.

Length of residence

It was important for this study that participants had been residing in London, Ontario for at least five years. This cut-off point, although chosen somewhat arbitrarily, ensured that participants had been exposed and immersed in the Canadian community for a sufficiently extended amount of time to not be considered newcomers. It was also chosen in considering similar studies, where a length of residence (LOR) of five years was considered the cut-off point for speakers to no longer be considered newcomers (Otheguy et al. 2007).

Following participant recruitment, in the end, a large majority of participants had been in Canada, and specifically living in London, Ontario, for over thirteen years at the time of their interview. In fact, only one participant, a first-generation female speaker, had been in Canada for less than ten years (IGF3 who had been in Canada for 9 years). In addition, due to the design of the study which includes only adult participants, second generation speakers in the

sample, who arrived as children under the age of seven or were born in Canada, would have been residing in Canada for at least ten years at the time of the interview (assuming an arrival age of eight, and participation in the study at the age of 18), and in fact, the actual LOR in Canada for second generation participants ranges from 13 to 18 years. For first generation speakers, the LOR ranges from 9 to 20 years. This 11 year range provides sufficient variation to consider LOR in respect to subject pronoun expression for this generation group. This factor, however, is not considered for second generation speakers because they do not have enough variation in relation to their length of residence with a range of only five years at most between them.

Length of residence is related to the variable of intensity of contact since the longer an individual has resided in English-speaking Canada, the more exposure the individual has had to English in a majority context, and the more personal relationships they may have developed with both Anglophone and Hispanic members of the community of London, Ontario. For first-generation speakers this will be more dependent on LOR, but for second generation speakers who have lived in the country since early childhood, the intensity of language contact may be better assessed as a function of their language use habits for each language. This, was addressed through a quantitative perspective, and more thoroughly through the qualitative analysis (see section 2.4).

Bilingual proficiency

The intensity of language contact is also connected to the measure of speaker's bilingual proficiency. From a bilingual perspective, it is considered that speakers themselves are the locus for language contact. Therefore, speakers with a higher proficiency in both languages are more likely to reflect, through their speech, transfer or convergence effects resulting from the contact between their two languages' grammars. Kaltsa et al. (2015) explain this through the notion of the activation threshold hypothesis, which states that once a grammar from an L1 has been fully acquired, a feature of this L1 will go into attrition only when there is a competing feature in the L2 which is used with a frequency that surpasses that of the L1 feature, and

which overcomes a certain threshold to become activated. At this point, the L1 feature may show signs of attrition and may begin to be replaced by the L2 competing feature. In the case of the current investigation, since subject pronouns in English are in many ways equivalent to subject pronouns in Spanish, we consider the use of English SPPs to be in competition with Spanish SPPs. For instance, both Spanish and English are SVO languages and therefore the position of the subject pronouns may be in the same syntactic node. The two languages then are similar at the surface level, while having distinct characteristics at underlying levels such as the discourse-pragmatic factors that condition the overt expression of subjects in Spanish but not in English. These surface similarities and underlying differences then make it possible to have language transfer effects according to the activation threshold hypothesis. If English is used with sufficient frequency, it is then expected that the discourse-pragmatic features of the use of subject pronouns in Spanish might begin to erode and be replaced instead by features of English SPP's use. For this reason, the level to which participants were proficient in both English and Spanish, as opposed to proficient in just one of the two languages, is considered as an additional indirect measure of language contact intensity at the speaker level.

Therefore, speakers' bilingual proficiency, as measured through the language questionnaire administered (Appendix C), is also included in the analysis. This questionnaire includes self-reported measures of proficiency in English and Spanish in regards to speakers' reading, writing, listening, and speaking abilities. Participants' proficiency in English and Spanish was calculated as the average of their self-reported scores for each of speaking, listening, writing, and reading, abilities in each language. For each of the four abilities, speakers rated their proficiency from "no difficulty", which was coded as equivalent to 4 points, to "a lot of difficulty", which was coded as equivalent to 1 point. Therefore, for each language, according to speakers' self-reports, speakers were rated on a scale from 4 to 16 points where a score between 4 and 7 represented speakers of low proficiency, a score between 8 and 12 represented speakers of intermediate proficiency, and a score between 13 and 16 represented speakers of high proficiency. Speakers were categorized into two groups, those with a high level of bilingualism,

which means they are speakers with an advanced proficiency level in both English and Spanish, and those with an advanced proficiency in only one language (whose language proficiency is limited to a low or intermediate level in one of the two languages).

Following recruitment, it was determined that variation in levels of bilingualism was only present in the first generation group, where some speakers have not yet acquired a high level of English proficiency. Since participation in the study required ability to communicate at a fairly high level of Spanish proficiency, speakers who are Spanish dominant (mostly first generation participants), but not English dominant (mostly second generation speakers with an insufficient level of Spanish speaking proficiency), were eligible to participate. Therefore, due to the fact that the interviews were conducted in Spanish, participants in the second generation group were all advanced in both their level of Spanish and English, and for this reason, this factor of bilingual proficiency could not be included in the analysis for this generation as there is no variability to draw a comparison between speakers. Nevertheless, it is not the case that all first generation speakers have low proficiency in English, and therefore, considering this as a factor within the investigation for first generation speakers can help in determining intra-generational effects due to language contact intensity as measured through language proficiency.

In contrast, second generation speakers are expected to vary in relation to their use of Spanish, and their Spanish language input (both quantity and quality), which would likely have had an influence in their acquisition of Spanish. For instance, it is expected that most, or all, of the input second generation Hispanic children receive is within their home and family contexts, and that each second generation speaker will therefore vary greatly in their experiences learning and using Spanish. For second generation speakers in this study, although language proficiency or length of residence factors may not provide an appropriate source of variation for comparison in relation to their level of language contact intensity, their reported experiences with and regular use of English and Spanish serve as a better source of information to determine the effects resulting from language contact intensity.

Language use (English-Spanish)

An analysis of speakers' self-reported language use habits as a measure of language contact intensity was also included. Speakers were asked to indicate the frequency with which they use Spanish in their daily lives within different contexts on a percentage scale of 5 levels: never (0% - 10%), rarely (10% - 39%), sometimes (40% - 59%), often (60% - 89%), and always (90-100%). Four main contexts were addressed in the questionnaire: home and immediate family context, local social context, international social context (including family abroad), and professional context (including both work and school contexts). For each domain there were between one and nine questions aimed at identifying the frequency of use within that domain (Appendix C). The actual number of questions for each domain varied according to generation and individual speakers since several questions did not apply for each generation. For instance, for the home domain there were nine questions included in the questionnaire, however, several questions addressed language use with a spouse or children which did not apply to any of the second generation speakers. In contrast, some of the questions in the home domain also addressed language use with parents or other members in the home, which did not apply to first generation speakers who no longer live with their parents. In this case, for many, but not all, first generation speakers their language use with parents was more characteristic of their language use within the international social context since in many cases their parents remain in Colombia. In all cases, speakers were assigned a score between zero and four for each question according to their response: a score of zero corresponded to a 'never' response, a score of one corresponded to a 'rarely' response, a score of two corresponded to a 'sometimes' response, a score of three corresponded to an 'often' response, and a score of four corresponded to an 'always' response. From these scores, a cumulative per context score was attributed by dividing the sum of the scores attributed to each question by the number of applicable questions for each speaker. Finally, for a total measure of language contact intensity, a cumulative score across all contexts was calculated by adding up the raw scores and dividing by the total number of questions across all contexts that each speaker answered. This procedure resulted in final scores for each speakers in percentage form for each domain, and for a measure of use of Spanish in

their daily life across domains. Speakers in each generation were then divided according to these scores as being in a situation of “high” language contact intensity (less everyday use of Spanish across all domains), or in a situation of “low” language contact intensity (more frequent use of Spanish across all domains). To create this division between speakers in “high” vs “low” situation of contact, I first determined the median value of language contact intensity for each generation. Speakers whose score of Spanish language use was equal to or greater than the median value were considered to be in a situation of “low” language contact intensity with English, while speakers whose score was lesser than the median value were considered to be in a situation of “high” language contact intensity. This process resulted in two groups in each generation: a high intensity group and a low intensity group. This factor was then included in the multivariate regression analysis. It is important to note that each generation was considered separately due to clear differences between generations, where first generation speakers had, in general, higher cumulative scores (representing lower contact intensity) than second generation speakers, as was expected. Therefore, I used this measure to distinguish intragenerational differences rather than to look at language contact intensity in general across both generations since doing so would have resulted in a generation effect rather than a true language contact intensity effect.

Spanish language use measures were also indirectly addressed through a qualitative analysis of the speakers’ experiences with both languages, including the different sources of input they receive, their habits of language use in different contexts, and their access to language learning resources (this analysis is further described below in section 2.6.1).

Finally, quantitative data including quantified measures of language use in each domain, qualitative data on language use gathered through the qualitative thematic analysis and quantitative data relating to the variable use of Spanish SPPs for individual speakers is used to conduct a mixed-methods analysis in order to examine possible effects of second generation speakers’ experiences and language habits with their two languages on their variable use of Spanish SPPs.

4 Interview modality

During data collection, speakers were given the opportunity to participate in the sociolinguistic interview via three different modalities: over the phone, over video-call, or in-person. Interview modality could have a potential impact on the production of subject personal pronouns in Spanish. For instance, interviews conducted over the phone involve voice communication only. Given that communication between two individuals involves not only the spoken components but also extra-linguistic components such as body language and facial gestures, the question arises whether this will significantly impact the quality of the conversation. According to Shuy (2001), phone interviews in social science research can be less successful than in-person interviews in achieving more natural conversations. Shuy explains that during in-person interviews there is a greater number of interactions that lead to more natural communication between the interviewer and the interviewee such as politeness checks, visual cues, small talk, and other “asides in which people can more fully express their humanity”. These aspects which are affected in phone interviews can influence the comfort of speakers during the interview, the complexity of their responses, and length of utterance, which can in turn influence the production of subject personal pronouns. For instance, according to Travis (2007) longer narratives tend to correlate with greater subject continuity and a higher use of null Spanish SPPs.

Some of these concerns could be diminished by using a video call, but as many of us discovered during the 2020-2021 year, video calls depend largely on the quality of internet connections and there can be lags in sound or image, and problems with connectivity on either side. Therefore, it was deemed important to take into consideration the modality of the interviews as a possible influencing factor in the current investigation.

2.6 Factors considered: Spanish language maintenance

Although the main focus of the current investigation is to determine the factors that condition the expression of subject personal pronouns in the Spanish of first and second generation Colombians living in Canada, a second consideration is the analysis of speakers' attitudes and language use habits which may or may not contribute to the maintenance of Spanish in the community.

This section is a valuable addition as it provides further information regarding the situation of Spanish-English contact within which the speakers in the sample interact every day. Also, by investigating the situation of language contact from a qualitative point of view, we can get an indirect measure of the intensity of the situation of language contact, which can in turn help us to better understand any evidence of language change, or any effects of language contact seen as a result of our analysis of the variable expression of Spanish SPPs. In my opinion, considering that every situation of language contact is different due to the socio-historical factors at play in each individual situation, an investigation which seeks to examine the possible effects of a language contact situation needs to be further based on and informed by an analysis of the community and the social factors that surround the community (or communities) being considered. It is this that I aim to achieve through the current analysis of language maintenance factors as they relate to our speakers' attitudes and habits of language use.

2.6.1 Language maintenance factors

The maintenance of a minority language by a community within a majority language context has been shown to be impacted by a number of different factors including: the community's own attitudes towards their own language; the perceived prestige of the language within the larger community; the contexts in which the language is used by speakers, which is influenced by its perceived prestige and at the same time influences the input younger speakers receive as well as the opportunities speakers in general have to practice the language; the efforts of the

local authorities and of the minority community itself to create opportunities to cultivate and practice the language; and the influx of new monolingual speakers to the community among others.

According to Pauwels (2016) the investigation of when, where, and with whom a minority language is used can help to better understand the status of the minority language in the community and the likelihood that language shift might occur. For instance, Pauwels (2016) explains that women within a minority language community might begin to lead a language shift process if they consider that through their own use of the majority language they are more likely to increase their own and their children's prestige status and social and financial success, downgrading the importance of minority language maintenance in the process. This shift is influenced by the larger community's attitudes and perceptions of their attitudes towards the minority language and the minority language community itself.

In this study, I will be considering the following factors as they relate to minority language maintenance through a qualitative thematic analysis:

1. Use of Spanish in social, professional contexts, and in public in Canada (at home, at work, with friends) - Perceived importance of Spanish for speakers themselves.
2. Attitudes towards inter-generational language transmission - Perceived importance of Spanish for future generations.
3. Attitudes toward endogamous/exogamous relationships/marriages for the speakers themselves or their children.
4. Access to/knowledge of community-led resources/activities that contribute to the maintenance (learning and practice) of Spanish.
5. Cultural association with Hispanicity or Colombia.

Each of these factors, which will be described in further detail in the sections that follow, has been noted to play a significant role in the likelihood that a minority language is maintained

inter-generationally (Pauwels 2016; Fang 2017). Through this analysis, I hope to gain a better understanding of the vitality of the Spanish language in Canada and the manner in which Colombian speakers use Spanish and English in London, Ontario.

1 Use of Spanish and perceived importance of Spanish for speakers

For instance, speakers may consider that using and learning the minority language is a valuable asset since it helps in securing better job opportunities or higher pay as a bilingual employee. Alternatively, speakers may consider the use of the minority language an important element in their social life as it allows them to communicate with close friends and family. However, as noted above, speakers may instead consider that it is more important to learn and use the majority language than the minority language to increase their chances to improve their socioeconomic status. In this case, we are more likely to see a quicker process of language loss and consequently language shift within two or three generations. This can be due to perceived financial opportunities associated with improving the use of the majority language, or due to perceived discrimination of the minority language in the workplace, at school, or in the larger community.

The perceived attitudes of the larger community towards the use of the minority language may also influence the attitudes and use of the minority language by its speakers. Minority speakers may reduce the contexts in which they use their language if they perceive it is unwelcome in the larger community. Instances of discrimination towards individuals speaking a minority language in public or at work are, unfortunately, not uncommon in the context of the U.S. and Canada, and thus some speakers may feel more comfortable speaking in the majority language in public, and encouraging their children to use the majority language instead in more contexts. For instance, in St. Thomas, Ontario, there was one reported instance of a man who attacked a Hispanic teenager at a strip mall parking lot when he heard them speaking in Spanish (Dubinski 2017). Although fortunately this is an extreme and relatively isolated case in Canada, other instances of discrimination, and more subtle pressures from work or school

to abandon the minority language in favour of the majority language are common in both the United States and Canada (Guardado 2002). This contributes to language shift in minority communities since actions taken to avoid instances of discrimination can have the side effect of further reducing the opportunities for children in these minority communities to hear, receive input, and practice the language.

In order to analyse these topics within the current analysis, the following questions were asked during the interviews: “Do you use Spanish at home?” “Do you use Spanish with your Hispanic/bilingual friends?” “Do you use Spanish at work/school?” “Do you prefer speaking Spanish or English in public?” “Do you maintain a close relationship with your Hispanic family abroad?”. In addition to these questions, some follow-up questions were occasionally asked. Specifically, following the question of whether Spanish was the language spoken at home, participants were asked who in their household they spoke Spanish with and how frequently. These were designed to address the importance attributed to the use of Spanish for individual speakers, as represented through their daily use of the language. When analysing the data, I divided the relevant codes that were identified and that were used to code the data into four categories according to the contexts they addressed: the home environment; social interactions; work environment; and in public. Within each of these, a number of different codes were used. For comments relating to the home context, the following codes were used: Spanish at home as a rule; Spanish with older generations only; mix of Spanish and English. For comments relating to the social context, the following codes were used: Spanish with friends; English only, or mix of Spanish and English with friends; depends on whether I want to be understood; Spanish only with friends who are new to Canada. When considering the work environment, the following codes were used: English only at work; Spanish at work informally; Spanish at work with benefit; sense of discrimination or negative attitudes about Spanish. Finally, the codes used to analyse speakers’ responses regarding their choice of language when in public were: Spanish preference; Spanish - no reason not to; Spanish - more comfortable; Spanish - so others won’t understand - depends on who I am with; English preference; English - to avoid

looks/discrimination; English - to be respectful to others; other (including limited comments about topic-dependent use of language).

2 Attitudes towards inter-generational language transmission

In connection with the previous topic, the community attitudes towards their own language can also influence parents' desire for their children to learn the minority language. For instance, while this is often not the case with languages such as Spanish, in the case of stigmatized languages such as Creoles, African American Vernacular English, and some Indigenous languages, parents may in fact consider that learning the minority heritage language is not important or even harmful to their children's future financial and professional success (Truman 2019; Lane 2010; Siegel 2007). For example, in an investigation on the attitudes of Mayan women in the Yucatan Peninsula, it was found that while Mayan women valued the use of their language, they also considered the learning of Spanish as more important for their children's future financial success and therefore put in minimal effort towards the goal of Mayan language transmission (Truman 2019). Therefore, pressures to communicate in a majority language can influence parents' efforts to pass on the minority language to their children if this is not considered a priority. For this reason, in the current analysis we consider this an important factor to include.

In order to analyse these topics within the current analysis, the following questions were asked during the interviews: "do you think it is important that the children of Hispanic families in Canada learn to speak/read/write Spanish? Why or why not?". These questions were designed to target the speakers' general attitudes towards the importance of maintaining the language in future generations and for their own children. From the responses provided by speakers to these questions, the data was coded using the following codes: bilingualism considerations; future opportunities (professional, romantic, etc); family considerations; culture, identity and pride; too difficult (Non-Hispanic parents, too much English, pressure to learn English); French importance; only speaking Spanish with first generation speakers; parents'

attempts to establish Spanish use in the home.

3 Attitudes towards linguistically endogamous/exogamous relationships

Another factor that may contribute to the inter-generational transmission of the minority language includes the attitudes towards linguistic endogamous and exogamous relationships. According to Pauwels (2016), exogamy, which refers to relationships with a partner that belongs to the majority (ethno)linguistic group or another minority individual with whom the majority language would be the main language for communication within the relationship, tends to be a major factor in driving language shift. This is due to the fact that marriages and spousal relationships determine to a great extent the language that is used in the home environment. Therefore, linguistically exogamous relationships, where a minority language speaker resides with a majority language speaker, bring the majority language, in this case English, into the home context and further reduce the linguistic input from the minority language for the children in the household.

This is an important consideration since the home context can be one of the few contexts in which speakers from a minority language community, in this case Spanish-speakers in Canada, practice and use the minority language. Pauwels (2016) explains that it is less likely that a minority language will be successfully acquired by second generation speakers in a household where only one parent speaks the language. However, it is important to note that it is not a given that the minority language will not be transmitted to the children in the home in these cases. That is, in some cases children growing up within a household where the parents are from different cultural or linguistic backgrounds can successfully acquire two or more languages and gain in addition a wider knowledge and appreciation for different cultures. Nevertheless, this case is less common and in the majority of cases the children acquire the minority language incompletely or do not acquire it at all. This will vary greatly and will depend on the quantity and quality of input that the child receives from all sources including the home and the community.

In our analysis, since exogamy is generally considered a risk factor favouring language

shift towards the majority language, this is an important factor to consider. Although it is not a clear determinant of language transmission, it will inform us about the existing risk factors observable within our sample of speakers, which may inform us about the possible status of Spanish in the community of London, Ontario. When analysing the data, the questions I used to address investigation of this subject were: “Do you think it is important to marry someone who is Latin@/Hispanic? Why or why not?” These questions were also slightly adapted for each speaker. For instance, in the case of first generation speakers with children, one question that was included was “Do you think it’s important that your children marry someone who is Latin@/Hispanic?”. From the responses provided by speakers the data was coded using the following codes: Yes, no, unsure/mixed feelings, cultural considerations, other considerations being more important, comprehension considerations, family considerations.

4 Access to language resources and community involvement

Previous research regarding the maintenance of minority languages has cited the availability of resources, and presence of a community committed to the maintenance of the minority language as important factors in favouring the maintenance of a minority language. According to Fang (2017), in fact, when a group works to ensure that their language is used in settings such as education and places of worship, this strongly increases the group’s chances of achieving success in maintaining their language. Fang (2017) goes on to explain the important role that policy put in place by government authorities can play in supporting the maintenance of minority languages. For instance, by helping minority communities access government services and resources, as well as formal education, in their own language. In Canada, this is seen in regards to the efforts put forth by the government to support maintenance of French, but little is seen in regards to maintaining other minority and immigrant languages in the country. Therefore, in this investigation, I focus on the effort of the minority-language community itself to create and provide Spanish learning resources and communities to help the next generation of Hispanics to acquire the languages. Providing educational resources and forming groups

throughout the community, other than the home, where young speakers can come to use and practice the language helps in language maintenance efforts by providing further sources of language input which does not just add to the quantity of input received by young Hispanics, but also to the quality of the input as there is more variation to the situational contexts where they can use their language. This helps in increasing their vocabulary as they need to apply their language to a wider range of contexts and situations, and their range of registers as they may be addressing individuals with different levels of formality (i.e., addressing other young Hispanics vs. addressing a religious group authority figure).

In this study, therefore, I consider the availability of resources and of community groups (or any other contexts where speakers' may find opportunities to use and practice Spanish) as a relevant factor in determining the likelihood that Spanish will be maintained with future generations. In order to assess this factor I included in the interview the following question: "Are there any resources and/or groups available to help you/your children learn and practice Spanish?". This question was in some cases followed up by asking more specifically: "Did you/your children have any opportunities outside the home to use or practice Spanish?". This was done to ensure that any possible resources, even if just informal in nature, were addressed by speakers. From the responses provided by speakers to these questions, the data was coded using the following codes: Yes; no; only at home/with family; and yes, but I did not use them.

5 Cultural association to Hispanicity and/or Colombia (2nd Generation speakers)

The ties of second generation speakers to their heritage culture has been cited by some authors as being a significant factor in language maintenance. For instance, in a study by Schecter and Bayley (1997), the authors investigated the experiences of four Mexican families living in the United States and examined, through a qualitative analysis, their attitudes towards their culture, and their heritage language. Ultimately, through their qualitative analysis, the authors found that the participants in their study associated proficiency in the heritage language, in this case Spanish, with cultural identity, and therefore loss of the heritage language was associated

with loss of cultural identity. This point, accentuating the strong inherent link between language and culture, and its importance in relation to maintenance of a minority language is also mentioned in a study by Guardado (2002). In his study, Guardado (2002) investigated a group of Hispanic families living in Vancouver, Canada, to determine the factors that parents of these families considered to be important for language maintenance. Through his investigation, he determined that the motivation of children to learn and maintain the heritage language was an important factor. That is, specifically, he found that among the four families interviewed, the two families that had been most successful at helping their children maintain the Spanish language and reducing language loss after arrival to Canada had emphasized the links to the Hispanic heritage culture for their children, motivating them to foster a connection in identity to their cultural roots. This strategy proved to be more successful than strategies focused on imposing rules regarding the use of the heritage language in the home. Therefore, children who felt a close connection to their heritage culture and incorporated this culture more strongly within their identity were more successful at maintaining their heritage language. One possible reason for this, explains Guardado (2002), is that imposing rules regarding the use of the heritage language could instead create a feeling of resentment towards the use of the language. Instead, making children become interested in the music, family history, cultural traditions, among other cultural aspects, while not necessarily enforcing the use of Spanish in the home, seems to be a stronger predictor of language maintenance in heritage speakers' homes.

These findings of previous studies, in addition to other literature, reiterate the importance of cultural identity formation for second-generation speakers (and other generations after that) in increasing the likelihood that the heritage language will be maintained (Pauwels 2016; Fang 2017). For these reasons, an investigation into the sense of association of the second generation speakers in the sample of the CoSLO corpus to their cultural heritage is an important factor to consider in the current investigation. This can help determine the motivations of these speakers, who were successful in maintaining their heritage language, Spanish, and the likelihood that they will also be successful in passing it on to their own children in the future. In this case,

only the second generation speakers were targeted because in the case of first generation speakers, their sense of identity as Colombian and as Hispanic is strong as they spent the majority of their lives in Colombia and still consider their status as immigrants rather than Canadian. The sense of identity as Colombian or Canadian, and the question regarding their ties to their cultural background, seems most relevant in the case of second-generation speakers. Further, motivations for teaching the language to future generations, where parents' attitudes and opinions are most important are already assessed in this investigation in the section regarding the attitudes towards inter-generational transmission of the language (see Section 2.6.1.2 in this chapter). In order to target speakers' attitudes and thoughts on this topic, the following question was asked: "Do you consider yourself more Colombian or Canadian? Why?". From the responses provided by speakers to these questions, the data was coded using the following codes: both/depends on who asks; divided (neither fully); connection to Colombia - culture and family; connection to Canada - culture and relationships; language considerations; labeling as outsider/discrimination.

2.7 Data Analysis

The data collected in this study was analyzed using quantitative and qualitative analysis methods according to the type of research questions that were being considered. Recall that in this work, I have divided the analyses into two sections. The first and primary part of this thesis work focuses on an investigation of the variable expression of Spanish subject personal pronouns as used by first and second generation immigrant Spanish speakers in London, Ontario. This data was coded as discussed in section 2.4 in order to account for the frequency of use of overt vs. null SPPs, and to determine how this variable is conditioned by different linguistic and social factors. For this analysis I used quantitative approaches which are described below. In contrast, the second part of this work focuses on an investigation of the factors that may contribute to the inter-generational maintenance of Spanish in London, Ontario. For this

investigation, determining the attitudes and language habits of speakers is essential, and therefore a qualitative thematic analysis was used. In the following sections, the quantitative and qualitative analysis methodologies are described in further detail.

2.7.1 Quantitative data analysis

In order to determine the social and linguistic factors that condition the variable expression of Spanish SPPs in the two groups of Colombian speakers (first and second generation speakers), the analysis of the data was carried out using Rbrul (Johnson 2009) through R Studio (RStudio Team 2020). Specifically, I conducted a multivariate logistic regression analysis (step-wise up and down) in order to determine which social and linguistic factors were most important in conditioning the variable expression of Spanish SPPs. That is, which factors, of those considered in the investigation, contribute to the construction of the best model which most successfully explains the observed variation in the data. Excel (2018) was used in order to organize the data, to carry out preliminary analyses including summary statistics, and to create the various charts needed to visually represent the data. In construction of the graphs used to represent the data, I used the proportion of use of overt SPPs or the proportion of pre-verbal SPPs, in percentage format, plus or minus the standard error of the mean (SEM) value which was used to create the error bars in all graphs.

Following upload of the data to Rbrul and completion of the multivariate logistic regression analysis, a X^2 test was conducted with all factors that Rbrul identified as contributing to the best logistic regression model. The significance of differences between groups was in all cases initially tested using a 0.01 level of significance, and in cases where the differences were not significant at this level, the differences were tested using a 0.05 level of significance. The p-values and level at which factors were found to be significant are indicated where appropriate. In the case of the analysis to determine differences between the two generations' use of pre-verbal and post-verbal subjects the two population proportions were compared using a X^2 test.

2.7.2 Qualitative data analysis

In applied sociolinguistics, interviews have been gaining in popularity in recent years as a means of investigating the beliefs and lived experiences of participants in response to certain phenomena (Talmy 2010). In this study, I conduct a qualitative analysis of the interview data collected in order to gain a better understanding of the context in which the language contact situation between English and Spanish occurs. This qualifies this study, according to Manzoor (2016), as an embedded mixed-methods analysis, where qualitative data is used as a secondary source of data to better inform the primary quantitative analysis. The thematic analysis conducted in this investigation is theoretical, or deductive, in nature, which indicates that decisions surrounding the identification of themes coded in the data resided on specific questions which were based on previous theoretical knowledge⁵. That is, I conducted the thematic analysis by focusing on specific questions which actively sought to address topics related to theory on the language maintenance factors outlined above by examining the interview data for explicit content which related to the language maintenance factors considered (Table 2.6). Specifically, I hypothesize that due to the fact that the presence of Spanish in Canada is relatively recent (in fact, much more recent than in the U.S.) and that there is little influence from Hispanic culture, among other factors, the use of Spanish in Canada is likely to result in language shift that could occur in two or three generations.

This hypothesis is addressed deductively by assessing the attitudes, experiences and overall comments regarding Spanish in Canada for all 20 speakers, and by determining whether these align with factors known to be associated with maintenance of a minority language in a situation of language contact.

In Table 2.6, I highlight the seven factors that I consider in this analysis and the specific questions that were asked to target these factors in the sociolinguistic interviews.

All factors considered as outlined in Table 2.6 which are included in the current analy-

⁵This is contrasted against an inductive approach which seeks to identify the research question after examining the data for themes (Braun and Clarke 2006)

Factor	Interview Question(s)
Cultural association with Hispanicity or Colombia (G2 only)	Do you consider yourself more Colombian or Canadian? What do you like or dislike about being Colombian/Canadian? Do you maintain a close relationship with Hispanic family abroad?
Perceived importance of Spanish for speakers.	Has Spanish helped you in your work? Do you use Spanish at home? with other family?
Perceived importance of the use of Spanish for future generations.	Do you think it is important that the children of Hispanic families in Canada learn to speak/read/write Spanish? Why or why not?
Attitudes toward endogamous /exogamous marriages.	Do you think it is important to marry someone who is Latin@/Hispanic? Why or why not?
Access to, or knowledge of community-led resources for the maintenance of Spanish.	Are there any resources/groups available to help you/your children learn and practice Spanish?
Use of Spanish in social and professional contexts.	Do you speak in Spanish with your Hispanic friends? Do you use Spanish at work/school? Has Spanish helped you with your work?
Perception of attitudes towards Spanish language use by majority community	Do you use Spanish at work/school? Do you use Spanish in public? Why or why not?

Table 2.6: Questions used to address factors contributing to language maintenance for qualitative thematic analysis.

sis, have been identified previously in the literature as important factors contributing to the maintenance of a minority language and are described above (see section 2.6.1).

Note as well that within Braun and Clark's (2006) description, the methods I used to determine the themes within the data were at the "semantic level", which in qualitative research indicates that I focused on the interview content as explicitly expressed by speakers without attempting to draw latent meaning from the statements in the data.

Further, during the analysis, some of the above factors were collapsed into a single group as these larger factor groups were better suited for the data collected. Notably, the factors relating to the perceived importance of the language to speakers were grouped together with the factor relating to the speakers' use of Spanish in social and professional contexts, and that of their perceptions of the larger community's attitudes towards their language. All of these factors seemed to relate to the speakers' use of Spanish in their daily life in Canada, and many similar themes were salient in the different contexts. Therefore, I determined that these could

be grouped together, and the five factor groups discussed in the following sections are:

1. Spanish language use in various contexts in Canada
2. Attitudes towards inter-generational language transmission
3. Community and Spanish-learning opportunities.
4. Attitudes towards linguistic exogamous/endogamous relationships
5. Cultural ties, background, and identity.

Following transcription of the interviews, data regarding each of the questions addressing the different language maintenance topics being addressed were extracted.

Given the semi-structured nature of the sociolinguistic interviews, some additional transcription was needed to gather the data since questions were sometimes asked in different order and therefore some answers to the key questions being considered were not initially transcribed in some of the interviews (i.e., in the case of long interviews where the questions were asked closer to the mid-point of the interview instead of at the end). Following the extraction of all the necessary data, I carefully organized the data for each topic using Microsoft Word and an Excel workbook composed of multiple spreadsheets (one spreadsheet for each question), and then proceeded to examine the data and highlight the most relevant or recurrent aspects for each speaker. From these selections, codes were developed for each topic according to the patterns found in the data. I then repeated the process, this time looking at all speaker data for each topic and identifying the most relevant and recurrent codes across speakers for each topic. Finally, after looking at final codes, codes that were similar were grouped together, some that were less prevalent in the data were excluded (i.e, codes on speakers' comments which were only present in one or two interviews), and remaining codes were organized into theme and sub-theme categories. It is important to note that in developing codes, themes, and subthemes in the data, theoretical considerations on each topic were always considered, which further classifies our analysis as deductive, and our theme selection as theory-grounded (Braun and Clarke 2006).

1 Qualitative study reliability measures

It is important at this point to address the measures taken to ensure the reliability of the current qualitative analysis. As qualitative approaches become more commonplace in fields of research outside of sociology, questions regarding the reliability and objectivity of qualitative analyses have been raised. Notably, it is argued that some level of subjectivity is present when creating codes and analysing data qualitatively, which is not present in quantitative research. According to Barusch et al. (2011), given the acknowledged subjectivity of qualitative research methods, it is important in qualitative studies that authors present a description of the researcher within the methodological approach of a given study in order to provide readers with a better understanding of the “researcher’s lens”. With this goal in mind, I present in what follows a description of myself as a researcher at the time of this investigation in order to provide the reader with a better understanding of the researcher’s lens used in this study.

At the time of this investigation, I am a Ph.D. candidate in linguistics at the University of Western Ontario. My research focuses primarily on topics of Spanish language variation, but I am also active in projects addressing the intersection of language and education, especially as these concern the education of students who are speakers of non-standard language varieties such as Creoles. Although my research interests nowadays focus in linguistics, my academic background also includes an Honours Bachelor of Science where I specialized in Biology. I also am a multilingual speaker who is fluent in Spanish, English and French. Perhaps the most relevant factor of myself in regards to this investigation is that I am a native speaker of Spanish, with native-like proficiency in English, and I am a Canadian citizen who immigrated from Colombia and arrived in London, Ontario, at the age of twelve. As an immigrant Spanish speaker, I consider myself to be part of the 1.5 generation as described by Guardado (2008), with different aspects of myself identifying with both first and second generation speakers. At this point, I would also like to address that although I do have pre-existing attitudes and opinions on the questions and topics addressed in this investigation, I did my best to ensure that none of these influenced the study in a way that biased the investigation.

In qualitative research, authors' biases and backgrounds can influence the patterns and/or interpretations given to themes that become most salient, as authors tend to write themselves into their work (Barusch et al. 2011). This can be circumvented by collaborating with a second author in the investigation to verify that multiple independent researchers arrive at similar conclusion given the same data set. However, in the case of the current investigation, since it is an independent project, this was not possible. Similarly as when considering the observer's paradox in sociolinguistics, there is no one specific fool-proof way to mitigate all possible effects from the researcher's influence in the analysis, but there are strategies that can be used to mitigate its effects by remaining aware of the influence of the situation at play. That is, just as in sociolinguistics we try to minimize observer's paradox effects by employing strategies that make speakers feel comfortable, or addressing topics which awaken more emotional reactions which helps speakers forget they are being observed (as much as possible), there are strategies that I used to mitigate and minimize the effects of my own biases during the qualitative analysis. As an investigator, I am aware of my own biases as a Colombian immigrant living in Canada, and I made a strong effort to ensure that in my analysis of the data I remained as unbiased and impartial as possible considering all possible patterns in the data including those that disagree or conflict with my own personal opinions or points of view. To further ensure that no personal biases influenced the analysis, I established and followed specific protocols in the coding of the data analysed, especially when dealing with topics where I may have had strong personal opinions such as my pre-existing notions regarding unavailability of Spanish language resources in London, Ontario, and hesitance of first generation speakers to use Spanish in public (from previous experiences and comments from personal connections). Specifically, the questions designed to address each of the specific topics were read exactly as written during the interview and answers by speakers to these pre-determined questions were fully transcribed and coded in totality. Further, by using a semantic rather than latent analysis of the data, patterns and codes that I identified in the data reflected only that which was specifically expressed by speakers and I used no latent meaning behind speakers' statements to

derive codes, themes or subthemes. Also, although traditional qualitative analysis methodologies tend to discourage the quantification of data (Braun and Clarke 2006), I relied on counts of codes, (within categories organized in Excel spreadsheets), in order to determine the most salient themes as gathered explicitly from the data. I also provided some quantifiable evidence within the qualitative analysis to ensure the reader can clearly see the extent to which a given theme was present among the speakers in the CoSLO corpus.

One further approach that was taken to ensure the reliability of the current qualitative analysis was to ensure that a theoretical basis for the development of topics addressed, codes, and themes was used from the beginning. This further follows the suggestions proposed by Barusch et al. (2011), minimizing the likelihood that any author-specific influences impacted the analysis, and ensuring a high standard of reliability in the qualitative analysis.

Chapter 3

Results

3.1 Quantitative analysis: Variable subject pronoun expression in Londombia

In this section, I present the results of the quantitative analysis of the current investigation as it pertains to the analysis of the linguistic, social and extra-linguistic factors that influence the use of overt as opposed to null subject pronouns in variable contexts in the Spanish of Colombian speakers living in London, Ontario.

I will begin by presenting an overview of the results while considering both generations of speakers consulted in this study. This will then be followed by independent analyses of each generation of speakers to determine whether there are differences in regards to the conditioning factors of the variable expression of SPPs between the two generations.

3.1.1 Considering both generations

A total of 2366 tokens were extracted from the data including both generations. These tokens were coded according to the different linguistic, social, and extra-linguistic factors considered. A multivariate logistic regression of the coded data was carried out using Rbrul (Johnson 2009), and the significant factors were then further analysed using a X^2 test for categorical

data.

The linguistic factors considered are switch reference (utterance initial, same referent, and switch of referent), grammatical person (1SG, 2SG, 3SG, 1PL, and 3PL), pronoun number (SG vs. PL)¹, clause type (main, coordinated, subordinate), verb mood, verb tense, semantic content of the verb (active, mental, and stative), and polarity of the clause ([+ negative], or [- negative]). The social factors considered are gender (male vs. female), age, generation of immigration (first vs. second), length of residence in Canada, and age of arrival in Canada. Lastly, one extra-linguistic factor is also considered which is the modality through which the interview was conducted (in person, telephone, or video call). For more details about these factors please refer to sections 2.5.1 and 2.5.2 in chapter 2 of this work. It is also important to note that in the analysis in Rbrul, speaker was included as a random intercept factor in order to increase reliability of the data and to minimize the likelihood of Type 1 errors due to over-representation of tokens by specific speakers (Johnson 2009; Brezina 2018). Speaker itself was not included in initial analyses as a fixed factor. However, at the end of the current section, the by-speaker data is presented and a short analysis of by-speaker trends is presented.

Following the Rbrul analysis, when considering both generations of Colombian Spanish speakers in London, Ontario grouped, four linguistic factors are found to be significant according to the best logistic regression model: Grammatical person, verb mood, switch reference, and polarity of the clause. It is important to note that pronoun number, although not selected as part of the best model, was significant when grammatical person was not considered. However, the model including grammatical person instead of pronoun number showed a lower AIC value and was therefore selected as the preferred model. Nevertheless, I consider both factors to be relevant and in this section I will comment on both. Further, no social or extra-linguistic factors were identified as significant following the multivariate logistic regression when considering both generations together. A table displaying the results from the multivariate logistic

¹Pronoun number and grammatical person were tested in separate multivariate regression models. This was done since including them together into a single test introduces a co-linearity that cannot be accounted for mathematically. The model including grammatical person showed the lowest AIC (Akaike Information Criterion) value and was therefore selected as the best fit model.

regression analysis ($R^2=0.15$) is included in Table (3.1) below.

	Overt SPP	N	Proportion	Factor Weight
Grammatical Person				
3SG	178	564	0.316	0.67
1SG	353	1121	0.315	0.62
2SG	20	79	0.253	0.53
3PL	54	287	0.188	0.44
1PL	32	315	0.102	0.26
Range				41
Verb Mood				
Conditional	23	41	0.561	0.65
Subjunctive	26	77	0.338	0.48
Indicative	589	2248	0.262	0.37
Range				28
Switch Reference				
- Continuity	368	1099	0.335	0.60
+ Continuity	269	1267	0.212	0.40
Range				20
Polarity				
- Negation	568	2074	0.274	0.54
+ Negation	68	292	0.233	0.46
Range				8

Table 3.1: Linguistic, social, and extra-linguistic factors conditioning the overt expression of Spanish SPPs among two generations of Colombian Spanish speakers living in London, Ontario ($N=2366$).

In Table (3.1), the results from the Rbrul analysis for both generations of speakers are displayed. In the sections that follow, I discuss the effects of these factors on the variable expression of Spanish SPPs following the order of factor effect size according to the Rbrul results.

1 Grammatical person and number

Results from the analysis suggest that the grammatical person of the pronouns in question is a significant factor in conditioning the overt vs. null expression of SPPs ($X^2(4, N = 2366) = 72.73, p < 0.001$) in variable contexts (Figure 3.1).

Notably, third person singular ($FW = 0.67, N = 564$), first person singular ($FW = 0.62, N$

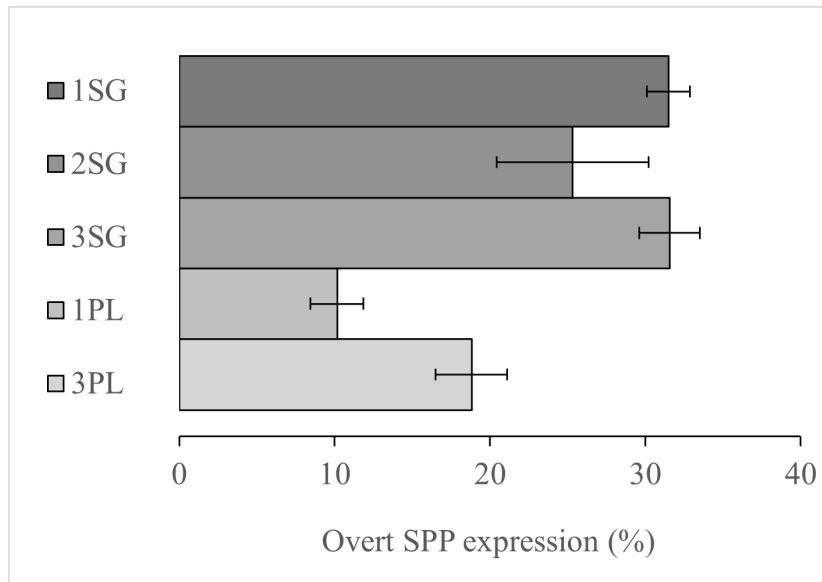


Figure 3.1: Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to 3SG ($N=564$), 1SG ($N=1121$), 2SG ($N=79$), 3PL ($N=287$), and 1PL ($N=315$) grammatical person.

= 1121), and second person singular ($FW = 0.53$, $N = 79$) constructions favour the use of overt SPPs while third person plural ($FW = 0.44$, $N = 287$) and first person plural ($FW = 0.26$, $N = 315$) constructions disfavour overt SPP expression.

From these results, we can see a clear division between singular and plural pronouns which points to the fact that pronoun number is an important factor in conditioning the overt expression of SPPs in this population. Therefore, although the pronoun number was excluded from the best-fit logistic regression model ², pronoun number remains an important conditioning factor for the variable in question (figure 3.2).

As shown in the figure above (Figure 3.2), a significant higher use of overt SPPs is associated with singular constructions (31.23%, $N = 1764$), than with plural constructions (14.28%, $N = 602$) ($X^2(1, N = 2366) = 65.54$, $p < 0.001$). Note that the majority of tokens extracted are in the first person singular and third person singular grammatical person. This isn't surprising given the nature of the interview style where speakers are more likely to be referring to themselves or other third persons in their speech.

²This was determined by selecting the model with the lowest AIC (Akaike Information Criterion) value.

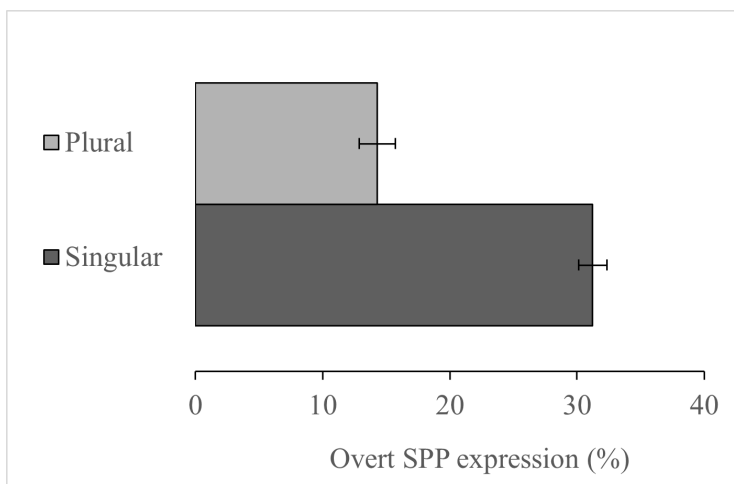


Figure 3.2: Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to singular ($N= 1764$) and plural ($N= 602$) pronoun forms.

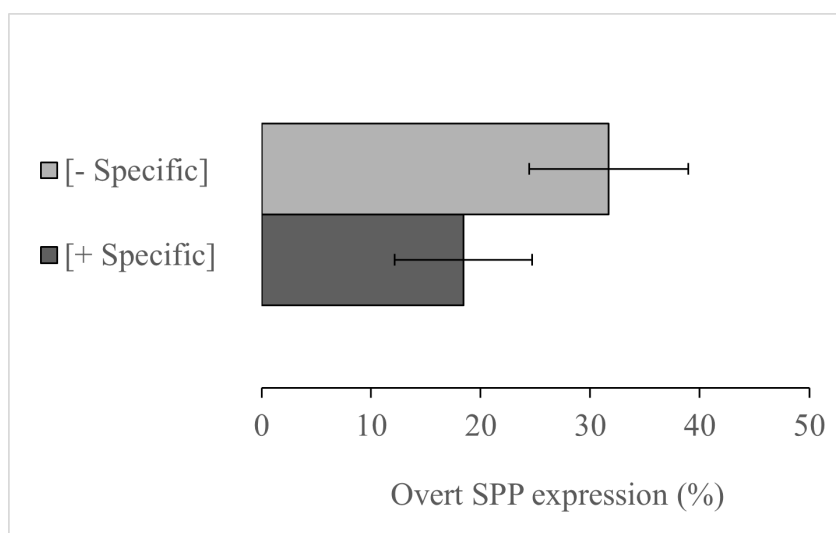


Figure 3.3: Frequency of use (+/- SEM) of 2SG overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to [+ specific] ($N=38$) or [- specific] ($N=41$) pronouns.

It is also important to highlight that within this analysis, second person tokens ($N = 79$) include both instances where the pronoun was used with a specific reading (in this case referring to the interviewer), and a general reading (i.e., using *tú* ‘you’ to refer to a non-specific person).

A X^2 analysis was carried out using these second person singular constructions to determine if there were significant differences in regards to how [+ specific] and [- specific] 2SG pronouns

condition the use of overt SPP expression to determine if there were any differences between the two in regards to variable SPP expression (Figure 3.3).

Although the difference is not significant ($X^2(1, N=79) = 1.84, p = 0.175$), there is an observable trend where [- specific] 2SG pronouns tend to correlate with higher instances of overt SPPs (31.7%, $N=41$), than [+ specific] 2SG pronouns (18.4%, $N=38$). It is likely that the difference between these two conditions did not reach significance due to the small sample size ($N=79$) which creates a large margin of error as seen in Figure (3.3) above.

2 Verb mood

Another linguistic factor that is significant in conditioning the overt expression of Spanish SPPs in the population of Colombian Spanish speakers in London, Ontario, is verbal mood.

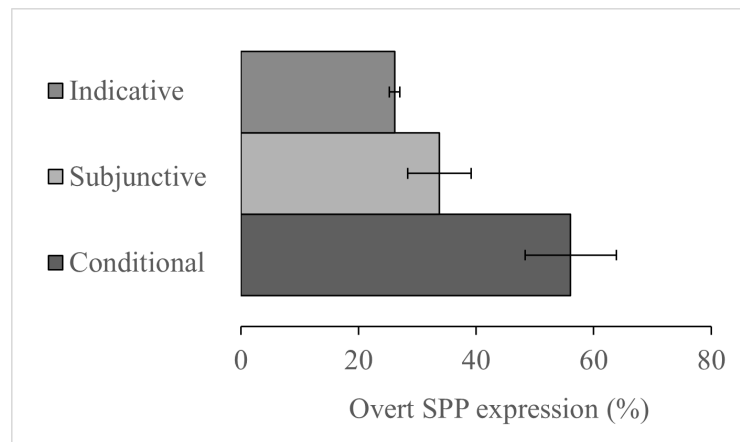


Figure 3.4: Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to indicative ($N=2258$), subjunctive ($N=77$), or conditional ($N=41$) verbal mood.

According to the Rbrul analysis conducted, constructions using the conditional mood ($FW=0.65, N=41$) favour the overt expression of SPPs while subjunctive ($FW=0.48, N=77$) and indicative ($FW=0.37, N=2258$) constructions disfavour it (Figure 3.4).

Note that conditional constructions use the highest proportion of overt vs. null SPPs (56.09%, $N=41$), followed by subjunctive constructions (33.76%, $N=77$), and indicative constructions (26.16%, $N=2258$). The difference between these proportions was found to be sig-

nificant according to the X^2 test conducted ($X^2(2, N = 2366) = 20.24, p < 0.001$).

3 Switch reference

Results of the multivariate logistic regression also suggest that the overt expression of subject personal pronouns (SPPs) in variable contexts among the population considered is conditioned significantly by the switch reference factor.

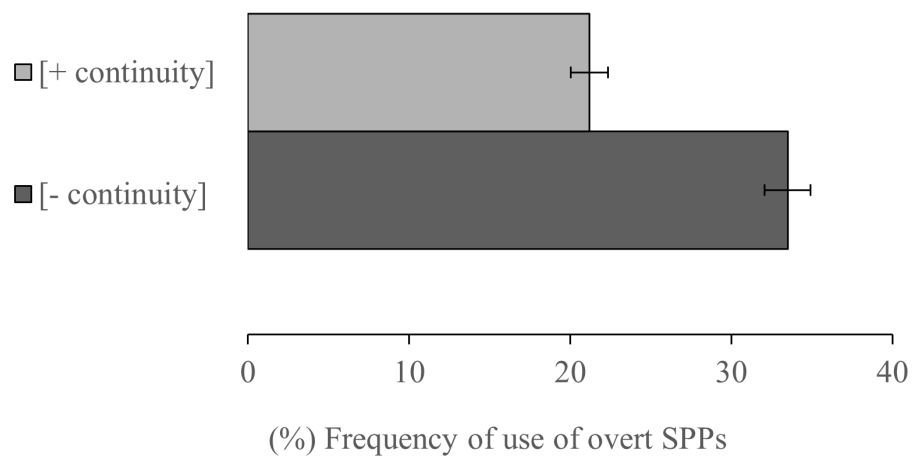


Figure 3.5: Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to the switch reference factor ([+ continuity], $N=1267$, [- continuity], $N=1099$)

Results show that overt SPP expression is favoured when there is no continuity of the referent in the discourse ($FW = 0.60, N = 1099$), while the overt expression of SPPs is disfavoured when there is continuity of the referent ($FW = 0.40, N = 1267$). The differences between the favouring and disfavouring conditions are found to be significant following a X^2 analysis ($X^2(2, N = 2366) = 42.24, p < 0.001$).

4 Clause polarity

Results of our analyses also showed that the polarity of the clause is a significant factor in determining the use of overt vs. null SPPs in variable contexts within our population. This

factor was coded according to whether the clause was negated [+ Negation] or not [- Negation]. Results show that the presence of negation disfavors the use of overt SPPs ($FW = 0.46$, $N = 292$) in variable contexts, while the absence of negation favours it ($FW = 0.54$, $N = 2074$) (Figure 3.6).

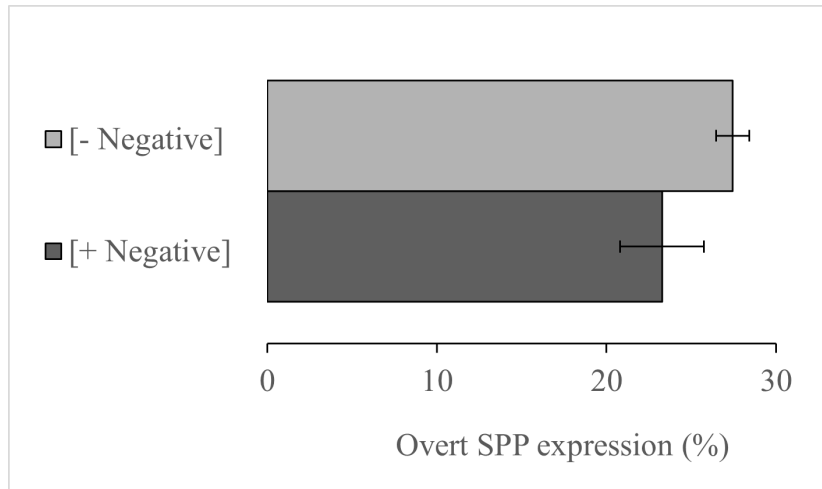


Figure 3.6: Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to whether the clause was [+ Negative] ($N = 292$) or [- Negative] ($N = 2074$).

Although a significant factor within the logistic regression model in Rbrul, note that the factor weights for the two conditions are almost neutral and centered around the 0.5 value. This suggests a very small association. Further, a subsequent X^2 test conducted determined that the difference between [- Negative] (27.4%) and [+ Negative] (23.29%) constructions as it relates to the expression of SPPs in variable contexts was non-significant³ ($X^2(1, N = 2366) = 2.2376$, $p = 1.347$).

5 Other factors - Non-significant or not selected as significant conditioning factors

In this section, I briefly discuss the remaining factors which were not selected as part of the best logistic regression model following a step-wise, up-and-down variable regression analysis

³Note that the absence of significance within a X^2 test does not mean that the factor is not important in conditioning the variable in question. Rather, that according to the specific test, the difference was not detected as deviating sufficiently from the mean. Therefore, the absence of significance in one test does not imply there is no effect, just that an effect was not detected by the specific statistical test used.

conducted in Rbrul.

The factors that are described here are: clause type, verb tense, position of the pronoun in relation to the verb, semantic content of the verb, speaker gender, and bilingual language proficiency.

1. Clause type

In the current investigation, clause type was considered as a possible factor contributing to the variable expression of Spanish SPPs. The factor was coded according to whether the tokens for overt and null SPPs were found in main, subordinate, or coordinate clauses. When considering data for both generations, it was determined that in variable contexts, overt SPPs were used most frequently in main clauses (29.8%, $N = 1081$), followed by subordinate clauses (26.5%, $N = 599$), and were least frequent in coordinate clauses (22.7%, $N = 686$). However, although there is a trend with main clauses being associated with higher frequency of overt SPPs, this seems to be a weak association.

2. Verb tense

The tense of verbs accompanying SPPs was a factor also considered in the current analysis. The broad categories of tenses considered in the analysis were present, past, imperfect, and future. When the data for both generations is considered, overt SPPs are found to be more frequently expressed with past (27.6%, $N = 550$), imperfect (27.8%, $N = 263$), and present (26.7%, $N = 1504$), while they are less frequent with verbs in the future tense (17.9%, $N = 39$).

3. Semantic verb type

Another factor considered in the analysis was the semantic content of the verbs. Verbs were coded according to three semantic categories in order to determine effects of the semantic content of the verb on variable expression of Spanish SPPs: Mental, Stative, and Action. Results showed that overt SPP production in variable contexts was most frequent with mental verbs when both generations of Colombian Spanish speakers were considered (29.0%, $N = 558$), followed by stative verbs (27.1%, $N = 711$), and finally action verbs (25.7%, $N = 1097$).

4. Gender

Speaker gender was not included within the best logistic regression model when considering both generations. In this case, this was due to the two genders considered, men and women, being similar in their use of overt vs. null SPPs. Nevertheless, in the sample collected, men used overt SPPs slightly more frequently (28.1%, $N = 1159$) than women (25.8%, $N = 1207$).

In the following section, I consider the results from the quantitative analysis considering first generation speakers, followed by the quantitative analysis of second generation speakers. An analysis of the two generation groups separately will allow us to better determine the way that SPPs are conditioned within each group, which could help in identifying differences between the two groups. This analysis of intra- and intergenerational differences can be instrumental in identifying any possible processes of change across generations.

3.1.2 First Generation Speakers

When considering the variable expression of SPPs in the group of first-generation speakers in our sample, a total of 1239 tokens were extracted from 10 interviews. In this section, I present the results of the statistical analyses conducted using this subset of tokens.

	Overt SPP	N	Proportion	Factor Weight
Grammatical Person				
2SG	17	50	0.340	0.64
1SG	136	474	0.287	0.61
3SG	96	379	0.253	0.60
3PL	26	150	0.173	0.43
1PL	16	186	0.086	0.24
Range				40
Switch Reference				
- Continuity	167	527	0.317	0.61
+ Continuity	124	712	0.174	0.39
Range				22

Table 3.2: Linguistic and social factors conditioning the overt expression of Spanish SPPs among first-generation Colombian Spanish speakers living in London, Ontario.

Following a multivariate logistic regression analysis conducted in Rbrul, it was determined that the variable expression of SPPs in first-generation speakers was significantly conditioned

by two linguistic factors: grammatical person of pronoun, and switch reference. The Rbrul results from this multiple logistic regression analysis ($R^2=0.149$) are shown in Table (3.2).

Notice that when compared to the results obtained when considering both generations data together, fewer factors seem to condition the variable expression of Spanish SPPs in the first generation group, with clause polarity and verbal mood not being considered significantly contributing factors. This could point to an effect primarily driven by the second generation group.

1 Grammatical person and number

The second factor that significantly conditions the variable expression of SPPs in the case of first-generation speakers was the grammatical person of the pronoun. According to the Rbrul results, overt SPPs are favoured, in order, by 2SG ($N= 50$, $FW=0.64$), 1SG ($N= 474$, $FW= 0.61$), and 3SG ($N= 379$, $FW= 0.60$) constructions. In contrast, the use of overt SPPs in variable contexts is disfavoured by 3PL ($N= 150$, $FW= 0.43$) and 1PL ($N= 186$, $FW= 0.24$) constructions. Note that there are insufficient tokens representing 2PL constructions and therefore these were excluded from the analysis, as was done with the analysis of both generations.

The use of overt SPPs in variable contexts was significantly more frequent in 2SG (34%), 1SG (28.7%), and 3SG (25.3%) contexts, while it was used less frequently in 3PL (17.3%) and 1PL (8.6%) contexts ($X^2(4, N=1239)=37.032$, $p < 0.001$) (Figure. 3.7).

Note as well, that there is a clear divide between singular and plural contexts. Although pronoun number was not selected as part of the best fit logistic regression model, this is an important factor and it is considered a significant factor in Rbrul results when grammatical person is excluded. As can be seen in Figure (3.8), in the case of variable contexts, the data from our sample shows that overt SPPs are used significantly more frequently in singular (27.8% $N= 897$), than in plural (12.3%, $N= 42$) constructions ($X^2(1, N= 1239)= 33.01$, $p < 0.001$).

Similarly to the situation observed when considering both generations of speakers together, grammatical person is a better predictor of the use of overt SPPs in variable contexts than

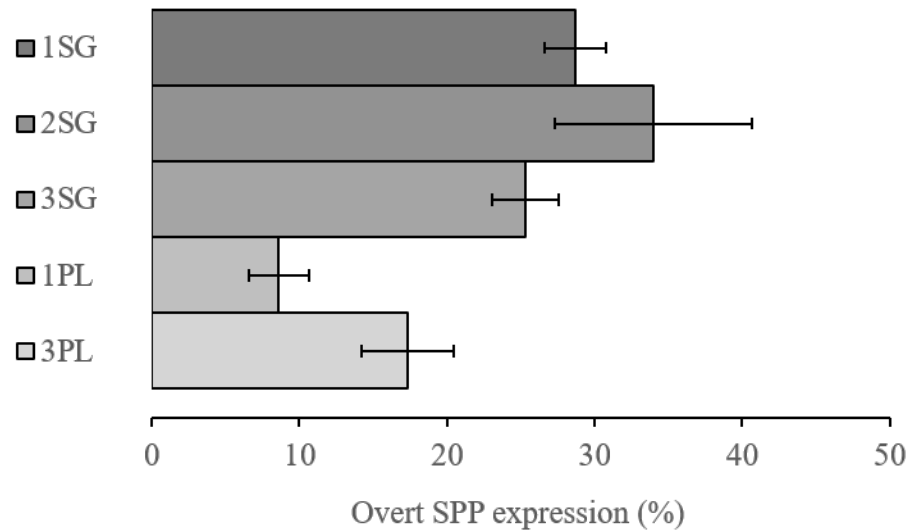


Figure 3.7: Frequency of use (+/- SEM) of overt SPPs by first-generation speakers of Colombian Spanish in London, Ontario according to 1SG ($N=474$), 2SG ($N=50$), 3SG ($N=379$), 1PL ($N=186$), and 3PL ($N=150$) grammatical person.

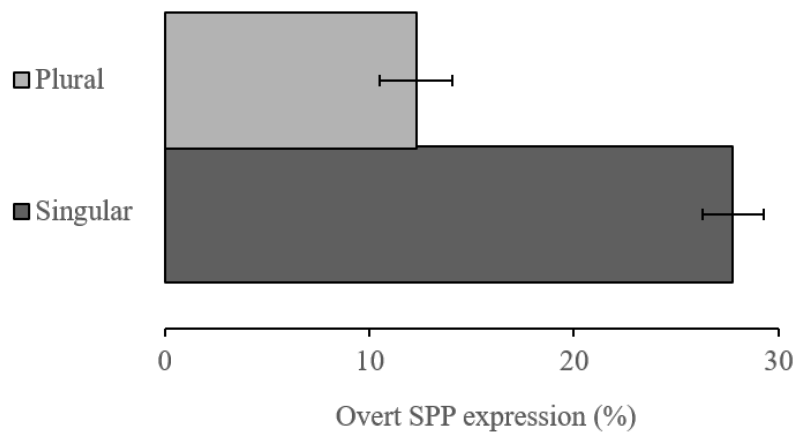


Figure 3.8: Frequency of use (+/- SEM) of overt SPPs by first-generation speakers of Colombian Spanish in London, Ontario according to singular ($N= 897$) and plural ($N= 42$) SPP constructions.

grammatical number (Section 3.1.1) according to the AIC value of the models. However, pronoun number remains an important conditioning factor for first generation Colombian Spanish speakers.

2 Switch reference

The tokens extracted from the sample were coded according to whether the pronoun in question represented a new referent or a switch in referent or whether the pronoun represented the same referent.

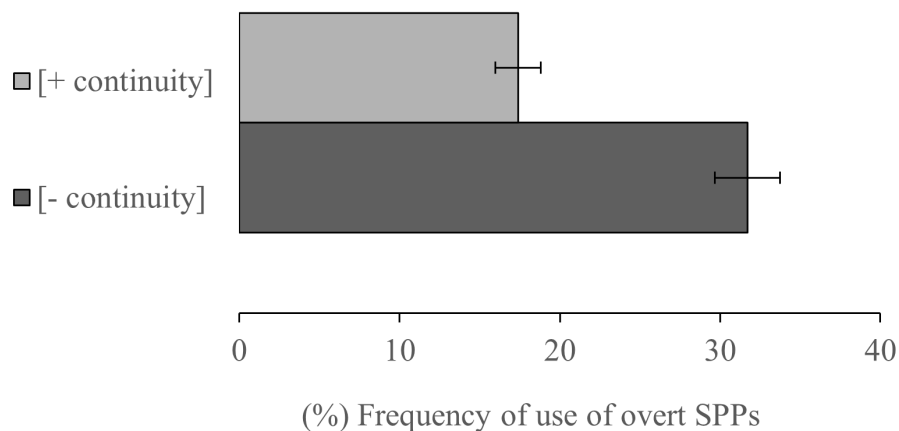


Figure 3.9: Frequency of use (+/- SEM) of overt SPPs by first-generation speakers of Colombian Spanish in London, Ontario according to whether there was a continuous referent ([+ continuity], $N=712$), or a non-continuous referent ([- continuity], $N=527$).

If the pronoun represented a new referent or a switch in referent, it was considered to denote a lack of referent continuity, ([- continuity]), and if the pronoun represented the same referent, it was considered to denote continuity ([+ continuity]) of the referent.

According to the results, the use of overt SPPs in variable contexts is favoured when there is no referent continuity in the discourse ($N = 527$, $FW = 0.61$). In contrast, SPPs which represented the same referent as in the previous clause, disfavoured the use of overt SPPs ($N=712$, $FW= 0.39$).

As can be seen in Figure 3.9, overt SPPs were used most frequently in variable contexts when the SPP was in utterance initial condition or when there was a change in referent, that is when there was no continuity of referent (31.7%). In contrast, overt SPP expression was used least frequently (17.4%) in cases where the pronoun represented a continuous referent. The

difference between the use of overt SPPs when there is no referent continuity, which favour the use of overt SPPs, and its use when there is a continuing referent, which disfavors them, is significant ($X^2(1, N=1239)=34.332, p < 0.001$).

3 Other factors - Non-significant or not selected as significant conditioning factors

In this section, as I did when considering data for both generations, I briefly discuss the remaining factors which were not selected as part of the best logistic regression model for data considering only the first generation of speakers following a step-wise, up and down variable regression analysis. The factors that are described are: clause type, verb tense, semantic content of the verb, clause polarity, interview modality, gender, length of residence, and bilingual language proficiency. I also include results for effects of proficiency in English, a factor only considered for first generation speakers since there was no variation for this factor with second generation speakers who were all highly proficient in English.

1. Clause type

In the current investigation, clause type was considered as a possible factor contributing to the variable expression of Spanish SPPs. The factor was coded according to whether the tokens for overt and null SPPs were found in main, subordinate, or coordinate clauses. When considering data for first generation speakers, it was determined that in variable contexts, overt SPPs were used most frequently in main clauses (25.7%, $N = 548$), followed by subordinate clauses (23.9%, $N = 310$), and were least frequently in coordinate clauses (19.9%, $N = 381$). Note that although there is a trend with main clauses being associated with higher frequency of overt SPPs, this seems to be a weak association.

2. Verb tense

The tense of verbs accompanying SPPs was a factor also considered in the current analysis. The broad categories of tenses considered in the analysis were present, past, imperfect, and future. When the data for only the first generation is considered, overt SPPs are found to be more frequently expressed with imperfect (25.4%, $N = 142$), and past (26%, $N = 331$), while

they are less frequent with present (22.1%, $N = 745$), and future tense verbs (19%, $N = 21$).

3. Verb mood

Verb mood is another factor that was excluded from the best logistic regression model selected through Rbrul's step-wise variable selection when considering first generation speakers. When this factor is considered for first generation speakers alone, overt SPPs are used most frequently with conditional (40%, $N = 15$) and subjunctive constructions (27.3%, $N = 55$), while they seem to be least frequent with indicative construction (23.1%, $N = 1169$).

4. Semantic verb type

Another factor considered in the analysis was the semantic content of the verbs. Verbs were coded according to three semantic categories in order to determine effects of the semantic content of the verb on variable expression of Spanish SPPs: mental, stative, and action. Results showed that overt SPP production in variable contexts was most frequent with mental verbs when first generation Colombian Spanish speakers were considered (28.0%, $N = 239$), followed by stative verbs (23.3%, $N = 420$), and finally action verbs (21.7%, $N = 580$).

5. Clause polarity

The polarity of the clause (presence or absence of negation) was also considered in the analysis conducted. Results for the first generation speakers' data show that Spanish SPPs were expressed overtly with higher frequency with [- negative] clauses (23.8%, $N = 1117$) than with [+ negative] clauses (20.5%, $N = 122$). This differs from results which considered data from both generations and data from second generation speakers. Nevertheless, when we consider the first generation speaker data only, we see a trend which favours overt SPPs in affirmative phrases which follows the pattern observed when both generations are considered together.

6. Gender

Speaker gender was not included within the best logistic regression model when we considered the first generation immigrant speakers. Similarly to when we considered both generations, in the case of first generation speakers, this was due to speakers from the two genders

considered, men and women, being similar in their variable expression of Spanish SPPs. In the sample collected, men used overt SPPs slightly more frequently (24.1%, $N = 607$) than women (22.9%, $N = 632$).

7. Intensity of language contact (Bilingual proficiency and English proficiency)

The proficiency of speakers in Spanish and English was assessed through the language questionnaire administered as part of the study procedures. Therefore, the information on language proficiency is based on speakers' self-reports of their proficiency in Spanish and English. Following data collection, I categorized speakers as having advanced proficiency in both languages, or having advanced proficiency in only one language. In the case of first generation speakers, all speakers were native speakers of Spanish, and all were advanced in their literacy skills in Spanish. However, first generation speakers varied in their level of bilingualism. In regards to their variable use of Spanish SPPs, only two speakers reported themselves as being highly proficient in both Spanish and English. These speakers seemed to use of overt SPPs in variable contexts at a lower frequency (18.8%, $N = 267$), than the remaining eight speakers who reported having advanced proficiency only in Spanish (24.5%, $N = 972$) and who seemed to use overt SPPs more frequently.

Following this analysis, I also looked at how variation in level of English proficiency in these speakers could influence their production of overt SPPs in variable contexts. For this, I used speakers' self-reports in regards to their English proficiency and categorized them as having low ($N = 3$), intermediate ($N = 5$), or advanced ($N = 2$) proficiency in English. When I considered this factor, results showed that overt SPPs in variable contexts are used with highest frequency by speakers whose proficiency in English is at an intermediate level (28.1%, $N = 590$), followed by speakers with advanced proficiency (19.9%, $N = 267$), and speakers with low proficiency in English (18.8%, $N = 382$).

8. Intensity of language contact (Frequency of use of Spanish)

An analysis of frequency of use of Spanish as an additional measure of language contact intensity was included in the quantitative analysis for first generation speakers. The measure

used was based on a cumulative score calculated from speakers' answers to questions in the language use questionnaire which addressed their frequency of use of Spanish in four different contexts: home/family, social-local, social-international, and professional (school and/or work). The scores from each of these contexts were added and this raw score was divided by the number of possible points that each speaker could have had according to the number of questions they answered. The values calculated according this method ranged from 0.70 to 0.93 for the first generation group, and the median value was 0.85. Speakers' whose score was above this 0.85 value and therefore had a higher frequency of use of Spanish were considered to be in a situation of "Low" intensity of contact with English as compared to speakers in this generation group who had scores below at or below 0.84, which were categorized as speakers in a situation of "High" intensity contact with English. This factor was included in the multivariate regression analysis and it was not selected as a significant factor. That being said, speakers with lower self-reported frequency of use of Spanish, which were categorized as being in "High" intensity contact with English had a higher use of overt SPPs (25.9%, $N = 632$) than speakers which reported a more frequent use of Spanish and which were considered as being in "Low" intensity contact with English (20.9%, $N = 607$).

9. Intensity of language contact (Length of residence)

Length of residence varied for first generation participants from a length of 9 to 21 years.

In order to consider the correlation between the LOR factor and the use of overt SPPs, I calculated the Spearman's Rho correlation coefficient. The results of this test showed that there was no significant correlation ($r_s = -0.28311$, $p = 0.42799$, $N = 10$).

3.1.3 Second Generation Speakers

In our analysis of second generation speakers, a total of 1127 tokens were extracted from the 10 interviews that compose this group within the corpus. In this section, I present the results of the statistical analyses conducted using only this subset of tokens.

Following an Rbrul analysis, it was determined that the use of overt SPPs in variable con-

	Overt SPP	N	Proportion	Factor Weight
Grammatical Person				
3SG	82	185	0.443	0.78
1SG	217	647	0.335	0.66
3PL	28	137	0.204	0.47
2SG	3	29	0.103	0.29
1PL	16	129	0.124	0.29
Range				49
Verb Mood				
Conditional	17	26	0.654	0.67
Subjunctive	11	22	0.500	0.53
Indicative	318	1079	0.295	0.30
Range				37
Interview Mode				
In person	169	446	0.379	0.59
Video call	69	211	0.327	0.50
Telephone	108	470	0.230	0.41
Range				18
Switch Reference				
- Continuity	119	572	0.351	0.58
+ Continuity	145	555	0.261	0.39
Range				19
Clause Polarity				
- Negation	303	957	0.317	0.55
+ Negation	43	170	0.253	0.45
Range				10

Table 3.3: Linguistic and social factors conditioning the overt expression of Spanish SPPs among second-generation Colombian Spanish speakers living in London, Ontario.

texts by second-generation speakers is significantly conditioned by four linguistic factors and one extra-linguistic factor. The linguistic factors that are significant are switch reference, grammatical person, verb mood, and polarity of the clause. In addition, only one extra-linguistic factor, interview mode, seems to significantly influence the use of overt SPPs in variable contexts. The results of the Rbrul analysis are shown above (Table 3.3).

In this section I describe the effects of these factors as determined through the statistical analyses used. The factors are discussed in this section in descending order according to their effect size as determined in the Rbrul results above.

1 Grammatical person and number

The factor that seems to condition the variable expression of SPPs in the case of second-generation speakers most strongly is the grammatical person of the subject pronoun.

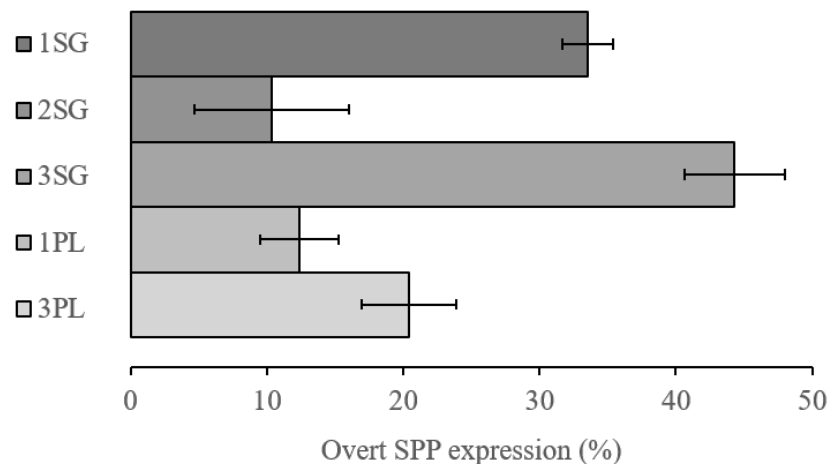


Figure 3.10: Frequency of use (+/- SEM) of overt SPPs by second-generation speakers of Colombian Spanish in London, Ontario according to 1SG ($N=647$), 2SG ($N=29$), 3SG ($N=185$), 1PL ($N=129$), and 3PL ($N=137$) grammatical person.

According to the Rbrul results, overt expression of SPPs is favoured, in order, by 3SG ($N=185$, $FW=0.78$), and 1SG ($N=647$, $FW=0.66$) constructions. In contrast, the use of overt SPPs in variable contexts is disfavoured by 3PL ($N=137$, $FW=0.47$), 2SG ($N=29$, $FW=0.29$) and 1PL ($N=129$, $FW=0.29$) constructions. Note that there was only one token representing 2PL constructions and therefore this grammatical person was excluded from the analysis.

The use of overt SPPs in variable contexts was significantly more frequent in 3SG (44.3%), 1SG (33.5%), and 3PL (20.4%) contexts, and it was used less frequently in 2SG (10.3%) and 1PL (12.4%) contexts ($X^2(4, N=1127)=51.319$, $p < 0.001$) (Figure 3.10).

Note that there seems to be a trend towards a divide between singular and plural constructions. The one exception seeming to be the apparent alignment of the 2SG forms with the frequencies of 1PL and 3PL forms. However, there were only a total of 29 tokens for 2SG

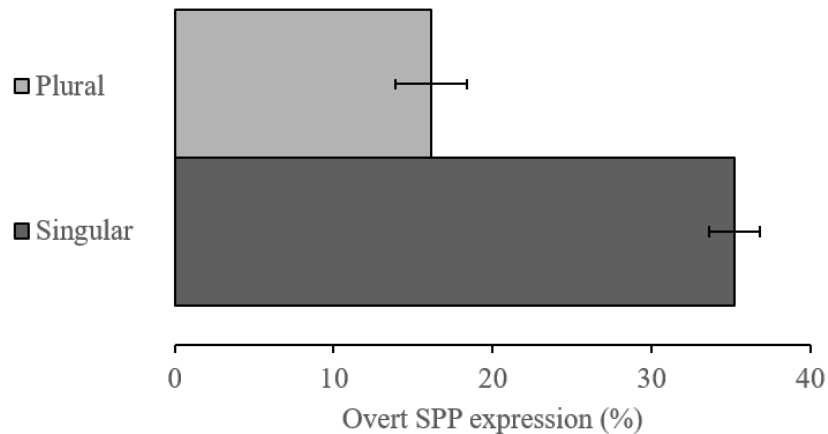


Figure 3.11: Frequency of use (+/- SEM) of overt SPPs by second-generation speakers of Colombian Spanish in London, Ontario according to singular ($N= 860$) and plural ($N= 267$) pronoun forms.

forms, and therefore there is very little that can be deduced from this with any certainty.

In addition, similarly to with the first-generation group, the 1SG and 3SG forms deviate significantly from the 1PL and 3PL forms (Figure 3.10). For this reason, although pronoun number was not selected as part of the best fit logistic regression model, this is an important factor to investigate.

The overt expression of SPPs is more frequent in singular constructions at 35% frequency, than in plural constructions where the frequency is only of 16.1% as seen in Figure (3.11). Despite the fact that this factor was not selected as part of the best-fit logistic regression model in Rbrul, this difference is significant according to a X^2 test ($X^2(1, N= 1127)= 35.037, p < 0.001$).

2 Verb Mood

Results from my analysis also revealed that when we consider second-generation speaker data, verb mood is a significant conditioning factor regarding the variable expression Spanish SPPs.

According to results from the multiple logistic regression analysis, the use of overt SPPs

is favoured with constructions which are in the conditional ($N=26$, $FW=0.67$) or subjunctive moods ($N=22$, $FW=0.53$), while it is disfavoured in indicative constructions ($N=1079$, $FW=0.30$) (Table 3.3). This can also be seen in Figure (3.12).

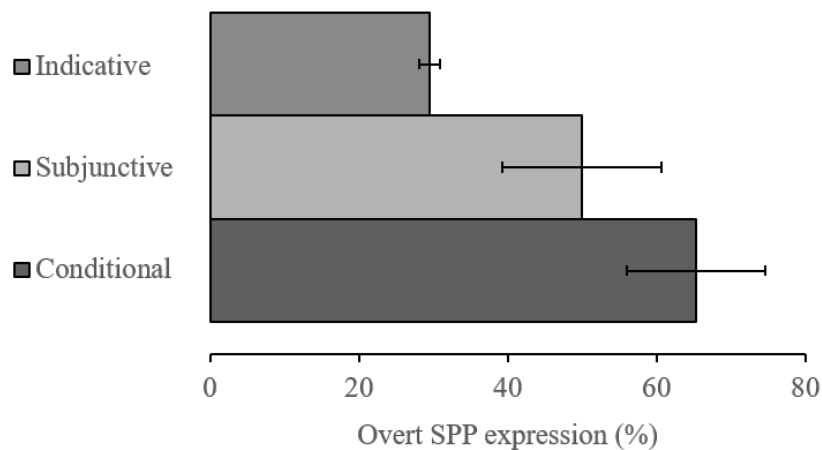


Figure 3.12: Frequency of use (+/- SEM) of overt SPPs by second-generation Colombian Spanish speakers in London, Ontario according to indicative ($N=1079$), subjunctive ($N=22$), or conditional ($N=26$) verbal mood.

As can be seen in Figure (3.12), overt SPPs are more frequent with conditional (65.4%) and subjunctive (50%) constructions, while they are significantly less frequent (29.5%) in indicative constructions ($X^2(2, N=1127)=19.318$, $p < 0.001$).

3 Interview mode

Results suggest that the mode of interview used to collect the data is a significant conditioning factor on the second-generation speakers' use of overt vs. null SPPs in variable contexts.

According to the data, speakers were significantly more likely to use overt SPPs in variable contexts when being interviewed in person and via video call, than when being interviewed over the phone ($X^2(2, N=1127)=24.412$, $p < 0.001$) (Figure 3.13).

However, it is important to note that there was no significant difference in the frequency of use of overt SPPs between in person (37.9%) and video call (32.7%) interviews ($X^2(1, N=657)=1.671$, $p=0.1962$). The observed difference above represents mostly a difference of

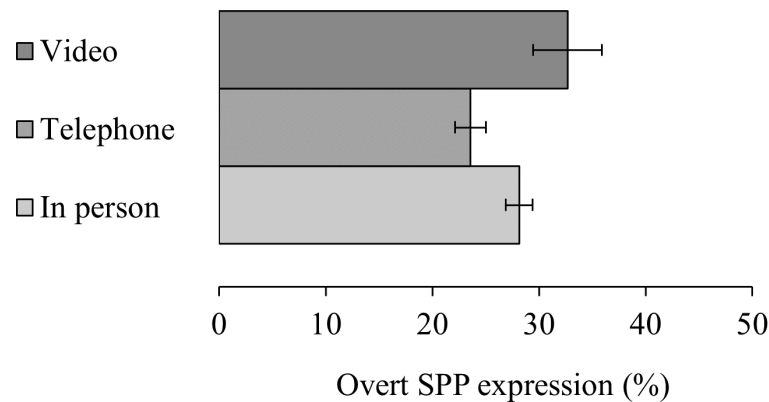


Figure 3.13: Frequency of use (+/- SEM) of overt SPPs by second-generation Colombian Spanish speakers in London, Ontario according to whether the interviews were in person ($N=446$), via telephone ($N=470$), or via video call ($N=211$).

either group in contrast with telephone interviews which show the lowest frequency of use of overt SPPs (22.9%).

Results from the logistic regression analysis using Rbrul corroborate this result indicating that in person ($N=446$, $FW=0.59$) and video call ($N=211$, $FW=0.50$) interviews favour the use of overt SPPs in variable contexts, while interviews via the telephone ($N=470$, $FW=0.41$) disfavour it (Table 3.3).

4 Switch reference

Results suggest that in the case of second generation speakers, switch reference is a significant conditioning factor for the variable expression of Spanish SPPs.

According to the results from the Rbrul analysis, the use of overt SPPs in variable contexts is favoured when there is a no continuity of referent ($N = 572$, $FW = 0.58$), while it is disfavoured in contexts where there is continuity of the previous referent ($N= 555$, $FW= 0.39$).

In addition, as shown in Figure (3.14), the use of overt SPPs vs, null SPPs in variable contexts is highest in the cases where there is no referent continuity (35.1%) and lowest in cases where there is continuity of the previous referent (26.1%).

This difference between the use of overt SPPs in utterance initial and switch reference con-

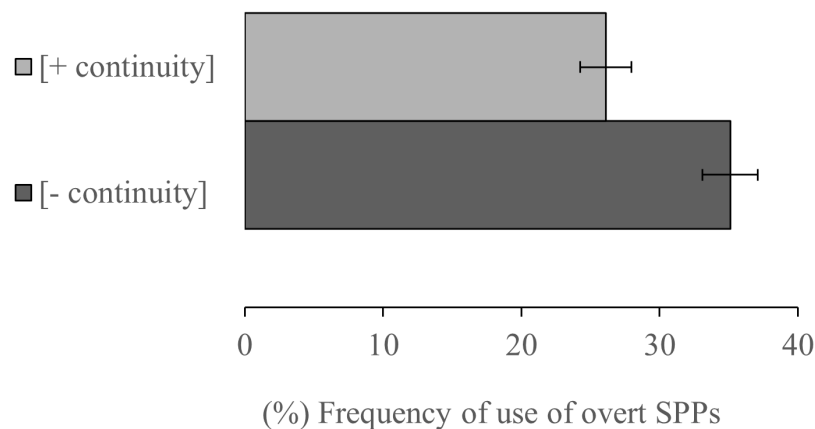


Figure 3.14: Frequency of use (+/- SEM) of overt SPPs by second-generation speakers of Colombian Spanish in London, Ontario according to whether there was a continuous referent ([+ continuity]: $N=555$), or whether the pronoun represented a new referent or a change in referent in the discourse ([- continuity] $N=572$).

texts vs. in same referent contexts is significant according to a X^2 test ($X^2(1, N= 1127)=10.757$, $p < 0.01$).

5 Clause polarity

The polarity of the clause was also found to be a significant conditioning factor which influences the variable expression of SPPs in second-generation speakers.

The Rbrul results show that the overt expression of SPPs is favoured in clauses which are affirmative and do not include a negation element ($N= 957$, $FW=0.55$), while it is disfavoured in clauses where there is negation present ($N= 170$, $FW= 0.45$) (Table 3.3).

As can be seen in Figure (3.15) above, clauses which are [- Negative] correlate with a higher frequency of use of overt SPPs in variable contexts (31.6%), which clauses which are [+ Negative] show a lower frequency of use (25.3%).

However, note that the difference in frequencies as described above is small, and it was found to be non-significant following a X^2 test ($X^2(1, N=1127)=2.750$, $p=0.0972$). This is

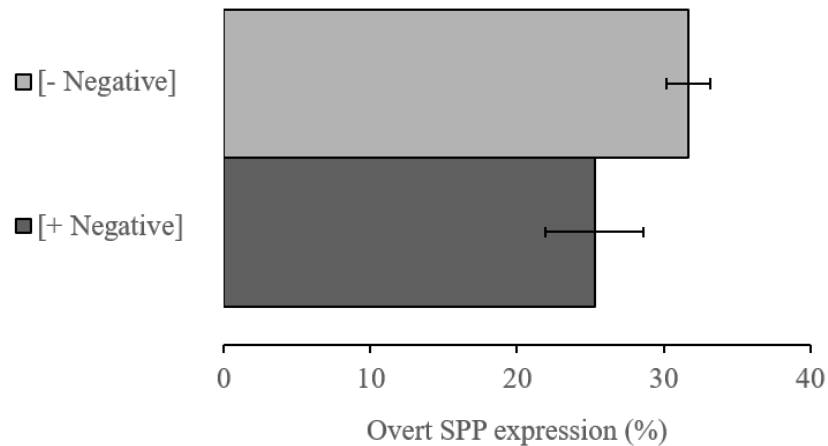


Figure 3.15: Frequency of use (+/- SEM) of overt SPPs by second-generation speakers of Colombian Spanish in London, Ontario according to whether the clause was [+ Negative] ($N=957$) or [- Negative] ($N=170$).

also reflected in the Rbrul results since the factor weights for both conditions center near the 0.5 value and both in fact point towards a small but significant effect of this factor over the variable.

6 Other factors - Not included in the multivariate logistic regression model

In this section, I briefly discuss the factors which were not selected as part of the best logistic regression model for data considering only the second generation of speakers. The factors that are described are: clause type, verb tense, semantic content of the verb, and gender. In the case of second generation speakers, since an advanced proficiency in Spanish was a requirement to participate in the interview, it's not surprising that all speakers in this generation self-reported as having an advanced proficiency level in both English and Spanish, with the exception of one speakers (2GM2) who reported his proficiency in Spanish as intermediate. For this reason, however, a comparison analysis based on proficiency was not possible for this group. Similarly no analysis on length of residence was conducted because all speakers had been living in Canada since childhood for a similar length of residence ranging from 13 to 18 years.

1. Clause type

In the current investigation, clause type was considered as a possible factor contributing to the variable expression of Spanish SPPs. The factor was coded according to whether the tokens for overt and null SPPs were found in main, subordinate, or coordinate clauses. When considering data for second generation speakers, it was determined that in variable contexts, overt SPPs were used most frequently in main clauses (34%, $N = 533$), followed by subordinate clauses (29.4%, $N = 289$), and were used least frequently in coordinate clauses (26.2%, $N = 305$). Note that although there is a trend with main clauses being associated with higher frequency of overt SPPs, this seems to be a weak association. It is also interesting to note that this trend aligns closely with the trend seen with the first generation speakers' data (see section 3.1.2.4)

2. Verb tense

The tense of verbs accompanying SPPs was a factor also considered in the current analysis. The broad categories of tenses considered in the analysis were present, past, imperfect, and future. When the data for only the second generation is considered, overt SPPs are found to be favoured and more frequently expressed with verbs in the imperfect (30.8%, $N = 130$), past (30%, $N = 220$), and present tenses (31.2%, $N = 759$), while they are disfavoured with future tense constructions (16.7%, $N = 18$).

3. Semantic verb type

Another factor considered in the analysis was the semantic content of the verbs. Verbs were coded according to three semantic categories in order to determine effects of the semantic content of the verb on variable expression of Spanish SPPs: mental, stative, and action. Results showed that overt SPP production in variable contexts was most frequent with stative verbs when second generation Colombian Spanish speakers were considered (32.6%, $N = 291$), followed by mental verbs (29.8%, $N = 319$), and finally action verbs (30.2%, $N = 517$).

4. Gender

Gender was not included within the best logistic regression model when considering the

first generation immigrant speakers. Similarly to in previous sections where either only first generation speakers or both generations' data were considered, in the case of second generation speakers, this was due to speakers from the two genders considered, men and women, being similar in their variable expression of Spanish SPPs. In the case of second generation speakers, men also used overt SPPs slightly more frequently (32.6%, $N = 552$) than women (28.9%, $N = 575$).

5. Intensity of language contact (Frequency of use of Spanish)

In order to have a measure of language contact intensity for second generation speakers, I used a cumulative score calculated from speakers' answers to questions regarding their frequency of use of Spanish in four different contexts: home/family, social-local, social-international, and professional. The scores from each of these contexts were added, and this raw score was divided by the number of possible points that each speaker could have had according to the number of questions they answered. The values calculated according to this ranged from 0.44 to 0.78 for the second generation group. Speakers' whose score was above 0.60 and therefore had higher frequency of use of Spanish were considered to be in a situation of lower intensity of language contact with English as compared to speakers who had scores at or below 0.59, which were categorized as speakers in a situation of "high" intensity contact with English. This factor was included in the multivariate regression analysis and it was not selected as a significant factor contributing to the model to fit the observed variation in use of overt SPPs. Nevertheless, the analysis showed that speakers with lower self-reported frequency of use of Spanish, which were categorized as being in "high" intensity contact with English showed a higher use of overt SPPs (38.9%, $N = 447$) than speakers which reported a more frequent use of Spanish and which were considered as being in "low" intensity contact with English (24.1%, $N = 572$). Note that these results report on data from only nine speakers since speaker 2GF5 failed to complete the online survey sent.

3.1.4 Summary of logistic regression analysis results

	First Generation		Second Generation	
	Proportion	Factor Weight	Proportion	Factor Weight
Grammatical Person				
1SG	0.287	0.61	0.335	0.66
2SG	0.340	0.64	0.103	0.29
3SG	0.253	0.60	0.443	0.78
1PL	0.086	0.24	0.124	0.29
3PL	0.173	0.43	0.204	0.47
Range		40		49
Verbal Mood				
Conditional	0.400	-	0.654	0.67
Subjunctive	0.273	-	0.500	0.53
Indicative	0.231	-	0.295	0.30
Range		-		37
Interview Mode				
In-person	0.231	-	0.379	0.59
Video	0	-	0.327	0.50
Telephone	0.243	-	0.230	0.41
Range		-		18
Switch Reference				
- Continuity	0.317	0.61	0.351	0.58
+ Continuity	0.174	0.39	0.261	0.39
Range		22		19
Clause Polarity				
- Negation	0.238	-	0.317	0.55
+ Negation	0.205	-	0.253	0.45
Range		-	-	10

Table 3.4: Linguistic and social factors conditioning the overt expression of Spanish SPPs among first and second-generation Colombian Spanish speakers living in London, Ontario.

3.1.5 Generation analysis

In this section, I carry out a comparison between the two generations of speakers with the goal of better understanding how speakers who are first generation immigrants differ from second generation immigrant speakers in regards to their use of Spanish Subject Personal Pronouns (SPPs) and more specifically the variable expression of Spanish SPPs.

Recall that all clauses where there was optionality for the use of an overt or a null subject pronoun were considered in my analyses of the data (refer to section 2.4.2). In total, 1239 tokens were extracted from interviews with first generation speakers, and 1127 from interviews with second generation speakers. The proportion of use of overt SPPs as opposed to null SPPs by speakers from each generation can be seen in figure (3.16) below.

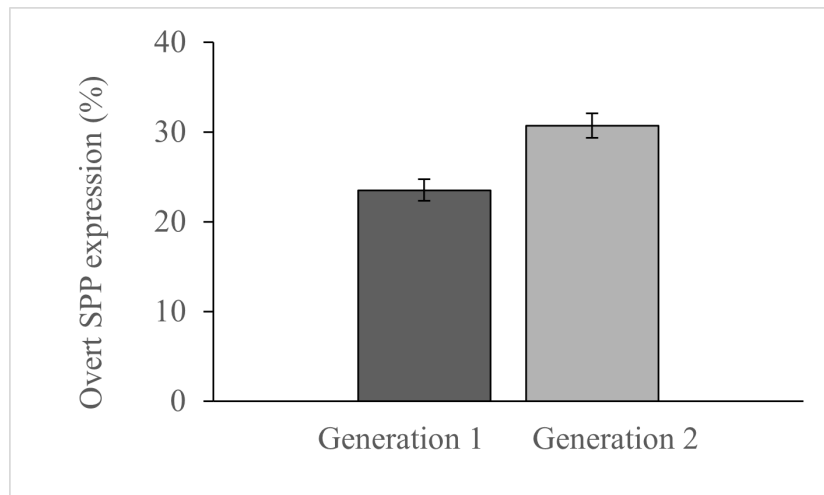


Figure 3.16: Frequency of use (+/- SEM) of overt SPPs by two generations of Colombian Spanish speakers in London, Ontario according to in first ($N= 1239$) and second ($N= 1127$) generation speakers.

As can be seen in figure (3.16), first generation speakers express SPPs overtly with a frequency of 23.5% in variable contexts, while second generation speakers do so with a higher frequency of 30.7%. This suggests a trend where second generation speakers tend to use overt SPPs with a higher frequency than first generation speakers. Further, following a X^2 analysis, the difference between the two generations groups in regards to their use of overt SPPs in variable contexts was found to be significant ($X^2(1, N = 2366) = 15.612, p < 0.001$). However, according to the multivariate regression analysis conducted in Rbrul, the difference between the two generations is not significant when the speaker factor is included as a random intercept. This suggests that the data from certain individuals may be impacting the average of one or both groups. In addition, it is important to note that when the speaker factor is not included as a random intercept in the Rbrul multivariate analysis, generation and interview mode are both

factors identified as significant in the data considering both generations. For this reason, in the following sub-section, I discuss the comparison of the two generations while further exploring the data from individual speakers for both generations.

1 Pronoun position in relation to verb

I conduct an analysis in regards to the position in which pronouns are more likely to be used by each of the two generations. In Spanish, in addition to being able to omit the phonetic realization of subject pronouns, subject pronouns when overtly used can be placed both pre-verbally and post-verbally.

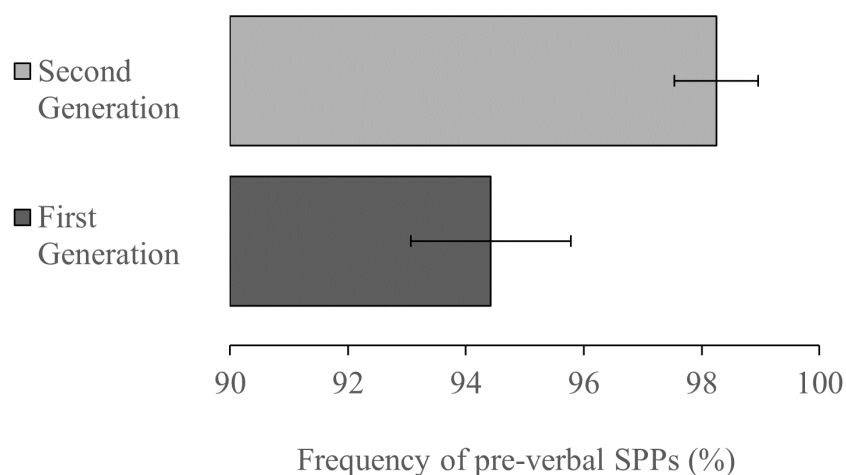


Figure 3.17: Frequency of use (+/- SEM) of pre-verbal SPPs by first ($N = 287$) and second ($N = 343$) generation Colombian speakers in London, Ontario.

An analysis of the placement of subject pronouns in relation to the verb according to generation shows that there is an increase in the use of pre-verbal subjects between the two generations of Colombian Hispanics in Canada. That is, while first generation speakers use pre-verbal subjects about 94.4% of the time, second generation speakers use pre-verbal subjects about 98.2% of the time.

As seen in (3.17), first generation speakers use pre-verbal subjects less frequently than sec-

ond generation speakers. The significance of this change between first and second generation was tested using a X^2 test which revealed that it was a significant ($X^2(1, N = 630) = 6.7858, p = 0.009$) at a significance level of $p < 0.05$.

2 Individual speakers' considerations and generation analysis

In Table (3.5), the number of tokens extracted for each interview, and the relative frequency of use of overt SPPs for all speakers whose data is included in the CoSLO corpus used in this investigation is presented. Further, the last column also includes the distance from the mean for the appropriate generation for each speakers' frequency values.

In order to further investigate the differences between the two generations, I ranked the frequency values listed in Table (3.5) and conducted a Mann-Whitney U test, which revealed that the difference between the two generations was non-significant ($Mann - Whitney U = 27, p = 0.1031$). This also suggests that it is possible that some extreme or outlying values in the data are in fact skewing the averages used in the X^2 test above which in turn showed a significant difference between the two group means.

In order to try and identify the values that could be skewing the average, I sought to test statistically for possible values that can be considered as outliers in the data when considering both generations. To do this, I used the interquartile range (IQR) of the data, and specifically the $IQR * 1.5$ rule to find outliers⁴. In order to do this I identified the first, $Q1 = 20.5$ and the third $Q3 = 31.98$ quartiles of the complete data set, and used these values to calculate the $IQR * 1.5$ value which in this case was of 17.21%. This value was then used to determine whether any values could be considered as deviating too far from the central tendency of the data. However, the range obtained was of 3.30-49.18, which does not identify any of our values as possible outliers, with the possible exception of the value of speaker 2GM1 who had a frequency of overt SPP expression of 49%. Nevertheless, even if data from speaker 2GM1 is excluded, the

⁴According to the $IQR * 1.5$ rule to find outliers, any value within a normally distributed data set that falls below the values for $Q1 - (IQR * 1.5)$ or above $Q3 - (IQR * 1.5)$ is considered to be too far from the central tendencies of the data and is therefore considered a potential outlier.

First Generation			
Speaker ID	N (tokens)	Frequency of overt SPPs	Distance from the mean
1GF1	136	10.3%	0.131
1GF2	136	28.7%	-0.052
1GF3	131	29.7%	-0.063
1GF4	111	27.0%	-0.036
1GF5	120	19.2%	0.043
1GM1	157	21.0%	0.024
1GM2	106	15.0%	0.083
1GM3	138	32.6%	-0.092
1GM4	105	28.6%	-0.051
1GM5	102	21.6%	0.019
1G-Average		23.43%	
Second Generation			
Speaker ID	N (tokens)	Frequency of overt SPPs	Distance from the mean
2GF1	115	42.6%	-0.119
2GF2	150	20.0%	0.107
2GF3	101	25.7%	0.049
2GF4	102	26.5%	0.042
2GF5	108	31.5%	-0.008
2GM1	113	49.0%	-0.184
2GM2	116	32.5%	-0.018
2GM3	108	34.2%	-0.036
2GM4	112	16.0%	0.146
2GM5	103	31.0%	-0.004
2G-Average		30.67%	

Table 3.5: Individual speaker data reported for all speakers in the first and second generation groups including number of tokens extracted from each interview, frequency of use of overt SPPs determined for each speaker, and the distance from the respective generation mean for each value.

results remain the same and according to a X^2 test, the difference between generations is still shown to be significant at a level of 0.05 ($X^2(1, 7.905, N = 2253, p < 0.05)$).

The observed difference between the two generations, however, could be due to factors not inherently related to the generation factor. For instance, when considering generational differences, second generation speakers used singular pronouns more frequently than first generation speakers. Since the use of overt SPPs is favoured with singular constructions, it is possible that the higher use of overt SPPs by second generation speakers is reflecting this difference between generations in use of singular vs. plural constructions. This same pattern where second gen-

eration speakers used constructions that favoured overt SPPs over null SPPs is also seen when considering the continuity of referent.

Second generation speakers in the sample had more instances of subject pronouns in utterance initial position and more instances of changes in referent in the discourse as compared to first generation speakers. Instances where there was no continuity of the referent favour the use of overt SPPs. Therefore, similarly to the pattern described in regards to the pronoun number factor, it is possible that the higher proportion of overt SPPs within the second generation is due in fact to this differences in use between generations whereby second generation speakers use fewer phrases with continuity of referent. This higher use of structures without continuity of referent may be due in part to shorter conversational turns. Anecdotally, I noticed that second generation speakers were less likely to enter longer narratives as part of their responses to questions in the interview. However, I did not formally account for length of utterance in this analysis, and this is a factor that will need to be considered in future studies.

3.1.6 Quantitative Results - Summary

In this first half of this chapter, I presented the results of a qualitative analysis of the data collected in this study, as well as the results of a generation analysis through which I compared the differences between the two generations of speakers in further detail.

In order to determine the social and linguistic factors that condition the variable expression of Spanish SPPs, I conducted a multivariate logistic regression analysis using data for both generations of speakers, as well as for the first generation, and for the second generation of speakers independently.

When we considered the grouped data from both generations, Rbrul results showed that only four linguistic factors are significant in contributing to the variable expression of SPPs (Table 3.1; Table 3.4). Specifically, when we consider speakers from both generations the most important factor is grammatical person of the pronoun (section 3.1.1.1), followed by verb mood (section 3.1.1.2), switch reference (section 3.1.1.3), and polarity of the clause (section 3.1.1.4).

Further, no social or extra-linguistic factors were found to be significant.

In contrast, when we considered data from first-generation Colombian Spanish speakers alone, results from the Rbrul analysis showed that only two linguistic factors, grammatical person (section 3.1.2.1) and switch reference (section 3.1.2.2), are significant in conditioning the variable expression of Spanish SPPs for this group (Table 3.2; Table 3.4). Similarly to when we considered data from both generations of speakers, no social factors were found to condition these speakers' use of overt vs. null Spanish SPPs.

Finally, when we considered data from second-generation Colombian Spanish speakers, the results showed that four linguistic factors and one extra-linguistic factor significantly condition their variable expression of Spanish SPPs (Table 3.3; Table 3.4). Similarly to the first generation, the factor of grammatical person (section 3.1.3.1) was the most important conditioning factor among this group of speakers. However, in this case, the next most significant factor was verbal mood (section 3.1.3.2), followed by interview modality (3.1.3.3), switch reference (section 3.1.3.4), and clause polarity (section 3.1.3.5). This represents not only an increase in the number of factors that condition the variable expression among second generation speakers as compared to their first generation counterparts, but also signals difference in the hierarchy of the conditioning factors.

The generational analysis revealed that second generation speakers in my sample used overt SPPs with a higher frequency than first generation speakers. However, this difference was found to be non-significant according to the Rbrul analysis, and the possibility of having a few speakers altering the average did not sufficiently account for the lack of significance. Instead, my analysis showed that other factors, unrelated to the generation factor, may be at play to explain the differences between generations. Specifically, second generation speakers used constructions that tend to favour the use of overt SPPs more than first generation speakers such as constructions without continuity of referent and shorter conversational turns. This in its own may be able to account for the higher frequency of use in this group of speakers.

In addition, in the generational analysis, I also considered differences according to speak-

ers' use of preverbal and postverbal subject pronouns. The results showed that second generation speakers use overt SPPs in preverbal position significantly more often than first generation speakers.

In conclusion, although the overall rate of use of overt SPPs was not found to differ significantly between generations, my quantitative results show important differences in the variable expression of Spanish SPPs between first and second generation Colombian Spanish speakers. Notably, second generation speakers' variable expression of Spanish SPPs is conditioned by more conditioning factors than their first generation counterparts, the hierarchy of the conditioning factors differs between generations, and second generation speakers use overt SPPs in preverbal position more often than first generation speakers. That being said, it is also important to highlight the similarities between the generations. For instance, in both generations the grammatical person of the pronoun and the continuity of the referent were important factors conditioning the use of overt Spanish SPPs.

In the following section, I will move towards a qualitative analysis of the data collected. This qualitative analysis will provide further information regarding the socio-cultural aspects surrounding the investigation, and it will inform the quantitative analysis allowing for a deeper analysis and understanding of the quantitative results just presented in this section.

3.2 Qualitative Analysis: Language Maintenance

The qualitative data analysis conducted for the current study is deductive in nature. I hypothesize that due to the fact that the presence of Spanish in Canada is much more recent than in the U.S., and that there is a lesser influence from Hispanic culture in Canada among other factors, the use of Spanish in Canada is likely to result in language shift that could occur in as little as two or three generations.

This hypothesis is addressed deductively by assessing the attitudes, experiences and overall comments regarding Spanish in Canada for all 20 speakers and determining whether these align with factors known to be associated with maintenance of a minority language in a situation of language contact (see section 2.6.1 in the Methodology chapter for detailed description of all factors considered). The specific factors considered were: use of Spanish in social and professional contexts; perception of the importance of Spanish by speakers themselves and by the wider community in London, Ontario; attitudes towards exogamous or endogamous relationships; and cultural association with Colombia or Hispanicity in general (in the case of second-generation speakers only). These factors and the questions used during the sociolinguistic interview to directly address them are outlined in Table 3.6 below, which is recreated from Table 2.6 in the methodology chapter.

After conducting the thematic analysis, a number of general themes and sub-themes were identified for each of the above factors. These themes inform us about the attitudes of Spanish speakers in London, Ontario in regards to their language, and their language use. Notably, the analysis revealed that speakers use Spanish mostly at home, but there are important differences on how first and second generation speakers use their Spanish and English language in the home, social, and public contexts, with second-generation speakers mostly using Spanish when communicating with first generation speakers. However, no generational trends were seen when considering the use of Spanish in professional contexts. The use of Spanish seems to vary instead with others' perceptions of Spanish in the workplace and with the opportunities to speak Spanish that are available due to informal interactions with other Hispanics.

Factor/Topic	Interview Question(s)
Use of Spanish in different contexts (home, social, professional, in public).	Do you use Spanish at home? Do you speak in Spanish with your Hispanic friends? Do you use Spanish at work/school? Has Spanish helped you with your work?
Attitudes towards inter-generational language transmission.	Do you think it is important that the children of Hispanic families in Canada learn to speak/read/write Spanish? Why or why not?
Attitudes toward endogamous/exogamous marriages.	Do you think it is important to marry someone who is Latin@/Hispanic? Why or why not?
Access to, or knowledge of community-led resources for the maintenance of Spanish.	Are there any resources/groups available to help you/your children learn and practice Spanish?
Cultural association with Hispanicity or Colombia (G2 only)	Do you consider yourself more Colombian or Canadian?

Table 3.6: Questions used to address factors contributing to language maintenance for the qualitative thematic analysis.

The trends that I identified in the analysis, while considering the attitudes of speakers towards intergenerational transmission, suggest that while the maintenance of the Spanish language is important to speakers, many deem this effort too difficult to achieve. On a connected theme, in regards to the availability of Spanish-learning opportunities in the community of London, Ontario, speakers varied in their responses regarding whether they knew about available resources or opportunities to learn and practice Spanish. However, even among speakers who were aware of such opportunities or resources, many mentioned not using them, and all speakers highlighted that the home was the primary source of language-learning opportunities for Spanish.

Further, in regards to the consideration of linguistically endogamous or exogamous relationships, speakers mostly considered that having a romantic partner who was linguistically endogamous was less important than having a partner who was culturally compatible or who possessed other personal values or qualities. Nevertheless, some second generation speakers consider linguistic endogamy important, but only in order to facilitate communication with their Hispanic family members. Finally, when I considered the connection of second-generation speakers to their cultural background, a major theme was a feeling of divided identity with

many speakers reporting a feeling of not belonging to either Colombian or Canadian communities fully. However, due to their increased contact and experiences with Canadian culture, as well as to their language abilities in English, the majority of second generation speakers concluded that they considered themselves to be more Canadian than Colombian. This perspective also seemed to be influenced by perspectives of others in the Hispanic or Canadian communities in which they interact, with instances of labeling or discrimination increasing or awakening their feelings of “not belonging” in either community. Further details of these results from the thematic analysis are presented in the sections that follow.

It is important to highlight that, within the investigation on the variable expression of Spanish Subject Personal Pronouns, the incorporation of a qualitative analysis further informs us about the community being considered in the investigation, and helps us to better understand the specific cultural and social aspects of the situation of language contact in London, Ontario. This is an essential consideration since language contact situations and connected linguistic effects vary according to socio-historic factors.

In addition, informing the quantitative analysis with qualitative data may also reveal patterns involving social and extra-linguistic factors that intersect or influence previously considered linguistic factors. I thus also adopt an explanatory mixed-methods approach whereby I consider information from the gathered qualitative data in order to explain results, trends, and patterns found at the quantitative analysis stage (Manzoor 2016). The explanatory mixed-methods analysis in regards to individual speaker variation will be included in Section (3.3) following the qualitative analysis.

3.2.1 Spanish language use in various contexts in Canada

In this section, I considered the speakers’ answers to questions that related to their use of Spanish and English at home, with friends, in the workplace, and in public. The main themes identified through this analysis are displayed in the table below.

Theme	Subtheme
At home	
Spanish with 1st Gen speakers.	- Spanish as a rule by 1st Gen. - English or “Spanglish” with siblings and younger children.
In social contexts	
Spanish with and among 1st Gen speakers	
Spanish rare among 2nd Gen speakers.	- English mostly with friends. - Spanish for specific purposes.
In professional contexts	
Spanish used informally at work.	
Spanish not beneficial at work.	- Spanish not needed for work. - Spanish discouraged at work.
In public	
Depends on company	
Considering others’ perceptions	- 1st Gen: English for respect - English for fear
2nd Gen: Spanish for specific purposes.	

Table 3.7: Themes identified through thematic analysis of conversations with first and second generation speakers regarding their use of Spanish at home, in social situations, in professional contexts, and in public.

In Table (3.7) I outline some of the most salient themes that emerge from the analysis of the responses by first and second generation speakers when asked about their use of Spanish and English in specific contexts.

1 The home environment

In the current analysis, the home context included all communication which occurred inside the home and across different generations. The questions asked were modified in each case according to the family with whom each individual resided. For instance, all second-generation speakers at the time of the interview resided with parents and/or siblings, while all first-generation speakers at the time of the interview resided with their spouse and children. The questions that addressed the home environment were adjusted accordingly in each case.

The analysis revealed one major theme and two sub-themes for speakers’ use of Spanish at home. The main theme identified was that Spanish is used almost exclusively when a first-

generation speaker is participating in the conversation. The two sub-themes identified were that first-generation speakers, the parents of second-generation speakers, consistently attempt to establish Spanish as the home language (40) and try to enforce it by disciplining first-generation speakers (41, 42). The second subtheme that is identified through analysis of the data is that English is the language most frequently used within the home when speaking to siblings or younger family members. In some exceptions, a mixture of English and Spanish is used as well.

Evidence of the first subtheme, where there are attempts by first-generation speakers to establish Spanish as the language spoken in the home can be seen in the excerpts below (40-42).

(40) *[...] desde chiquitos, mi mamá siempre quiso que el español fuera el idioma que se hablara en la casa. (2GM3).*

“[...] since we were little, my mom always wanted Spanish to be the language that would be spoken at home.”

(41) *[En la casa] Español. Mis hijos yo los regañé siempre cuando hablaban entre ellos en inglés. (1GF4)*

“[At home] Spanish. My children, I always scolded them when they spoke English amongst themselves.”

(42) *Pues en mi casa no se habla inglés. No se puede. (2GF5)*

“Well, in my house English isn’t spoken. It’s not allowed.”

As the examples above show (40-42), both first and second-generation speakers explain that Spanish is considered the language of the home. There is also evidence that first-generation speakers, the parents of second-generation speakers, attempt to enforce this by disciplining them (41) or making it clear that the use of English is not allowed (42).

Despite this, however, both first and second-generation participants noted that second-generation speakers use Spanish mostly when communicating with, or when being supervised by first-generation speakers. This is the primary theme identified through analysis of the home environment. For instance, the same speaker who noted that Spanish was not allowed at home in example (42) above, also noted that she communicates in English with her brother and that he only sometimes speaks in Spanish with their parents (43).

(43) Investigadora: *¿Y tu hermano le habla en español o en inglés a tus papás?*

2GF5: *A veces en inglés. Sí, a veces en inglés y a veces en español.*

Investigadora: *¿Y con [tu hermano] se hablan en español o en inglés?*

2GF5: *En inglés.*

Interviewer: “And your brother speaks to your parents in Spanish or English?”

2GF5: “Sometimes in English. Yes, sometimes in English and sometimes in Spanish.”

Interviewer: “And with [your brother] do you speak in Spanish or in English?”

2GF5: “In English.”

When asked about the language used between second-generation speakers, first-generation and second-generation speakers alike mentioned that the language used between second-generation speakers, for instance between siblings (as in example 43 above), was English or a combination of English and Spanish. This was the second sub-theme identified and can be evidenced in various excerpts in the data (43-47). In the excerpt below, a second-generation speaker explains this clearly by stating that in general they speak in English between siblings and only when the parents are present do they use Spanish.

(44) *Con mis hermanos y hermanas hablamos casi todo el tiempo. Diría el noventa y cinco por ciento en inglés y el cinco por ciento en español. Pero ese cinco por ciento es*

porque estamos con mi papas más que todo [...] Pero por lo general inglés. (2GF4)

“With my brothers and sisters we speak almost all the time. I’d say ninety-five percent in English and five percent in Spanish. But that five percent is because we are with my parents mostly [...] But in general English.”

It is important to note, that this type of situation in the home environment where parents (first-generation speakers) attempt to enforce Spanish use in the home, but children (second-generation speakers) use Spanish only with them and other first-generation speakers while opting to communicate in English with younger family members was described by all speakers in both generations. For this reason, this was identified as a primary theme in the analysis of language use in the home environment.

Some variation was observed in regards to reports of the quantity of English used in the home. Notably, while some second-generation speakers reported using English exclusively with their siblings when not in the presence of the parents (44), others reported using a combination of English and Spanish, or “Spanglish”, with both the parents and siblings (45-46).

(45) *Se vuelve como un revuelto a veces [...] mi hermana y yo contestamos en inglés, mi mamá nos dice algo en español y es así. (2GM3)*

“It turns into like a mix sometimes [...] my sister and I answer in English, my mom says something to us in Spanish and it’s like that.”

(46) *Más que todo hablamos en español cuando estamos con mis papás [...] y también Spanglish, ¿No? cuando nos comunicamos. (2GF3)*

“Mostly we speak in Spanish when we are with my parents [...] and also Spanglish, no? When we communicate.”

This sub-theme extends further when we consider conversations held with younger children and third-generation speakers in the family. For instance, when speaking to nieces or nephews, second-generation speakers reported that they usually end up speaking with them in English to

ensure they'll understand (47), or that the third-generation children will not reply in Spanish when addressed (48).

(47) *Pues, trato de hablar en español para que aprenda, pero o sea cuando necesito que me entienda en inglés porque ella, sí, no habla mucho español. Sabe algunas palabras y nos llama 'tío' pero... (2GM4)*

“Well, I try to speak in Spanish so that she will learn, but in English when I need her to understand, because she, yeah, she doesn't speak much Spanish. She knows a few words and calls us 'uncle' but...”

(48) *Pues [mis sobrinas], ellas entienden el español. No contestan en español pero lo entienden porque mi hermana siempre les habla en español, su esposo les habla en español, pero no, no las esfuerzan, no las esfuerzan a responder en español. (2GF1)*

“Well [my nieces], they understand Spanish. They don't reply in Spanish, but they understand it because my sister always speaks to them in Spanish, and her husband speaks to them in Spanish. But no, they don't push them, they don't push them to answer in Spanish.”

Communication with third-generation speakers in English more than in Spanish was similarly reported by first-generation speakers communicating with their grandchildren. In total, six speakers mentioned at some point during the interview that a third-generation child in their family had limited or no speaking abilities in Spanish. It is important to note that this represents all mentions of third-generation Hispanic children in the sample.

The pattern uncovered by this theme described above suggests a gradual decrease in the use of Spanish at home, with most Spanish being used only when addressing first-generation speakers who either enforce it or need it due to limited English-speaking abilities. In some cases, it is already creating a generational divide where first-generation speakers are unable to communicate with their grandchildren (49).

(49) *Tengo que hacerle la fuerza al inglés, porque [mis nietos] les trato de hablar español, medio entienden, pero me contestan en inglés, o sea que si yo quiero que realmente me entiendan, me toca hablarles en inglés. Y cuando me quieren contar algo, están emocionados y entonces me hablan en inglés y me toca entender qué me están diciendo. (1GF5)*

“I have to push hard with English, because [my grandchildren] I try to speak to them in Spanish, they kind of understand, but they reply in English, so if I want them to actually understand me, I have to speak to them in English. And when they want to tell me something, they are excited and they speak to me in English, and I have to understand what they’re telling me.”

The fact that within our sample all mentions of third-generation speakers revealed limited or no Spanish speaking abilities and that all second-generation speakers mentioned rarely use Spanish with siblings, may suggest that despite the attempts of first-generation speakers, Spanish is only used to a limited extent in the speakers’ homes.

2 Social contexts

Similar themes were identified when social contexts were considered. There were two main themes identified. The first theme is that Spanish is used mostly with and among first-generation speakers, which follows the pattern seen when analysing the home environment, and the second theme is that Spanish use in social contexts is rare among second-generation speakers. Within this second theme, two sub-themes were also identified. Notably, that English is the language used by second-generation speakers in social situations, and that Spanish is reserved by these speakers for use for specific purposes or specific situations.

The first theme identified in the social context follows the pattern seen in the home context. That is, that Spanish is used mostly with and among first-generation speakers.

For instance, first-generation speakers communicate solely in Spanish with friends and tend

to have social groups that are composed mostly or entirely of Hispanic friends (50).

(50) *Cuando llegamos acá, hicimos un grupo de amigos que, pues, se volvieron nuestra familia [...] Todos pues colombianos. [...] y todo lo que sucede ahí con ellos es en español.* (1GF3)

“When we arrived here, we made a group of friends which, well, they became our family [...] All, well, Colombian. [...] and everything that happens there with them is in Spanish..”

This situation was reported by six out of the ten first-generation participants who were consulted. All other first-generation speakers reported not having time for social activities, or being close only with family. This then highlights that first-generation speakers tend to form close friendships mostly with other first-generation Spanish speakers and in these encounters Spanish is the language used.

In contrast, second-generation speakers who at the time of the interview were all under the age of twenty-five, reported having large social circles, and their social connections varied from one individual to the next in regards to their cultural and language background. This is unsurprising given the multi cultural nature of the larger community in London, Ontario. However, regardless of whether their social circle was composed of English or Spanish-speaking friends, second-generation speakers expressed that they tend to communicate in English when in social situations (51).

(51) *Inglés casi todo el tiempo. Pues entre nosotros y entre, pues, nuestros amigos que también son latinos. Pues porque yo no sé. Desde pequeños uno empieza a hablar inglés en el colegio y después ya cuando uno no está en el colegio, ya está acostumbrado.* (2GM1)

“English almost all the time between us and between our friends who are also Latino.

Well, because, I don't know, ever since you're little you start speaking English at school and then when you're not in school you're already used to it."

Within our sample, eight out of ten second-generation speakers mentioned having a social group which included other Hispanic friends. The remaining two noted that their social group is made up mostly of friends from different nationalities who don't speak Spanish. More importantly, however, is that all speakers reported speaking mostly English with some exceptions.

These reported language habits thus highlight the second theme in our analysis which notes the rarity of Spanish use among second-generation speakers.

Among the exceptional cases where Spanish is used by second-generation speakers, two were most frequently mentioned in our data. Firstly, out of the eight second-generation speakers who reported interacting with Hispanic friends, four specifically mentioned that while they use mostly English, they will use Spanish with Hispanic friends when these friends are newcomers to Canada, and thus first-generation speakers (52 - 53).

(52) *...entonces ellos llegaron hace como dos años, un año. Entonces el inglés de ellos no es tan, tan avanzado, ¿right? Entonces, ¿para mí?, yo me defiendo más con el inglés, entonces cuando estoy con ellos me toca como esforzarme más en español.*
(2GF2)

"...so they arrived about two years ago, or one year. So their English is not as, as advanced, right? so, for me? I can defend myself better with English, so when I am with them I have to make a stronger effort in Spanish."

(53) *...porque yo sé más inglés que español, y si ellos saben también más inglés que español, pa' qué hacernos pa-, los payasos. Pero, pero sí. Si son hispanos y es para, para poder entendernos mejor, en español.* (2GF1)

"...because I know more English than Spanish, and if they also know more English than

Spanish, why act like clowns. But, but yes. If they are Hispanic and it's so that, so that we can understand each other better, in Spanish.”

These examples further highlight the first theme above that Spanish is only used when communicating with or among first-generation speakers. Secondly, second-generation speakers reported using Spanish for specific purposes.

The use of Spanish by second-generation speakers for specific purposes represents our second sub-theme for second-generation speakers. Notably, all second-generation speakers mentioned that they use Spanish to communicate with Hispanic friends when they want to ensure others nearby will not understand their conversation.

(54) *...a menos de que estemos hablando en público de algo privado, o algo que no queramos que la demás gente entienda, pero más que todo siempre es en inglés. (2GF3)*

“...unless we are talking in public about something private, or something that we don't want that other people understands, but mostly it's always in English.”

(55) *...o cuando digamos estoy hablando por teléfono y no quiero la gente al lado mío entienda lo que estoy diciendo entonces hablo en español con [mis amigos] pero, pero por lo general inglés. (2GF4)*

“...or when, let's say I'm speaking on the phone and I don't want that people beside me to understand what I'm saying, then I speak in Spanish with [my friends] but, but in general English.”

This use of Spanish for privacy purposes was expressed by all second-generation speakers. In addition to using Spanish for privacy, two second-generation speakers also stated that they use Spanish with other Hispanic speakers since they see this as an opportunity to practice their language skills as shown below in 56.

(56) *Cuando conozco a alguien que es hispano me fascina porque puedo hablar con*

ellos en español. (2GF4)

“When I meet someone who is Hispanic, I love it because I can speak with them in Spanish.”

This therefore shows the two primary cases when second-generation speakers reported using Spanish when in social situations. Second-generation speakers will tend to communicate in English, unless in the presence of a first-generation speaker, or when it is needed for specific purposes which can include increased privacy or to get practice with the heritage language.

3 Professional Contexts

In the case of professional environments, whether this represented a place of work, or an educational institution (College and University, except for one case of a grade 12 student still in high school) in the case of younger first-generation speakers, the primary theme that emerged is that Spanish is used only informally in places of work. The two subthemes that emerged from this reflect the reasons why the use of Spanish is uncommon. Primarily, that Spanish is not needed for the jobs held by participants, and that the use of Spanish in some cases is considered undesirable in professional contexts where participants operate.

The main theme that emerged from our data is that Spanish is used only informally at work. That is, while most speakers reported that Spanish is not a needed skill in their work, Spanish is used regularly when communicating with other Hispanic co-workers or colleagues. This was reported by six first-generation speakers and two second-generation speakers.

(57) *Donde trabajamos, [...] ahí tengo bastante-, tengo varios del staff que son hispano-parlantes y con ellos se tiende a hablar en español. (1GM4)*

“Where we work, [...] there I have a lot-, I have several on staff that are Hispanic and with them the tendency is to speak in Spanish.”

(58) *Entonces obviamente cuando, cuando son estudiantes que hablan español pues uno*

termina, termina practicamente hablando con ellos en español. (1GM5)

“...so obviously when, when they are students who speak Spanish well, one ends up, ends up practically speaking in Spanish with them.”

The informal use of Spanish with other Hispanics with whom participants interact professionally on a daily basis was reported by eight speakers in total. This represents the first main theme in our data.

However, only two speakers reported that Spanish was beneficial in their work, or that they had benefited professionally due to their knowledge of Spanish through additional pay for bilingual work, or through increased opportunities at work related to their bilingualism (59, 60).

(59) *Sí. Yo tuve un empleo en que ac-, eh, bueno actually dos. Me pagaban más porque hablaba español. (2GF3)*

“Yes. I had a job in which-, um, well actually two. They paid me more because I spoke Spanish.”

(60) *Yo soy un cajero aquí [...], y ahí tiene sus beneficios. Unas veces cuando viene gente que es-, habla español, pues yo les hablo en español, o alguien que, que no sabe cómo hablar en es-, en inglés viene y pues a mi me llaman porque todo el mundo sabe que yo hablo español. Me llaman y pues yo les ayudo. (2GM2)*

“I am a cashier here [...], and there it has its benefits. Sometimes when people come who is-, speak Spanish, well I speak to them in Spanish, or someone who, who doesn't know how to speak sp-, in English comes and well they call me because everyone knows I speak Spanish. They call me and well I help them.”

Only three speakers reported receiving some kind of benefit at work thanks to their knowledge of Spanish. Therefore, the rare use of Spanish due to it not being beneficial at work rep-

resents the second main theme in the analysis for professional contexts. The two sub-themes that emerge reflect the two primary reasons why Spanish at work is not seen to be beneficial by most Hispanic speakers. For instance, some speakers, from both first and second generation groups, reported not using Spanish at work because they consider it an undesirable behaviour at work, or because they experienced some discrimination due to their use of Spanish at work (61, 62).

(61) *No es algo que.. de hecho me parece mucho que a mi jefe y a otras personas en la oficina no les agrada que uno hable en otro idioma. ¿No? (1GF5)*

“It’s not something that... in fact, it really seems to me that my boss and other people in the office don’t like that one speaks in a different language. No?”

(62) *Una de las managers allá se me acercó y me dijo ah, 'no nos gusta que tú hables español en el almacén porque el resto de nosotros no entendemos.' (2GF1)*

“...one of the managers there came up to me and said ah, 'we don’t like that you speak Spanish in the shop because the rest of us don’t understand’ ”

Avoidance of use of Spanish at work to avoid discrimination or disapproval from others in the place of work was reported only by three speakers (two first-generation speakers and one first-generation speaker). The remainder of participants explained that they did not use Spanish at work very frequently because it was simply not used for the specific job. Note that in the cases where speakers reported not using Spanish at work, there were no follow up questions directly addressing the extent of opportunities (whether formal or informal) to use Spanish at work, and therefore it remains unclear whether these speakers have opportunities and do not use them, or simply have no opportunities to use Spanish at work.

4 Public contexts

When analysing the data touching on participants' answers to questions regarding their use of Spanish in public contexts, I hoped to gain further information regarding Spanish speakers' perceptions of the larger community's attitudes towards the use of Spanish. After completing the analysis on this data, three major themes emerged: use of Spanish in public depends on company, includes consideration of others' perceptions, and for second generation speakers it is done for specific purposes only.

The first major theme that emerged from the analysis is that the use of Spanish by Spanish speakers depends largely on who the speakers are with. Generally, both first and second generation speakers considered this to be the most important factor in determining whether English or Spanish would be used in a public context (63-65).

(63) *No sé, depende. [...] Normalmente yo me acomodo a que la otra persona también se sienta cómoda. ¿No? (2GM3)*

"I don't know, it depends. [...] Normally, I adapt so that the other person also feels comfortable. No? "

(64) *Si uno está o sea en el, en el círculo en que estamos nosotros cuatro solamente y hay más gente alrededor, pues sí hablamos en español. Pero si ya dentro de nuestro círculo está una persona que, que no habla español, pues ya hablamos inglés. (1GM5)*

"If one is, like in the, in the circle in which us four are and there is more people around, well yes, we speak in Spanish. But if now within our circle there is a person who, who doesn't speak Spanish, well then we speak in English. "

(65) *Si hablo en inglés es porque hay alguien con el que yo hablo que solo hable en inglés. Pero si yo estoy con la gente la que habla español, en un lugar público ¿y que me pongo a hablar en es-, en inglés? No. (1GF4)*

“If I speak in English it’s because there is someone with whom I’m speaking who only speaks English. But if I am with people that speaks Spanish, in a public place and that I begin to speak in Sp-, in English? No. ”

This viewpoint highlighted through the excerpts above (63-65) is one of the most common viewpoints held by both first and second generation speakers in our sample, with a total of six out of twenty speakers citing this as the most important consideration in deciding which language to use in public, and thus represents the first major theme identified in the analysis of public contexts.

Alternatively, however, several speakers also reported that they consider the perception, attitudes, or feelings of those around them when deciding which language to use in public settings. This represents our second largest group and is sub-divided into two sub-themes. The first sub-theme highlights an opinion held by only first-generation speakers. First-generation speakers stated that they consider that English needs to be used in public to be respectful to others who may not understand their conversation, or who may feel uncomfortable or misunderstand gestures due to their inability to understand what is being said. This represents the first sub-theme and can be evidenced in the following excerpts (66, 67).

(66) *Me preocupa que la gente que está al lado mío piense, si uno se ríe, que se está burlando de ellos. (1GF5)*

“It worries me that people who are beside me think, if one laughs, that one is mocking them.”

(67) *Se debe hacer el esfuerzo de hablar en inglés simplemente por respeto a la persona que está al lado. (1GM4)*

“The effort must be made to speak in English simply out of respect towards the person who is besides you.”

Although this was a common viewpoint identified within our data, with four of ten first-

generation speakers reporting it as their primary consideration in deciding which language to use in public, it is important to note that no second-generation speakers reported this as a major consideration.

The second sub-theme identified within the theme of considering others' perceptions when in public, highlights some speakers' avoidance of Spanish in this context due to fear of being discriminated against or due to feeling uncomfortable doing so.

(68) *Por la tanta discriminación que hay [...] el que hable español pues a veces tiene problemas.* (1GM2)

“Because of how much discrimination that there is [...] the one who speaks Spanish well sometimes has problems.”

(69) *Hoy en día hay lo que... han habido casos entonces me siento un poquito nerviosa a veces en hablando el español.* (2GF4)

“Nowadays there is the... there have been cases, so I feel a little bit uncomfortable sometimes in speaking Spanish.”

The use of English to avoid discrimination was only reported in three interviews. That being said, in at least four more interviews, speakers reported feeling uncomfortable speaking in Spanish, but deciding to continue using Spanish. This was the case of first-generation speakers who expressed wishing they were more proficient in English, but who feel more comfortable in Spanish (70), and also of those who simply preferred the use of Spanish and saw no reason to not continue to use it in public (71).

(70) *Yo preferiría hablar en inglés en un lugar público. Realmente, no me gusta hablar en español. Pero hablo en español. Sí. Porque es muy fácil para mí, si estoy con mi esposo [...] Igual si vamos a hablar en inglés, se va a notar raro.* (1GF5)

“I would prefer to speak in English in a public place. Really, I don’t like speaking in Spanish, but I speak Spanish. Yes. Because for me is very easy if I am with my husband [...] If we are going to speak in English, it will seem strange anyways.”

This was the case reported by two more first-generation speakers. However, some speakers reported not being concerned with others’ attitudes towards their use of a foreign language and choosing to continue speaking in Spanish regardless of the context in which they are.

(71) *A mi no me da pena que me vean hablando español. Pues que ellos aprendan el español si tanto quieren saber qué estamos hablando. (1GF4)*

“I’m not embarrassed to be seen speaking Spanish. Well, let them learn Spanish if they so badly want to know what we are talking about.”

Additionally, in the case of second-generation speakers, the major theme that emerged is that they tend to use English in public (72), unless in specific situations or for specific purposes. Notably, most second-generation speakers reported using Spanish only to get more privacy in their conversations (73).

(72) *Yo creo que yo tengo tendencia a decir con el inglés en un lugar público. (2GM5)*

“I think I have a tendency to say with English in a public place.”

(73) *A veces pues también empezamos a hablar en español por... porque sí, o si es alguien que no queremos que entienda lo que estamos diciendo pues en español. (2GM1)*

“Sometimes, well, we also begin to speak in Spanish because... just because, or if it’s someone who we don’t want to understand what we are saying, then, in Spanish.”

This last sub-theme identified is similar to the one identified in regards to the analysis of Spanish use in social contexts. I consider, therefore, that this sub-theme, where second-generation speakers use Spanish as a means to avoid being understood by others around them,

is a wider theme which describes a pattern for second-generation speakers in regards to their language use.

3.2.2 Attitudes towards intergenerational language transmission

In this section, I discuss the results of the thematic analysis conducted regarding the topic of intergenerational language transmission.

The attitudes towards intergenerational transmission of the heritage language of first and second-generation Colombian Spanish speakers in London, Ontario is analyzed by considering the answers of speakers when questioned about the importance of speaking Spanish for the following generations.

The thematic analysis conducted on this set of data revealed two main themes. These are shown in the table below (Table 3.8).

Theme	Sub-theme
The maintenance of Spanish is important.	- Value of bilingualism. - Family, culture, and identity.
Maintenance of Spanish is difficult	- Value of French. - Strong English influence. - Pressure on 1st Gen to learn English.

Table 3.8: Themes identified through thematic analysis of conversations with first and second generation speakers regarding their attitudes towards the importance of heritage language maintenance.

In this section I discuss each of the identified themes and corresponding sub-themes. Two main themes emerged during analysis. Specifically, both first and second generation speakers reported that maintaining the use of Spanish in future generations for reasons regarding either the value of bilingualism in general, or the importance of maintaining family, cultural, and identity connections. However, the second theme identified relates to speakers' comments on the difficulty of realistically achieving intergenerational language transmission to maintain the use of Spanish in future generations. Speakers reported reasons tied to the importance of French as a second language, the strong influence of English, and the pressure felt by first

generation speakers to learn and use English.

1 Maintenance of Spanish is important

Within the analysed sample, all speakers, from both first and second generations, reported that they considered that it was important that their children, or the children of other Hispanics in Canada, learn and maintain the use of Spanish (74-76).

(74) *A mi me parece importantísimo.* (1GF4)

“To me it seems very important.”

(75) *Me parece muy importante y necesario.* (1GM1)

“It seems to me very important and necessary.”

(76) *Por supuesto [...] que puedan hablar y entender español. Sí. Eso es prioridad número uno.* (2GM5)

“Of course [...] that they can speak and understand Spanish. Yes. That is priority number one.”

All speakers reported maintenance of Spanish in future generations as an ideal goal. However, only first-generation speakers expressed unanimously that it was important that future generations learn to read and write in Spanish. Within the second generation, speakers explained that while being able to read and write in Spanish would be ideal, it wasn't necessary.

(77) *Pues si aprenden a escribir [español] chévere. Si no quieren, ¿Quién los obliga?*
(2GF5)

“Well, if they learn to write [Spanish], cool. If they don't want to, who's gonna force them?”

- (78) *No es que se me haga súper importante que mis-, que en el futuro mis hijos aprendan a leer y escribir [en español]. (2GM3)*

“It doesn’t seem to super important that my-, that in the future my children learn to read and write [in Spanish].”

It is important to note, however, that this attitude, although present within the second-generation of speakers, was not unanimous and was only reported by three second-generation speakers. In contrast, some second generation speakers aligned more with first generation speakers in noting that learning to read and write was as important as learning to speak the language.

- (79) *La ventaja mía es que yo leo, entonces me ha ayudado a mantener un poco de mi, de mi vocabulario o a extender un poco mi voca-, mi vocabulario. (2GF2)*

“My advantage is that I read, so it has helped me to maintain some of my, of my vocabulary or to extend a bit my voc-, my vocabulary.”

- (80) *Es muy importante si uno-, para mantener el lenguaje, si uno lo sabe hablar, escribir y leer. Y así uno lo mantiene mejor y, y sí, uno como que forma esa conexión. (2GF3)*

“It is very important is one-, to maintain the language, if one knows how to speak, write, and read. And in that way one maintains it better and, and yes, one kind of forms that connection.”

- (81) *Si uno va a hablar un idioma, es importante que lo hable bien y que lo lea y lo escriba. (1GF1)*

“If one is going to speak a language, it is important that one speaks it well, and that one reads it and writes it.”

This attitude, in fact was the most popular. With 16 speakers, spanning both generations,

reporting that it was important to learn to read and write in addition to learning to speak (79-81). In fact, as seen above, some second-generation speakers reported that being able to read and write has helped them in maintaining their own use of Spanish (79, 80).

Sub-themes that emerged within this primary theme were related to the reasons for maintaining Spanish. Specifically, these are divided into two main groups. While some speakers noted various reasons related to the value of bilingualism in general, such as cognitive and job-related benefits, others highlighted the need to maintain a close connection with family, culture, and identity through the use of Spanish.

(82) *Saber como más de un lenguaje es muy bueno para el cerebro y yo he oído que niños que hablan más de un lenguaje o idioma [...] creo que aprenden más facil. (2GM2)*

“Knowing like more than one language is very good for the brain and I have heard that children that speak more than one language [...] I think that they learn more easily.”

(83) *El saberlo les va abrir a ellos puertas. Simplemente desde el punto de vista laboral creo que lo deben tener. (1GM4)*

“The fact of knowing it will open doors for them. Simply from a job-related point of view, I think they should have it.”

Reasons for maintaining Spanish that were related to cognitive benefits or to benefits in the job market were reported by a total of twelve speakers, making this an important consideration for speakers of both generations, but not the only consideration.

Speakers from both generations also reported that maintaining close ties with Hispanic family, culture, and identity was an important reason to maintain the use of Spanish in future generations.

(84) *Si [mis papás] no me enseñaron el español, yo no podía hablar con mi familia. (2GF2)*

“If [my parents] didn’t teach me Spanish, I wouldn’t be able to speak with my family.”

- (85) *¿Cómo no poderle hablar a mi tío, a mi prima, a mi abuela? No poder expresarme con ellos efectivamente ¿No? (2GM3)*

“How could I not be able to speak to my uncle, to my cousin, to my grandmother? To not be able to express myself with them effectively. No? ”

- (86) *El idioma es una de las cosas que te conectan a tu cultura y la cultura es muy importante tener y es bueno ser orgulloso de la cultura de uno aunque un-, no conozcas la cultura cien por ciento. (2GF1)*

“Language is one of the things that connect you to your culture and it is important to have culture, and it’s good to be proud of one’s culture even though one-, you don’t know the culture a hundred percent.”

- (87) *Si tienes claro tu identidad de esa manera para ti va a ser más fácil tratar ser feliz. (1GF4)*

“If you are clear on your identity for you it will be easier to try to be happy.”

As seen in examples (84-87), speakers reported connections with family, culture and identity as important reasons to maintain the use of Spanish among the children of Hispanics in Canada. These points of view were mentioned by 13 speakers in our sample.

2 Language maintenance too difficult

Many speakers, however, also reported that although maintaining Spanish would be important and ideal, the reality is that it may be a goal too difficult to attain. This is the second major theme identified and was noted in both first and second generation speakers.

- (88) *Sí, yo diría es importante, sí [...] Si no pasa, no pasa ¿Sí? Pero es que es muy difícil.* (2GF3)

“Yes, I would say it is important, yes. [...] If it doesn’t happen, it doesn’t happen. Yes? But it’s just that it’s too difficult.”

As seen above (88), some speakers reported that language maintenance of Spanish in future generations would be too difficult to achieve. This in particular was reported only by three speakers explicitly. However, a total of 12 speakers commented on difficulties that they have faced in achieving this goal.

The main difficulties reported, which represent the sub-themes identified, include the importance of French as a second language, the strong influence of English, and the pressure felt by first-generation speakers to learn and use English in Ontario (89-91).

- (89) *A veces yo estaba aprendiendo inglés con el niño viendo los programas entonces...*
(1GM1)

“Sometimes I was learning English with the boy watching programs so...”

- (90) *Porque después de que se pongan a aprender el inglés ya no van a querer tocar el español. No van a querer y no van a tener tiempo. Todos sus amigos van a hablarles en inglés y los profesores más inglés todavía.* (2GM5)

“Because after they begin to learn English they will not want to touch Spanish. They will not want to and they will not have time. All their friends will speak to them in English and the teachers even more English.”

- (91) *Es más probable tal vez que necesiten el francés aquí en Canadá.* (1GM4)

“It’s more likely perhaps that they need French here in Canada.”

These examples which represent the sub-themes identified in the data, reflect some of the

difficulties experienced by some first and second generation Colombian Spanish speakers in achieving maintenance of the Spanish language.

It is important to note that although all speakers reported the maintenance of Spanish an important goal, a majority of speakers expressed difficulties in achieving this goal and three speakers noted that these difficulties make the maintenance of Spanish into the next generation unlikely.

3.2.3 Attitudes towards exogamous/endogamous relationships

The attitudes of speakers towards linguistically endogamous vs. exogamous relationships can have a significant impact on the likelihood of maintaining a heritage language, in this case Spanish, within the next generation. When speakers were asked about their opinions regarding the importance of having a partner who was also Hispanic or Latin@, two main themes emerged from their responses (Table 3.9).

Theme	Sub-theme
Language is not a main consideration in choosing a partner.	- Other considerations are more important. - Importance of cultural affinity.
2nd Generation wishes to facilitate communication with family.	

Table 3.9: Themes identified through thematic analysis of conversations with first and second generation speakers regarding their attitudes towards endogamous and exogamous relationships.

Following thematic analysis, the two main themes that emerged highlighted that: 1) Language is not a main consideration when selecting a partner for the speakers within our sample; and 2) Second-generation speakers may seek to find a Spanish-speaking partner in order to help facilitate communication with other Hispanic family members. In what follows I will briefly describe each of these themes.

1 Language is not a main consideration in choosing a partner

Speakers in both generations in general reported that being Hispanic or Latin@ is not a main consideration in choosing a romantic partner (92-96).

(92) *Si es una persona que, que es buena y trabajadora, pues qué importa que sea canadiense, americana, sudafricana... como sea, eso no... a mi no... no me interesa [...]*

(1GM2)

“If it’s a person who, who is good and hard-working, then what does it matter that they be Canadian, American, South-African... as it is, not that... to me, no... I don’t care.”

(93) *No. Lo importante es que los comprendan y que se entiendan y que... que sean un complemento. Eso sería lo importante. (1GM5)*

“No. What matters is that they understand them and that they understand each other and that, that they be a complement to each other. That’s what would be important.”

(94) *No, no, no, no. Definitivamente eso no importa. No. (1GF5)*

“No, no, no, no. Definitely, that does not matter. No.”

(95) *No. Como-, me va y me viene de cualquier lado. Eh, obviamente pues, si tengo la oportunidad de conocer alguien que habla español, pues sí, genial. (2GM3)*

“No. Like-, I don’t mind either way. Eh, obviously well, if I have the opportunity to meet someone who speaks Spanish, then yes, great.”

(96) *No, a mí no me importa. La verdad no necesariamente, yo podría casarme hasta con un canadiense. Yo no tengo problema. (2GF5)*

“No, to me that does not matter. In truth, not necessarily, I could even marry a Canadian. It’s not a problem for me.”

As evidenced in the examples above, the majority of speakers noted that marrying a Hispanic or Latin@ is not a main consideration and some speakers stated that in that decision other considerations are more important, including virtues, and overall compatibility of the couple (92, 93). Ultimately, many speakers consider that finding someone who speaks Spanish would be a bonus but not a priority (95).

Within this theme another sub-theme emerged. Notably, most speakers considered that the cultural compatibility of the couple was more important than the language spoken by either party (97-98).

(97) *Mucho mejor, digamos en el sentido de la cultura como tal, complementarse con alguien que tenga como la misma afinidad y cultura, la misma música, las mismas cosas-. Es como-. Hay más afinidad ¿No? (1GM3)*

“Much better, let’s say in the cultural sense as such, to complement each other with someone who has the same affinities and culture, the same music, the same things-. It’s like-. There is more affinity. No?”

(98) *[Mi novio] es muy abierto a las diferentes culturas, entonces es mucho más fácil [...] En el pasado, con otras personas sí ha sido muy difícil, o porque hay diferencias en la cultura o porque tenemos-, es decir, diferencias en nosotros. (2GF1)*

“[My boyfriend] is very open to different cultures, so it is much easier [...] In the past, with other persons it has been very difficult, either because there are cultural differences or because we have-, that is to say, differences between us.”

It seems therefore that in selecting a romantic partner, Colombians in London, Ontario focus on considerations other than language such as personal virtues and cultural affinity. Speakers are willing to consider a partner of a different linguistic background as long as the person possesses other desirable virtues and as long as their cultural background is compatible with their own.

2 Second generation wishes to facilitate communication with family

A second major theme that emerged within our data is specifically focused on the second generation of speakers. Despite the fact that only one first-generation speaker explicitly expressed that it was important that their children marry a Spanish-speaking partner, many second-generation speakers reported that they considered that marrying a Spanish-speaking partner would be important to facilitate communication with their parents and other Hispanic family members who are not comfortable communicating in English.

(99) *El inglés de [mi mamá] todavía sigue siendo muy enredado entonces ella se frustra especialmente con la novia de mi hermano, por ejemplo, ahm, ella no habla en español tampoco. Pero, pero es más por mis papás más que todo. (2GF4)*

“[My mom’s] English is still very muddled so she gets frustrated, especially with my brother’s girlfriend, for example, ahm, she doesn’t speak Spanish either. But, but it’s more for my parents mostly.”

(100) *La verdad no, no me importa casi nada, pero es chévere que-, pues con mis papás y todo. (2GM4)*

“In truth, no, it doesn’t matter almost at all, but it is cool that-, well with my parents and all.”

As shown in the examples above (99, 100), second-generation speakers themselves may seek linguistically endogamous relationships, but may do so partly in order to facilitate interactions with their parents or other Hispanic family members, including the consideration of their own future children to whom they wish to pass the use of Spanish language.

3.2.4 Spanish language-learning opportunities

Another important factor that contributes to the maintenance of a heritage language is the availability of resources or of a community through which young speakers can continue to learn and practice their language. When asked about the availability of resources to learn Spanish in London, Ontario, the sampled group of participants are divided in their responses. While some speakers expressed not being aware of many available resources, others mentioned knowing about resources available in the community. Ultimately, however, there was an unanimous consensus expressing that the home is the main source of Spanish language learning and practice. These observations give rise to the main three themes identified within the data (Table 3.10).

Theme	Sub-theme
Few/No resources available.	
Resources are available.	- Not used.
Spanish is learned at home	

Table 3.10: Themes identified through thematic analysis of conversations with first and second generation speakers regarding their knowledge of Spanish-learning resources in the community.

The three main themes identified in Table 3.10 above, are: 1) There are few or no resources; 2) Resources are available but not used; and 3) Spanish is learned in the home (Table 3.10). In this section these themes are described in further detail and examples are provided to evidence each one.

1 Few/No resources available

A total of ten speakers reported not being aware about availability of resources to learn or practice Spanish in the community. This is evidenced in the examples below (101, 102).

(101) *No, no mucho, no. yo creo que los recursos aquí en Canadá son-, o las opciones son muy bajas para aprender el idioma. (2GM5)*

“No, not much, no. I think that the resources here in Canada are-, or the options are very low to learn the language.”

(102) *No hay muchas, pienso yo.* (1GF1)

“There aren’t many, I think.”

Within our sample, ten speakers (including speakers from both generations), thus reported not being aware of many resources available, or not being sure of how they could access them. However, in contrast with the examples listed above, the remaining ten speakers reported knowing about some resources in the community. This gives rise to our second main theme for this factor.

2 Resources are available

A total of ten out of 20 speakers explicitly mentioned knowing about or having used resources in the community to learn and practice Spanish.

The resources reported included extra-curricular activities for youth such as dance groups, and church-related activities, as well as resources to receive formal education in Spanish such as Saturday morning programs, high-school courses, and university courses (103, 104).

(103) *En nuestra religión, íbamos a una, a pues a una iglesia hispana, y hablábamos todo, todo en español y, y eso pues nos ayudó en el sentido de, de aprender a leer y a escribir.* (2GF3)

“In our religion, we went to a, well to a Hispanic church, and we spoke everything, everything in Spanish and, and well that helped us in the sense of, of learning to read and write.”

(104) *En el grupo [...] baila. (I: Ah super. Sí, eso es bonito, ¿No?) Y ella lleva cuatro años bailando, y pues me encanta porque pues todas las canciones son en español, la profe-*

sora le está hablando todo el tiempo en español y me encanta que ella ame sus culturas latinas. (1GF2)

“In the group [...] she dances. (I: Ah great. Yes, that is nice, ¿No?) And she has been dancing for four years, and well I love it because all the songs are in Spanish, the teacher speaks to her in Spanish all the time, and I love that she loves her Latin cultures.”

However, out of the speakers who reported knowing of available resources, 60% of them reported never having used them and/or choosing to not use them for a number of reasons (105-107).

(105) *Yo no, no creo que yo haya voluntariamente buscado que ellos, eh tuvieran una enseñanza en español. (2GM4)*

“I don’t, don’t think that I voluntarily sought that they, eh, had an education in Spanish.”

(106) *Yo me acuerdo cuando era chiquito mi, mi mamá y mi papá me llevaban a un colegio de es-, español para aprender el es-, el español como cada sábado o domingo y... pero no me gustó. (2GM2)*

“I remember when I was little my, my mom and my dad used to take me to a Spanish school to learn Spanish like every Saturday or every Sunday and... but I didn’t like it.”

(107) *Bueno, en el bachillerato ellos-, o en la secundaria, ellos dictaban cursos de español si quería tomarlos, pero no me interesó. (2GM5)*

“Well, in high school they-, or in secondary, they offered Spanish courses if I wanted to take them but it didn’t interest me.”

The speakers who expressed that they did not use the resources available, reported that the resources available were too basic for their level of Spanish, too small or difficult to access, or

in some cases just not finding them interesting (106-109).

(108) *No, usualmente en nuestros cursos de español en la high school era lo más básico.*

Entonces yo en sí estaba muy avanzada. (2GF3)

“No, usually in our Spanish courses in the high school it was the most basic content. So I in fact was too advanced.”

(109) *Ofrecían clases para los niños en español pero es demasiado pequeño. No es*

muy, muy conocido. (1GM1)

“They offered classes in Spanish for the children but it’s too small. It’s not very well known.”

As shown in the examples above, even in the instances where speakers report being aware of available resources, the majority express not using them or using them but not finding them challenging enough as was the case of speaker 2GF3 who found the courses available to them were too basic for her level. There were also mentions, not evidenced here, of speakers who took high school or university courses in Spanish simply knowing that they could pass the class easily.

3 Spanish is learned in the home

In general, speakers in both generations noted that Spanish in London, Ontario is principally learned and practiced in the home (110, 111). This represents the last important theme in the data when we consider the factor relating to the available resources to learn Spanish in the community.

(110) *Pues mis, mis papás, mi mamá fue la que realmente se esforzó mucho en que noso-*

tros no perdie.. perdiéramos el español. (2GM4)

“Well my, my parents, my mom was who really made a strong effort so that we wouldn’t lo-, lose our Spanish.”

(111) *Todo en la casa. Algo muy importante, de nuestra familia, mi esposa trajo películas en español, infantiles, de todo tipo. (1GM3)*

“Everything at home. Something very important in our family, my wife brought movies in Spanish, for children, of all kinds.”

As evidenced in examples (110) and (111) most speakers report having learned or having their children learn Spanish mostly in the home with little or no help from outside sources. Parents often report using Spanish-language resources such as books and movies brought from Colombia.

Therefore, it seems that the population sampled is evenly divided between those who are aware of available resources to learn and practice Spanish, and those who are not aware of availability. Nevertheless, resources when available are rarely used and most speakers report the home as the main source of Spanish input and Spanish-language resources.

3.2.5 Identification with cultural background (Second generation only)

In this section, I consider only data from the ten second-generation speakers in the sample and analyse their ties to and identification with Canadian and Colombian cultural backgrounds. Having a sense of connection to one’s cultural background can be an important element in language maintenance as it serves as a strong motivator in efforts to learn and practice the language.

The thematic analysis of the interview data from conversations with second-generation speakers regarding their sense of Colombian or Canadian cultural identity highlighted three main themes, which are displayed in the table below (Table 3.11).

As shown above, I identified three main themes through my analysis: 1) Most second-generation speakers identify as more Canadian than Colombian, in large part due to their sense

Theme	Sub-theme
More Canadian than Colombian.	- Language considerations.
Other's perspectives influencing identity.	- Labeling or discrimination experiences.
Blended or divided identity.	

Table 3.11: Themes identified through thematic analysis of conversations with second generation speakers regarding their cultural identity.

that their language skills are superior in English than in Spanish; 2) Perspectives of Colombian and Canadian community members around them contribute to their identity construction due mainly to influences from labeling or experiences of discrimination; 3) Most second-generation speakers tend to express feelings of having a blended or divided cultural identity. In this section, I will provide examples and briefly describe each of the themes identified.

1 More Canadian than Colombian

The analysis revealed that eight out of ten second-generation speakers identify as being more Canadian than Colombian, with differing weights given to each of these identities according to various factors. This is a main theme in our data and is evidenced by the examples below (112-113).

(112) *Yo me considero más canadiense que colombiano.* (2GM3)

“I consider myself more Canadian than Colombian.”

(113) *Hoy en día yo di-, diría que soy más..., un poquito más canadiense que, que colombiana* (2GF4)

“Nowadays I wou-, I would say that I am more..., a little more Canadian than, than Colombian.”

In regards to why these speakers identified as more Canadian than Colombian, many reported having a closer affinity to Canadian than Colombian customs. However, an important contributing factor to the feeling of being disconnected from a Colombian identity is their language skills as shown in the examples below (114-115).

- (114) *Si alguien me pregunta como ¿de dónde es usted? de... Yo, yo digo 'colombiano', pero eh... Yo hablo mejor inglés que español y eh, leo mejor y entiendo mejor en inglés y he vivido acá mucho más tiempo que en Colombia entonces [...] cuando pienso, creo que soy como más canadiense. (2GM2)*

“If someone asks me like “where are you from? of... I, I say ‘Colombian’, but eh... I speak more English than Spanish and eh, I read better and understand English better and I have lived here much longer than in Colombia so [...] when I think, I believe I am like more Canadian.”

- (115) *Hay cosas que yo digo y hago, y gestos que son muy pero muy colombianos y se nota [...] Pero en otro sentido sí, a veces veo por ejemplo, hablo mucho en inglés. (2GF5)*

“There are things I say and do, and gestures that are very, very Colombian and it’s noticeable [...] but in another sense yes, sometimes I see for example, I speak in English a lot.”

There is therefore, among second-generation speakers, an association between their language skills and their cultural identity. As seen above, several speakers reported that they felt less Colombian due to being more comfortable with the English language than with Spanish.

2 Other’s perceptions influencing identity

Another factor which seems to contribute to the identity formation of second-generation speakers is the perceptions of other Colombian and Canadian community members.

- (116) *Para una persona colombiana, yo creo que nosotros seríamos canadienses. (2GM1)*

“To a Colombian person, I think that we would be Canadian.”

- (117) *Mi mamá y mi papá también dicen como que.. que creen-, creo que ellos creen que*

yo soy como más canadiense. (2GM2)

“My mom and my dad also say like, that they believe-, I think that they think that I am like more Canadian.”

The examples above, show that within the Hispanic community, second-generation speakers are sometimes labeled by first-generation speakers as being “more Canadian”, which contributes to their construction of a “more Canadian” cultural identity (116-117). Similarly, comments from members of the Canadian community can have a comparable effect in the opposite direction, making speakers feel less Canadian (118).

(118) *Yo cuando era chiquita no me-, como que no me sentía diferente del resto de la, de la población canadiense y ya cuando me estaba volviendo adulta y hablaba español en público, ahí fue cuando habian unas instancias donde me hicieron darme cuenta en que ‘oh, yo no soy completamente canadiense’. (2GF1)*

“When I was little I didn’t-, like I didn’t feel different from the rest of the, of the Canadian population and then as I began to become an adult and I spoke in Spanish in public, it’s there that there were some instances where they made me realize that ‘oh, I am not entirely Canadian’.”

It is interesting to note that the perceptions of others around them from the larger community seem to contribute to the lack of connection with Canadian culture in the case of the two second-generation speakers who consider themselves to be more Colombian than Canadian (119-120).

(119) *Yo, yo me identifico como latina. Entonces, colombiana. (2GF2)*

“I identify as Latina. So, Colombian.”

(120) *Colombiano. O sea yo, yo nunca digo-, o sea, si alguien me pregunta de dónde soy,*

nunca digo 'canadiense'. [...] Soy colombiano, pero vivo en Canadá. (2GM4)

“Colombian. I mean, I, I never say-, I mean, if someone asks me where I am from, I never say ‘Canadian’. [...] I am Colombian, but I live in Canada.”

It is important to note that these two speakers who were either born in Canada or arrived before the age of five, have sensed or experienced instances of discrimination against themselves or others in their family. This may have contributed to a disconnect from a Canadian identity as they feel they have been labeled as outsiders due to their language use or other personal characteristics (121-122).

(121) *Creo que la gente automáticamente piensa que yo podía ser nacida aquí y yo soy canadiense. Pero cuando me ven ya saben que yo no soy de aquí. ¿Me entiendes? Como... Ya cuando te miran, ya saben que no eres de aquí. (2GF2)*

“I think that people automatically think that I could be born here and I am Canadian. But when they see me they know that I’m not from here. Do you understand me? Like... Once they look at you, they know that you’re not from here.”

(122) *[...] Entonces estaba caminando y había un hombre en la calle, [...] y despues estaba como gritando ‘Hey! Hey!’, entonces lo ignoré obviamente y, y despues empezó a decir groserías y despues ‘ah no me entiendes’, eh, o sea, ‘Regresa a tu, a tu país, bla, bla, bla’ [...] y yo ‘Oh my God’. (2GM4)*

“[...] So I was walking and there was a man in the street, [...] and after he was like screaming ‘Hey! Hey!’ so I ignored him obviously and, and then he began to curse and then ‘oh you don’t understand me’, uh, like ‘Go back to your, to your country, blah, blah, blah’ [...] and I was like ‘Oh my God’.”

The case of these two speakers who identified as more Colombian than Canadian were outside the general norm, however. A majority of second-generation speakers instead reported

having a sense of divided or blended identity.

3 Blended or divided identity

The complexity of factors that contribute to the construction of the cultural identity of second-generation speakers, seems to lead, in most cases within our sample, to feelings of a divided or blended identity (123-125).

(123) *Es algo que lo hace a uno sentirse como... ni ahí, ni allá, pero en la mitad.* (2GM3)

“It’s something that makes one feel like... neither here not there, but in the middle.”

(124) *Tampoco es que me siento súper comfortable en la cultura canadiense, pero tampoco es que me siento súper comfortable con un grupo de colombianos. Uno no se siente como que pertenece a un grupo específico. ¿Sí me entiendes?* (2GF4)

“I don’t feel super comfortable either in the Canadian culture, but neither is it that I feel super comfortable with a group of Colombians. One doesn’t feel like one belongs to a specific group. Do you understand me?”

(125) *Depende quién me está preguntando. No sé, estando acá me gusta decir que soy colombiano, pero estando en Colombia me gusta decir que soy canadiense.* (2GM5)

“Depends who is asking me. I don’t know, being here I like to say that I am Colombian, but being in Colombia I like to say that I am Canadian.”

Note in example (125), that the speaker reports that they express an identity opposite to that of the individuals with whom they are interacting, or the location where they are at the moment. If they interact with a Colombian individual who has recently arrived or who is a first-generation speaker, or when interacting with individuals while in Colombia, they express feeling less Colombian and therefore more Canadian. Similarly, when dealing with Canadian individuals while in Canada, they express feeling less Canadian than them and therefore more

Colombian. This further reflects the points of view expressed by the speakers in examples (123) and (124), where they express never feeling fully comfortable or belonging in a given group.

This is the most common point of view expressed by second generation speakers, with a total of eight speakers reporting mixed feelings about their cultural identity with similar statements, and represents the second major theme in the data in regards to my analysis of this factor.

3.2.6 Qualitative results - Summary

In this section, I conducted a qualitative analysis of the data collected for first and second generation speakers. I based my deductive analysis on the speakers' responses to a number of different questions which were specifically designed to address a number of factors which are known to influence minority language maintenance success, such as language use, attitudes towards intergenerational language transmission, attitudes towards endogamous/exogamous marriages, access to resources for language maintenance, and cultural association with cultural heritage or Hispanicity (see section 3.2; Table 3.6).

I began my analysis by considering how speakers in my sample use Spanish in their daily lives. Following my qualitative analysis, I identified that some of the main themes that emerge are: Spanish is used primarily with and among first generation speakers across all contexts; second generation speakers rarely use it among themselves except for specific purposes such as when wishing to have a private conversation; Spanish is not considered to be beneficial at most workplaces and is mostly used informally; and when in public, for the most part, speakers' choice of language will largely depend on the language preference of any people they are with. However, it is also important to highlight that the public context showed variability among speakers. While some speakers feel comfortable using both languages in public, others prefer using English only. This preference for English is most commonly due to a desire to be respectful to others around them, but in some cases also due to fear of discrimination.

When I considered the qualitative analysis for the set of questions which examined the speakers' attitudes towards intergenerational language transmission, the results revealed that both first and second generation speakers consider the transmission of Spanish to the next generation to be very important. However, speakers differed about the reasons for this attributed value. While some speakers considered transmission of Spanish important due to the importance of bilingualism, other speakers considered that Spanish specifically was important to maintain family ties and to foster a connection to the heritage Hispanic and Colombian culture. Despite the almost unanimous high level of importance attributed to the intergenerational transmission of the heritage language, a second major theme that I identified is that speakers consider this task to be too difficult. Speakers highlighted, among others, the higher value of French in Canada, the strong influence of English in the day-to-day life, and the pressure felt by first generation speakers to learn English.

Following the qualitative analysis of questions relating to speakers' attitudes towards linguistically exogamous and endogamous relationships, the primary theme that I identified is that the group of Colombian individuals sampled do not consider language as a primary consideration when finding a romantic partner for themselves or for their children. Both first and second generation speakers agreed that while it may be ideal to find a partner who speaks the same heritage language, other considerations such as cultural affinity are more important. That being said, some second generation participants expressed that finding a partner who spoke the same language was an important, albeit not primary, consideration in order to facilitate communication with their parents and immediate family.

I also conducted a qualitative analysis regarding speakers' knowledge and use of Spanish language-learning opportunities. In this case, speakers were divided in their responses; half of the participants surveyed reported not being aware of many or any Spanish language-learning resources or opportunities, and half reported having some knowledge of resources and opportunities available. That being said, in the case of speakers who reported knowing of available resources or opportunities, the majority confessed to not using them. For the majority of speak-

ers, ultimately, the home represents the primary medium where Spanish is learned, with little to no input from the community outside the home.

Finally, I considered the cultural association to Hispanicity or Colombia of second generation speakers. In this case, the question asked, “Do you consider yourself more Canadian or Colombian, and why?” was particularly difficult to answer confidently for most second generation speakers. The majority of the second generation speakers struggled with the response and talked through their feelings and their reasoning of why they may identify more with one identity or the other before reaching a decision. Ultimately, although it was not an easy answer for most of the speakers, the majority of second generation speakers concluded that they were more Canadian than Colombian. One major consideration in their arrival at this conclusion was the language dimension and specifically the fact that they felt more comfortable in English than in Spanish. However, another major theme that came up was the consideration of others’ perspectives of them. That is, many speakers based their answer regarding their national identity on an external source such as what their parents have said or what a given group might perceive them to be. In the end, although many speakers identified as being more Canadian than Colombian, most speakers also reported a feeling of divided or blended cultural identity.

In this results section, the qualitative analysis revealed a picture of how Spanish speakers in my sample use their Spanish language and how their identity as members of the Hispanic and Colombian community in London, Ontario influences their choices in different aspects of their daily lives ranging from the language used at home, to their social interactions with others, and their sense of cultural and national identity.

In the following section, I show the results of a mixed-methods analysis which incorporates the qualitative results just outlined here with the quantitative results from the previous section (section 3.1).

3.3 Explanatory mixed methods analysis

In this section, I continue the investigation of the expression of overt SPPs by individual speakers, while considering both the quantitative and qualitative data collected. Although the current investigation adheres to a more embedded mixed-methods design where, as explained by Manzoor (2016), qualitative data is collected simultaneously as a secondary component of the principal quantitative research design in order to inform the research about the larger context of the investigation or address secondary research questions, this section adopts aspects of an explanatory mixed-methods approach as well. This represents an explanatory mixed-methods approach in that I use qualitative data in order to find explanations to patterns observed through the primary quantitative analysis. Therefore, in the section that follows, I consider individual speaker variation, and analyse reasons why certain individuals within each of the two generations seemingly pattern differently than other members of the same generation.

3.3.1 Mixed methods approach to individual speaker analysis

In this section, I present a mixed-methods analysis by focusing on individual speakers in order to identify connections between the qualitative and quantitative data analysed in this study. In the table below, I present the results for speaker data from the Rbrul multivariate analysis conducted including both generations of Colombian speakers in London, Ontario (3.12).

The data presented show which speakers seem to favour the use of overt SPPs and which seem to disfavour it. Note that when we consider the eight speakers with factor weight values below 0.5, which seem to disfavour the use of overt SPPs (in comparison to other speakers' sampled use in the data), all but two of these speakers are from the first generation group. These second generation speakers are 2GM4 and 2GF2.

In addition, following qualitative analysis of the data, results revealed that when we consider second generation speakers, most of the individuals included in our sample identified themselves as having a dual or divided identity, and the majority ultimately considered them-

	Overt SPP	N	Proportion	Factor Weight
Speaker				
2GM1	55	113	0.487	0.67
2GF1	49	115	0.426	0.63
2GM2	38	116	0.328	0.57
1GF3	39	131	0.298	0.57
2GM3	37	108	0.343	0.56
1GM4	30	105	0.286	0.55
1GM3	45	138	0.326	0.53
2GM5	32	103	0.311	0.52
2GF3	26	100	0.260	0.52
2GF4	27	102	0.265	0.52
2GF5	34	108	0.315	0.51
1GF4	30	109	0.275	0.51
1GF2	39	136	0.287	0.48
1GM5	22	102	0.216	0.47
1GM1	33	157	0.210	0.43
2GF2	30	150	0.200	0.42
1GF5	23	120	0.192	0.42
2GM4	18	112	0.161	0.38
1GM2	16	105	0.152	0.38
1GF1	14	136	0.103	0.34
Range				33

Table 3.12: Rbrul results for individual speaker data ($N = 2366$). Data presented in descending order according to FW values for overall overt SPP usage.

selves to be more Canadian than Colombian. As previously discussed, this affinity to being closer to a Canadian or Colombian identity seems to be composed of multiple factors. It is interesting to note however, that the two speakers who identified as feeling more Colombian than Canadian (2GF2 and 2GM4) also showed the lowest rates of overt expression of subject personal pronouns, and patterned much closer to the first generation average of 23.43% than the second generation average of 30.6%, and they are the same two speakers identified above, (Table. 3.12), as the only two second generation speakers disavouring the use of overt SPPs in variable contexts. Specifically, one of these second generation speakers, a female (2GF2), used overt SPPs in 20% of all variable contexts, and the other second generation speaker, a male (2GM4), used overt SPPs in only 16% of variable contexts. It is also important to note that both of these speakers were either born in Canada, or arrived in Canada before the age of five,

representing two of the speakers with the least direct experience with the Hispanic country of origin and monolingual Spanish speakers.

Speaker	Home	Social-local	Social-int'l	Professional	Cumulative score	FW
2GM2	9/12	1/8	4/8	0/4	14/32 (0.44)*	0.57
2GF1	6/8	1/8	4/4	0/4	11/24 (0.46)*	0.63
2GM1	8/12	2/8	4/4	0/4	12/28 (0.50)*	0.67
2GM5	8/12	2/8	4/4	0/4	14/28 (0.58)*	0.52
2GF3	9/12	2/8	8/8	1/4	20/32 (0.63)	0.52
2GM4	10/12	2/8	8/8	0/4	20/32 (0.63)	0.38
2GF4	9/12	5/8	5/8	3/4	22/32 (0.69)	0.52
2GM3	11/12	5/8	4/4	1/4	21/28 (0.75)	0.56
2GF2	9/12	7/8	8/8	1/4	25/32 (0.78)	0.42
1GF1	16/24	7/8	8/8	0/4	31/44 (0.71)*	0.34
1GM2	19/20	6/8	4/4	0/4	29/36 (0.73)*	0.38
1GF5	20/20	4/8	8/8	0/4	32/40 (0.80)*	0.42
1GM3	24/24	4/8	8/8	0/4	36/44 (0.82)*	0.53
1GF4	24/24	5/8	6/8	1/4	36/44 (0.82)*	0.51
1GM4	19/20	8/8	6/8	1/4	34/40 (0.85)	0.55
1GF2	18/20	8/8	8/8	0/4	34/40 (0.85)	0.48
1GM1	16/20	7/8	8/8	3/4	34/40 (0.85)	0.43
1GF3	20/20	7/8	8/8	0/4	35/40 (0.88)	0.57
1GM5	20/20	7/8	8/8	2/4	37/40 (0.93)	0.47

Table 3.13: Quantified language use scores according to frequency of use of Spanish across four different domains (home/family, social-local, social-international, and work) as reported in language questionnaire for first generation ($N = 10$) and second generation ($N = 9$) speakers, and factor weight values for their use of overt SPPs. Asterisks (*) denote speakers in high intensity of contact for each generation group (see section 2.3.3).

However, it is also interesting to note that according to the quantified analysis of the language use questionnaire responses, these two speakers were also among the speakers in their generation to use Spanish most frequently and who therefore were considered among the speakers with less intense situation of language contact within their generation group. In fact, 2GF2 was the first generation speaker with most frequent self-reported use of Spanish across different domains (Table 3.13)

In addition, while these speakers (2GF2 and 2GM4) expressed a desire to maintain Spanish for improving work or other opportunities, they both quoted a desire to remain connected with family and the heritage culture as a motivation for the maintenance of Spanish.

Speaker 2GF2 was also one of only three speakers who expressed that finding a partner who was at least able to communicate in Spanish was an important consideration in order to facilitate communication with Hispanic family members (126).

(126) *Creo que sí. Yo creo que es importante que hacen el esfuerzo para comunicarse con la familia. Nosotros conocemos gente que se casaron con un canadiense, y el canadiense no habla nada de español y no puede comunicarse con la familia de la señora y es difícil. (2GF2).*

“[...] I think yes. I think that it is important that they make the effort to communicate with the family. We know people that have married a Canadian, and the Canadian doesn’t speak any Spanish and can’t communicate with the family of the wife and it’s difficult.”

There are two other speakers (2GF3 and 2GF4), whose use of overt subject pronouns patterns closely with that of first-generation speakers with variable expression frequencies of 25% and 26% each respectively. It is important to note that these two speakers are related and both of these speakers reported having been immersed in multiple activities with Hispanic groups in London, Ontario, including formal education courses and frequent participation in Hispanic religious groups (127-128).

(127) *Nosotros teníamos que ir a reuniones cristianas dos veces a la semana y sobretodo en español todo. Entonces yo tenía que leer bien el español, todas las escrituras que nos hacían hacer cada semana y comentábamos también en español. (2GF4)*

“[...] We had to go to Christian meetings twice a week and mostly in Spanish everything. So I had to read well in Spanish, all the scriptures we had to read, and we would comment as well in Spanish.”

(128) *También tomamos, tomamos cursos, o sea clases de español en high school. Yo hice*

todos los tres años de español que me ayudó a, a conocer un poquito más de la gramática. (2GF3)

“We also took, took courses, like Spanish classes in high school. I did all three years of Spanish which helped me to, to know a little more about grammar.”

In the case of these speakers, although they both reported identifying as more Canadian than Colombian, the increased quantity and variety of Spanish language input likely contributed to their language maintenance and may have influenced their use of subject pronouns as they engaged with more first-generation speakers and other Hispanic individuals. This is further confirmed by the fact that according to the quantitative measure of their language use, both of these speakers (2GF3 and 2GF4) were among the five speakers in their generation considered to experience a lower level of language contact intensity as assessed through their self-reported more frequent use of Spanish in different contexts (Table 3.13).

It is also important to note that in the case of these speakers, (2GF3 and 2GF4), the factor of family and culture is also an important consideration. They both reported a desire to communicate with family as a motivator for maintaining their use of Spanish, and they expressed, similarly to the 2GF2 speaker, that having a Hispanic or Latino partner was an important consideration for them in order to facilitate communication with their family.

(129) *Sí, sí, sí. Porque, por el mismo sentido, para que [mis hijos en un futuro] puedan comunicarse bien con sus abuelos, ¿no? (2GF4)*

“Yes, yes, yes. Because, in the same sense, so that [my children in the future] can communicate well with their grandparents, no?”

(130) *Pues mi cuñada es canadiense, entonces yo veo lo difícil que es para mis papás y, y a veces para mi cuñada. Que mi cuñada es excelente [...] Eh, pero la barrera existe y es difícil para las dos personas. Entonces, ahora, eh sí estoy depronto pensando*

que de pronto mi próxima pareja sea hispano. (2GF3)

“Well my sister-in-law is Canadian, so I see how difficult it is for my parents and, and sometimes for my sister-in-law. That my sister-in-law is excellent [...] Eh, but the barrier exists and it’s difficult for both people. So, now, eh, yes I am maybe thinking that maybe my next partner be Hispanic.”

Although these two speakers, (2GF3 and 2GF4), have equal factor weights which seem to favour the use of overt SPPs in variable contexts, note that the value of 0.52 indicates a weak favouring tendency (Table 3.12).

In contrast, the two speakers within the second generation group who showed the highest frequencies of use of overt SPPs in variable contexts (2GF1 and 2GM1) with frequencies of 42% and 49% respectively and who also favour the use of overt SPPs most strongly according to Rbrul results with factor weights of 0.63 and 0.67 respectively (Table 3.12), also reported a higher use of English when speaking with friends and with having fewer contexts in which they use Spanish.

(131) *Honestamente yo creo que la única oportunidad es cuando uno está pues en la casa porque pues acá no se habla nada más sino el español. (2GM1)*

“Honestly, I think that the only opportunity is when one is, well, at home because well here we don’t speak anything else but Spanish.”

This more frequent use of English is reflected as well in their self-reported frequency of use of Spanish. Speakers 2GF1 and 2GM1 are among the three speakers with the least frequent use of Spanish, and among the four speakers in the generation who are considered to be in a situation of high intensity contact based on their regular use of Spanish (Table 3.13).

Interestingly, when asked about their feelings towards their Colombian or Canadian identity, both speakers also reported not feeling comfortable within Colombian social groups, or having been considered “more Canadian” by other Hispanic members of the community.

- (132) *Con los canadienses no me siento cien por ciento uno de ellos, pero tampoco me siento cien por ciento con los latinos que vivieron y crecieron allá, y que tienen toda su familia allá, y tienen toda la cultura asociada a allá.* (2GF1)

“[...] With the Canadians I don’t feel a hundred percent as one of them, but I don’t feel a hundred percent with the Latinos that lived and grew up there, and that have all their family there, and have all the culture associated with over there.”

- (133) *Para una persona colombiana, yo creo que nosotros seríamos canadienses.* (2GM1)

“[...] To a Colombian person, I think that we would be Canadian.”

These results point to a relationship between frequency of use of Spanish, family connection, identification with the heritage culture, and the expression of subject personal pronouns in Spanish. Notably, I observed that within my sample, speakers who saw family as a primary motivation to use and learn Spanish and who more closely identified as Colombian, or who felt as part of that cultural group, showed a frequency of use of overt SPPs which more closely resembled that of first generation speakers. It is possible that these speakers feel more comfortable speaking Spanish and look for more opportunities to continue to learn and practice Spanish. In fact, in addition to this factor, speakers who used Spanish more frequently were also more likely to show frequencies of use of overt SPPs that more closely resembled first generation speakers.

Note as well that of the 12 speakers that seem to favour the use of overt SPPs only four are from the first generation group, with all other speakers being from the second generation group (Table 3.12). These speakers are 1GF3, 1GM4, 1GM3, and 1GF4. In the case of these specific speakers, no clear connecting factor seems to be present. For instance, while speaker 1GM4 had been in Canada for 20 years at the time of the interview, the second longest LOR in the CoSLO corpus, speaker 1GF3 had been in Canada for nine years, the shortest length of residence period in the sample at the time of the interview, and speakers 1GM3 and 1GF4

had been in Canada for 15 years. One pattern that is notable, however, is that all speakers in the first generation who favour the overt expression of Spanish SPPs, are considered to have intermediate or advanced levels of English language proficiency according to their self-reported scores. In addition, the speakers with the highest rates of use of overt SPPs, 1GM3 (32.6%) and 1GF3 (29.8%), both reported feeling a high level of pressure to learn English upon their arrival either due to experiences of discrimination, or due to pressure in regards to their professional success or their ability to be a model for their children (134-135).

(134) *Bueno el idioma es lo más difícil. La otra cosa es conseguir trabajo. Que la discriminación... A pesar de que yo tengo pinta de, de gringo, pues de Europeo o lo que sea, la discriminación... apenas la persona nota que uno tiene un acento empieza a discriminar.*

“Well the language is the most difficult. The other thing is finding work. That the discrimination... Even though I look like, like gringo, well European or whatever it is, the discrimination... As soon as the person notices that you have an accent, [they] begin to discriminate.”

(135) *Si queríamos hacer algo que nos brindara como una satisfacción personal y además de eso también poderle hablar con cierta autoridad a los hijos en términos de lo que deberían hacer y ser un modelo para ellos, pues sentimos que teníamos que, eh, aprender el idioma lo más rápido posible.*

“If we wanted to do something that gave us like some personal satisfaction and in addition to that to also be able to speak to the children with certain authority in terms of what they should do and be a model for them, well we felt that we had to, eh, learn the language as fast as possible.”

In addition, speaker 1GM4, who had been in Canada for the longest period of time, reported

during the qualitative analysis that he uses Spanish with possibly a higher frequency than he reported in the questionnaire. While in the questionnaire he reported that his children used Spanish between 60-89% of the time when addressing him, during the interview, the speaker reported that his children address him in English about 70% of the time (136).

(136) *Yo la mayoría, el noventa por ciento les hablo en español, pero ellos el setenta por ciento me lo hablan en inglés.*

“I for the most part, ninety percent I speak to them in Spanish, but they about seventy percent speak to me in English.”

This difference in score would translate to a change in frequency of use of Spanish from 0.85 to 0.82, which would in turn categorize this speaker as being in a situation of high language contact intensity. This would then show that three of the four speakers in the first generation who favour the use of overt SPPs, (1GF3, 1GM3, and 1GM4), are in a situation of high language contact intensity and that two of them have experienced high pressure to learn and adapt to an English-language environment.

Therefore, although there does not seem to be a clear pattern explaining the wide variation among first generation speakers' in regards to their use of overt SPPs (10.3% - 32.6%), it is possible that an effect of language dominance related to a higher level of pressure to learn English faster is at play in these situations. This, however, cannot be confirmed in the current study and would need to be assessed in future studies which directly address language dominance as a factor.

It is important to highlight as well that when we take individual speakers' factor weights into consideration, only two speakers show factor weights over the 0.6 level, and these speakers, 2GM1 and 2GF1, are both second generation speakers (recall as well, however, that 2GM1 was identified in section 2 as a possible outlier). In addition, these speakers are also the two speakers with the lowest reported frequency of use of Spanish (table 3.13), further indicating an effect of language contact intensity as measured through frequency of use of the minority

vs. the majority language.

The patterns apparent through analysis of the Rbrul results for the speaker factor (Table 3.12) suggest there may be a tendency for second generation speakers to use overt SPPs more frequently than first generation speakers, with most speakers who disfavour the use of overt SPPs (75%) being from the first generation group, and most speakers who favour the use of overt SPPs (66%) being from the second generation group. However, other factors influencing this pattern are suggested to be extra-linguistic including language contact intensity as determined through frequency of use of the minority language, as well as factors tied to language attitudes and sense of national identity.

3.3.2 Mixed-methods analysis - Summary

In this section, I conducted an explanatory mixed-methods analysis by informing the patterns and trends observed in the quantitative results with the themes identified in the qualitative analysis. This analysis allowed me to identify the important link between identity, language use and the variable expression of Spanish SPPs. Specifically, the results showed that while there is a trend where second generation speakers use overt SPPs with a higher frequency than first generation speakers, there are a few second generation speakers who in fact pattern closer to the average use of overt SPPs by first generation speakers. Among these speakers, the ones with the lowest rate of use of overt SPPs, (2GF2 and 2GM4), were also the only two speakers in their generation group to identify as more Colombian than Canadian (section 3.2), and they were the two speakers who reported the highest use of Spanish in their day-to-day life. This result was observed despite the fact that these speakers were also among the speakers in the corpus to have had the least experience in a Spanish monolingual country, with one of the speakers being born in Canada.

In addition to these speakers, the remaining second-generation speakers who also patterned closer to the first generation average, disfavoured the use of overt SPPs, were two speakers, (2GF3, 2GF4), who also reported some of the highest levels of Spanish use in their daily lives

and who reported family being an important consideration for learning and maintaining the use of Spanish.

In contrast, the speakers in the second generation group who reported the highest use of English in their daily lives were also the speakers in the sample who showed the highest frequency of use of overt SPPs. These speakers (2GM1, 2GF1) also more closely identified with a Canadian identity and they both reported instances where they had been labeled as “more Canadian than Colombian” by others or where they felt uncomfortable when interacting only with other Colombian individuals.

In conclusion, the results of this analysis allowed us to identify a link between heritage speakers’ variable use of Spanish SPPs and their cultural identity, frequency of use of Spanish, and sense of *familismo*. This is an important finding which could not have been found without employing the mixed-methods approach which was adopted in this study.

Chapter 4

Discussion

In this chapter, I discuss the results from the current investigation and consider how these results contribute to our understanding of the use of Spanish Subject Personal Pronouns in bilingual immigrant contexts. The chapter is organized to mirror the sections of the preceding results chapter with results pertaining to the quantitative investigation presented first, followed by a discussion of the qualitative investigation, and ending with a discussion of the explanatory mixed-methods analysis regarding individual speakers. Recall that while the quantitative investigation conducted sought to determine the social and linguistic factors that condition the variable expression of Spanish Subject Personal Pronouns (SPPs), the qualitative investigation sought to answer questions regarding the attitudes of Colombian Spanish speakers in London, Ontario and how these aligned with known factors that contribute to the maintenance of minority languages, in order to discuss the likelihood that Spanish will be maintained in future generations.

4.1 Variable subject pronoun expression in Løndombia across two generations

The primary goal of this dissertation is to determine the social and linguistic factors that condition the variable expression of Spanish SPPs in the Spanish of Colombian speakers, from the region of Cundinamarca, living in London, Ontario. The variable expression of Spanish SPPs is a well studied characteristic of Spanish which has been investigated by researchers in both monolingual (Cameron 1993; Travis 2007; Orozco 2015) and bilingual contexts (Silva-Corvalán 1994b; Flores-Ferrán 2004; Hurtado 2005; Montrul 2011; Otheguy and Zentella 2012). That being said, most of the studies concerning Spanish-English bilinguals in North America have to date been focused on communities of the United States. Through this investigation, I aim to expand the discussion on the variable use of Spanish SPPs by including new data from the use of Spanish in contact with English from a Canadian context. Specifically, I collected language data from two generations of Colombian Spanish speakers, building the CoSLO corpus of Spanish, and then proceeded to determine, via a quantitative analysis, the linguistic and social factors that condition the use of Spanish SPPs in this population and each of the two generations. This investigation thus, not only provides information on this characteristic of Spanish within a situation of Spanish-English language contact in a Canadian context, but also adds to the discussion regarding how the use of Spanish SPPs changes across generations in a bilingual community where Spanish is a minority immigrant language.

My investigation revealed that the factor which most strongly conditions the variable expression of Spanish SPPs for the population of Colombian Spanish speakers in London Ontario was pronoun type (grammatical person and number). This was the first factor with the greatest impact for first and second generation groups alike. Further, the only other factor that was significant in conditioning the variation for both first and second generation speakers was the change of referent (switch reference) in the discourse. These results align well with investigations regarding the variable expression of Spanish SPPs in both monolingual and bilingual

communities, which have shown that although different varieties of Spanish vary on the frequency of use of overt vs. null SPPs, most varieties show that the variable expression of Spanish SPPs is conditioned by a similar hierarchy of linguistic factors with grammatical person and change of referent being among the most consistently significant conditioning factors from one variety to the next (Flores-Ferrán 2004; Montrul 2004; Otheguy and Zentella 2012; Orozco 2015).

The pronoun type and change of referent were the only two linguistic factors that conditioned the use of overt SPPs for both first and second generation speakers, and the only two linguistic factors to significantly condition variable SPP expression for the first generation group. However, in addition to pronoun type and change of referent, the variable expression of Spanish SPPs in second generation speakers was found to be conditioned significantly as well by two more linguistic factors (verbal mood and clause polarity), and one extra-linguistic factor (interview mode).

It is important to highlight as well that this study did not find other social factors such as gender or bilingual proficiency to be significant in conditioning this variable in the Spanish of Colombian speakers in London Ontario. Previous studies have been somewhat divided regarding the role of social factors in conditioning the variable expression of Spanish SPPs. For instance, according to Otheguy and Zentella (2012), when considering the Hispanic population of New York, they found that gender and socioeconomic level were both significant factors conditioning variable expression of Spanish SPPs. Specifically, the authors found that females and individuals from higher socioeconomic levels showed higher rates of use of overt SPPs as opposed to males and individuals from lower socioeconomic status in New York City. These results, explain Otheguy and Zentella (2012), are somewhat counterintuitive because higher rates of overt SPP usage are associated with a lower level of prestige, and therefore it would be expected that males and speakers from lower socioeconomic levels would be more likely to show the increase in overt SPPs. Nevertheless, one explanation they propose is that within the context of New York, speakers of higher socioeconomic status are undergoing a faster

process of attrition since as they incorporate themselves into the community they also tend to use Spanish less frequently. The process of attrition then leads to faster erosion of discourse-pragmatic features in these speakers and to higher rates of use of overt SPPs consequently. This effect of social factors on the variable expression of Spanish SPPs, however, was not found in other similar studies which concluded that the variable use of SPPs was not conditioned by social factors (Bentivoglio 1987; Cameron 1993). While my study's results do find some evidence of extra-linguistic factors influencing the variable expression of Spanish SPPs, such as the interview modality, social factors such as gender were not found to be significant. It is my opinion that when it comes to social factors' influence on this variable feature of Spanish, each community might be different. These, may be dependent on those factors which are specific to each community such as gender roles (i.e., are women staying home more than men?), the sense of minority language community leading to increased opportunities to use the language in more contexts, prestige norms towards Spanish and the specific Spanish varieties in the community, among others.

4.1.1 Grammatical person and number

Similarly to previous investigations which have focused on Spanish-English bilingual communities, results of my investigation showed that the Colombian speakers in the current study favoured the use of overt SPPs with singular pronoun forms (Silva-Corvalán 1994b; Otheguy and Zentella 2012; Orozco 2015). Specifically, when considering data for both generation groups together (section 3.1.1.1), Colombian speakers in London used overt SPPs most frequently with third person singular pronouns followed by first and second person singular pronouns ($3SG > 1SG > 2SG$) (Table 3.1). In contrast, plural constructions were shown to disfavour the use of overt SPPs. This aligns with results of other investigations considering Spanish-English bilingual speakers in a situation of language contact. For instance, Flores-Ferrán (2004) found that for a population of Puerto-Rican speakers in New York, first, second, and third person singular constructions were significantly associated with a higher use of overt

SPPs. These results evidence the general tendency for singular pronouns to favour overt SPPs (Flores-Ferrán 2004; Orozco 2015; Otheguy et al. 2007).

However, the order in which singular pronouns favour the use of overt SPPs according to grammatical person can differ between different varieties. For instance, results of my investigation differ from results presented in Montrul's (2004) investigation where overt SPPs were favoured by second person singular pronouns followed by first, and finally third person singular pronouns. Differences in the order in which specific subject pronouns favour overt vs. null SPPs is not uncommon when we investigate different communities of Spanish speakers and different varieties of Spanish. In fact, Otheguy et al. (2007) showed differences in terms of the pronoun types with the highest frequencies of overt SPPs within the community of New York City according to both region of origin (Caribbean vs. Latin American Mainland) and length of residence in New York (newcomer vs. New York born and raised). Specifically, when considering Caribbean Spanish-speaking newcomers in New York vs. Caribbean heritage speakers born and raised in New York (NYBR), the authors noted a change between generations in the hierarchical order of pronouns favouring the use of overt SPPs. That is, Caribbean NYBR speakers used overt SPPs most frequently with third person singular pronouns, followed by first person singular pronouns, while Caribbean newcomer speakers showed a flipped tendency using overt SPPs most frequently with first person singular pronouns followed by third person singular pronouns.

The result presented by Otheguy et al. (2007) thus points to possible differences between generations within a community of speakers when in a situation of language contact in regards to the order of pronoun types with which overt SPPs are used most frequently. The results of my investigation mirror these results as well as they show an inversion between 1SG and 3SG pronouns in relation to the use of overt SPPs between generations. In my investigation, the results of the separate analyses of the first (section 3.1.2.1) and second generation (section 3.1.3.1) speakers showed that first generation speakers favour the use of overt SPPs with second person singular constructions, followed by first person singular and third person sin-

gular constructions ($2SG > 1SG > 3SG$), while second generation speakers favour the use of overt SPPs with third person singular pronouns, followed by first person singular pronouns ($3SG > 1SG$).

Otheguy et al. (2007) suggest that the change observed in their data between generations is a sign of dialect leveling due to contact with Spanish speakers from Latin American mainland varieties such as those from Mexico, Colombia and Ecuador. This was suggested due to the fact that the Caribbean NYBR speakers' use of overt SPPs according to the different pronouns mirrored the frequencies of mainland speakers in their sample. In other words, both mainland speakers and Caribbean NYBR speakers used overt SPPs most frequently with third person singular pronouns, followed by first person singular pronouns, while Caribbean newcomer speakers showed a flipped tendency using overt SPPs most frequently with first person singular pronouns followed by third person singular pronouns. In addition, there was no change between generations for mainland speakers in New York City, which further suggests that this change was led by dialect contact. However, since in London Ontario the spoken varieties of Spanish are mostly mainland Latin American varieties (Table 2.1), with only about 5% of the Spanish-speaking population reporting being from Caribbean Hispanic nations or territories, there is no anticipated dialect leveling effect in this aspect of language. Therefore the change between generations in the case of the current study must be related to different factors.

One possible explanation is related to ambiguity of the verb forms associated with the pronouns in question. That is, 3SG pronouns have the highest level of ambiguity across all other pronouns since the same 3SG verbal forms can be used to refer to the pronouns *él* 'he', *ella* 'she', and formal 2SG *usted* 'you'. It is possible that second generation speakers use overt SPPs with higher frequency with 3SG in order to decrease ambiguity. This is further supported by the fact that verbal mood was higher in ranking as a conditioning factor for the variable expression of Spanish SPPs among second generation speakers than among first generation speakers (3.3) since verbal mood is another factor which, as described in section 2.5.2, can also introduce further ambiguity depending on the verbal mood employed in the construction.

Another possible explanation is related to the possibility of incomplete acquisition of features at the discourse-syntax interface in these second generation speakers. Shin and Smith Cairns (2009) examined the acquisition of sensitivity to continuity of referent in monolingual Spanish-speaking children in regards to their use of null vs. overt SPPs. The authors explain that a much misunderstood aspect of acquisition of sensitivity to continuity is the duality of this factor. They emphasize that sensitivity to referent continuity relates not only to the ability to reduce ambiguity by including overt SPPs when there is a change in referent, but also to the ability to reduce redundancy by using null SPPs when there is continuity of the referent. These two distinct aspects of continuity, develop at different stages. While the ability to reduce ambiguity by employing an overt SPP develops at around the age of nine, the ability to reduce redundancy does not develop to adult-like levels until the age of fourteen. Therefore, children between the ages of nine and fourteen tend to overuse overt SPPs in contexts where there is continuity of the referent. They found evidence of this pattern in their study where children between the ages of 7 and 14 showed higher rates of use of overt SPPs in same referent contexts than their adult counterparts, especially in the case of 3SG pronouns which, the authors explain, are associated with the most ambiguous forms. In the case of my study, therefore, it is possible that second-generation speakers, who all arrived in Canada before the age of nine, may not have fully developed the sensitivity to the continuity of referent factor as it relates to the variable expression of SPPs, particularly the later developed sensitivity to redundancy in discourse and that this in turn results in a higher use of overt SPPs, especially in the case of 3SG pronouns. This, however, should be considered in future studies by including a comparison of subject pronoun expression between grammatical person and referent continuity.

It is now important to address the variable expression of SPPs by second generation speakers in the current investigation in regards to second person singular pronouns. While it's not clear why second person singular constructions did not significantly favour the use of overt SPPs in the second generation group, one possibility may lie with the differences between specific and non-specific second person singular pronouns. Previous investigations have shown

that the frequency of use of overt SPPs can vary according the specificity of second person singular pronouns (Cameron 1993; Otheguy et al. 2007; Flores-Ferrán 2004). For instance, in Flores-Ferrán (2004), the author showed that while frequencies of use of overt SPPs were high with both [+ specific] and [- specific] second personal singular pronouns, which surpassed all frequencies of use of overt SPPs for plural constructions, the rate of use of overt SPPs differed according to specificity. She explained that speakers in her sample were more likely to use an overt SPP with [+ specific] than [- specific] second person singular pronouns. Within my data, I was not able to tease apart this effect for individual generations due to the token count of second person pronouns being too low¹. However, an analysis of [+ specific] vs. [- specific] second person singular pronouns while considering the aggregate data for both generation groups showed the opposite pattern (section 3.1.1.1). Speakers within my sample tended to use overt SPPs more when using second person singular pronouns with a [- specific] reading. This trend seems to align more with the results presented by Cameron (1993) for the Madrid speakers, while going contrary to the patterns of the speakers from San Juan. The results of this investigation thus provide further evidence that different varieties of Spanish can vary in regards to whether the specificity of the second person singular pronouns favours or disfavors the expression of Spanish SPPs. It may also serve to show lack of influence from Caribbean Spanish varieties in Canada, as well as a contrast with Spanish speaking populations in the United States.

Another explanation for why second person singular pronouns seem to disfavour the use of overt SPPs in the data for the second generation group may simply be the low token count for these constructions. In total, there were only 32 tokens for second person singular constructions extracted for the second generation group, and only three of these tokens were overt instances. This low token count may have ultimately influenced the analysis resulting in a frequency of use of overt SPPs for second person singular pronouns that is not entirely representative of the generation's use. This will need to be confirmed in future investigations, however, with

¹When considering the second generation of speakers, there were only 29 second person singular constructions identified and included in the analysis, where only three tokens were instances of overt SPPs.

a higher token count for second person singular constructions which allows for a more representative sample of the population in regards to use of overt SPPs with this subject pronoun, and which allows for a more detailed analysis according to the [+/- specific] character of these tokens. Note that the low token count for second person pronouns in the current investigation, although not ideal, is common when the data is retrieved from sociolinguistic interviews since the participants are less likely to ask questions to or include the interviewer within their statements. In fact, it's interesting to note that at the end of one of the interviews that I conducted, one participant expressed that he had wanted to ask me a question about my opinion on one of the topics during the interview but felt he should wait until after the interview to ask.

Nevertheless, results in this investigation agree with general tendencies observed in previous studies which show that the use of overt Spanish SPPs is favoured with singular personal pronouns, particularly with third and first person singular pronouns, and disfavoured with plural personal pronouns. This can be due to verb forms which accompany singular pronouns showing an increased level of ambiguity. For instance, in the case of the verb *cantar* 'to sing', the first person singular form, in the indicative present, would be [*yo*] *canto* '[I] sing' which is much closer in form to the third person singular construction [*él/ella*] *canta* '[he/she] sings' than to any of the plural forms such as the first person plural [*nosotros*] *cantamos* '[we] sing', or the third person plural [*ellos/ellas*] *cantan* '[they] sing'. This idea also connects to the concept of "morphological richness" by which it is hypothesized that speakers of null subject languages, such as Spanish, are more likely to use null SPPs when there is sufficient information in the clause to recover the subject of the clause, which is in turn facilitated by a more saliently different verbal form which allows speakers to use morphological cues in the verb to determine the subject of the clause². This is also related to the functional compensation hypothesis of (Hochberg 1986), which explains that increased ambiguity of certain verbal forms can lead to speakers being more likely to include an overt pronoun in order to clarify the state-

²Recall, however, that this idea of "morphological richness" is contested in the literature since it does not apply well when considering other null subject languages which do not have a rich morphological verbal paradigm (Poplack 1980; Cameron 1993; Huang 1994; Roberts and Holmberg 2010).

ment (Hochberg 1986; Cameron 1993). Following this idea, the more saliently distinguishable verbal forms are, the less likely that there will be the need for disambiguation through the use of an overt SPP. Therefore, since verbal forms for first, second, and third person singular tend to not be as saliently different as plural forms when considering regular verbs then it follows that according to the functional compensation hypothesis, constructions accompanying singular pronouns will favour overt SPPs to reduce ambiguity and plural forms will disfavour them.

The results of my investigation seem to support this idea since it's in particular third and first person singular forms (the forms shown above to introduce the most ambiguity to a statement), that favour the use of overt SPPs most significantly in the population of London Ontario. This is true when the data for both generations is considered together, where although all singular pronouns favour the use of overt SPPs, first and third person singular pronouns show the highest frequencies with 28.7% and 25.3% frequency of use of overt SPP use correspondingly, as well as for the second generation group data where third and first person singular pronouns show the highest rates with 44.3% and 33.5% frequency of use of overt SPPs correspondingly. When first generation speaker data is considered separately, the highest frequency of use of overt SPPs is for second person singular pronouns, followed by first and third person singular pronouns. However, it is important to note again that the token count for second person singular constructions is also low within this generation group, with only 50 tokens, and there is a large margin of error ($SEM = 6.69\%$) in how well the frequency estimate based on the sample collected represents the true population average. In fact, according to the calculated error margins, the averages for first, second and third person singular constructions are not significantly different from each other when considering this generation's speaker data.

4.1.2 Change of referent

The second most significant linguistic factor that conditions the variable expression of Spanish SPPs for the Colombian speakers in my data is the change of referent in the dis-

course. Specifically, results showed that overt expression of Spanish subject personal pronouns is favoured when there is a change in referent in the discourse, or at the beginning of an utterance. Overt SPPs are disfavoured, however, when the referent in the considered clause was the same as in the previous clause. This pattern was observed for both generations equally when they were considered together (section 3.1.1.3) and separately (sections 3.1.2.2 and 3.1.3.4).

This result mirrors findings in previous research investigating the variable expression of Spanish SPPs in both monolingual and bilingual populations. Studies considering populations from Mexico, San Juan, New York, and Colombia, among others, all have found that a change in referent is a significant conditioning factor in the variable expression of Spanish SPPs, such that a new or changed referent in the discourse favours the overt expression of SPPs (Silva-Corvalán 1994a; Montrul 2004; Orozco 2004; Flores-Ferrán 2004; Otheguy et al. 2007; Orozco and Guy 2008). In fact, this result has been shown to be among the most stable conditioning factors of the variable expression of Spanish SPPs regardless of the linguistic varieties considered. Although variation is observed from one Spanish variety to the next in regards to the general rate of use of overt Spanish SPPs, there is some consistency in regards to the factors that condition the use of overt vs null SPPs with grammatical person and switch reference being among the most common ones to be significant from one community to the next (Flores-Ferrán 2004).

The importance of the switch reference factor in conditioning the expression of SPPs may also be tied to a need by speakers to reduce ambiguity (Montrul 2004; Otheguy and Zentella 2012; Orozco 2015). Since a null SPP is more likely to be used when information about the subject of the clause can be easily recovered from the context, one way in which speakers can reduce the ambiguity of their statements is through the use of overt SPPs when a new referent is introduced or changed. That is, a person may be more likely to use a null subject pronoun if the referent remains unchanged in the discourse since mentioning it again is unnecessary or redundant. However, when the referent changes, there is a higher need to use a pronoun overtly, in order to signal and clarify in the conversation who the referent is and minimize

possible ambiguity.

In a study by Montrul (2004), in which she considered both monolingual Spanish speakers and heritage English-Spanish bilingual speakers living in the United States, heritage speakers showed higher uses of overt SPPs as compared to monolingual Spanish speakers. Montrul (2004) explains that these bilingual speakers were more likely to use overt SPPs in contexts where they were redundant. In addition, she also documents errors by these bilingual speakers where they failed to use overt SPPs when there was a change in referent and an overt SPP would have therefore been preferred. She suggests that this is likely due to L1 attrition through a process where an erosion of [+ interpretable] discourse-pragmatic features, but not [- interpretable] syntactic features occurs. Although in the current study I do not consider the specific errors committed by speakers, I do see differences between the two generations which could be signalling that this erosion of discourse-pragmatic features reported by Montrul (2004) may be taking place among the second generation speakers in my data. Notably, second generation speakers used overt SPPs more frequently in cases where there was no change in referent (26.1%) than first generation speakers (17.4%). This difference between the two generations may indicate an erosion of discourse pragmatic features in the grammar of second generation speakers leading to the introduction of overt SPPs in contexts where a null SPP would be preferred, which leads to a higher frequency of overt expression of pronouns in contexts where there is referent continuity from the previous clause. It's important to note as well that this effect may be more advanced in other second generation speakers in the wider community of London Ontario who have a lower proficiency in Spanish. Unfortunately, however, speakers who considered that their level of Spanish was not sufficient to participate, chose not to participate in the study, and therefore, data in my study likely represents Spanish pronoun usage by some of the more advanced second generation Spanish speakers who may not be undergoing L1 attrition as quickly as other members of the community.

4.1.3 Verbal mood

According to the analysis of the data collected for this investigation, verbal mood was a significant factor conditioning the variable expression of Spanish SPPs only when considering data for both generations together, or when considering only the second generation of speakers. In my data, speakers from the second generation were significantly more likely to express SPPs overtly when the verb in the clause was in the subjunctive or conditional mood, with overt SPP usage in conditional verb forms outnumbering that in subjunctive forms. In addition, although verbal mood was not identified for the first generation speakers as a significant factor in the logistic regression model, the data for this first generation speakers still shows a trend following the same distribution as that seen with second generation speakers. That is, towards the increased use of overt SPPs with subjunctive and conditional phrases. This observed trend, coupled with the fact that the factor is selected as significant when both generations are grouped together suggests that although the factor did not reach significance in the analysis for first generation speaker data, verbal mood may nevertheless still be a contributing factor for this generation.

Verbal mood as a conditioning factor of null vs. overt SPPs, just as when considering pronoun type, can also be considered from the perspective of the need for speakers to reduce ambiguity. In Spanish, although there is a rich morphological verbal system, not all verbal modes and tenses are equally (un)ambiguous. Specifically relevant to this investigation is the fact that conditional and subjunctive conjugations introduce a higher level of ambiguity than indicative forms. For instance, if we return to our example of the verb *cantar* ‘to sing’ and consider the present indicative conjugations as compared to the present subjunctive and present conditional conjugations of the verb, we see that in the present indicative there are fewer forms that are identical to each other morphologically and therefore less chance of ambiguity arising in this verbal mood. This can be seen in the Table (4.1) below (Table re-created from section 1.5.2.1 in the introductory chapter). Notice for instance that in the conditional and subjunctive verbal paradigms there is a three way ambiguity between first person singular, third person

singular, and second person singular *usted* forms. In contrast, in the indicative the first person singular form is distinct morphologically and therefore it shows only a two way ambiguity between third person singular and second person singular *usted* forms.

Subject Pronoun	Indicative	Conditional	Subjunctive
Yo (1SG)	<i>canto</i>	<i>cantaría</i>	<i>cante</i>
Tú (2SG)	<i>cantas</i>	<i>cantarías</i>	<i>cantes</i>
Ústed (2SG)	<i>canta</i>	<i>cantaría</i>	<i>cante</i>
Él/Ella (3SG)	<i>canta</i>	<i>cantaría</i>	<i>cante</i>
Nosotros (1PL)	<i>cantamos</i>	<i>cantaríamos</i>	<i>cantemos</i>
Ustedes (2PL)	<i>cantan</i>	<i>cantarían</i>	<i>canten</i>
Ellos/Ellas (3PL)	<i>cantan</i>	<i>cantarían</i>	<i>canten</i>

Table 4.1: Present tense conjugations of the verb *cantar* ‘to sing’ in the indicative, conditional, and subjunctive moods.

Therefore, the patterns observed in the Table 4.1 explain how using the conditional or subjunctive moods increase the possible ambiguity in discourse which can lead to increased need of overt SPPs. This in turns aligns with my results for both generations of speakers which show a tendency to use a higher frequency of use of overt SPPs with conditional and subjunctive constructions. Further, considering that first person singular pronouns were the most frequently used pronouns in the interviews coded for this investigation, it is then this exact increased level of ambiguity, where first person singular forms are ambiguous in constructions using conditional and subjunctive but not indicative verbal mood, that plays a role when the subjunctive and conditional moods are used by Spanish speakers in London Ontario.

In discussing the verbal mood used by speakers in the current data, it is interesting to see as well the use of subjunctive and conditional moods across both generations. According to Silva-Corvalán (1994b), in the process of language attrition in a situation of language contact such as the one considered in this investigation, verbal moods such as the subjunctive and the conditional are among the first ones to decrease in frequency of use among heritage speakers of the language as simplification of the grammar begins to occur as a result of language attrition. In the population of Colombian Spanish speakers, however, we still see the use of subjunctive and conditional verbal moods among both first and second generation speakers. This may

indicate that at least for those speakers who participated in the study, the process of language attrition is not yet greatly reflected in simplification of the grammatical system.

4.1.4 Clause polarity (negation)

Clause polarity was a factor identified as significantly conditioning the overt expression of Spanish SPPs for speakers of the second generation (Table 3.3, section 3.1.3.5), but not for speakers in the first generation (Table 3.2, section 3.1.2). Spanish SPPs were more frequently expressed overtly when a clause was [- negative] than when it was [+ negative]. This result agrees with some previous studies, such as that by Lastra and Butragueño (2015) who also determined that negative clauses were more likely to be accompanied by null subjects than overt subjects in their analysis of the Spanish of Mexico city.

According to Lastra and Butragueño (2015), the increased likelihood that speakers will omit the overt expression of a subject pronoun in negated clauses has to do with the fact that negated statements tend to be followed by repetitions. That is, a speaker is likely to repeat themselves in regards to the statement they are negating. The consecutive repetitive clauses, leads to higher incidence of null subjects because the referent of the clauses repeated remains unchanged and therefore is easily recoverable from the immediate context, rendering overt SPPs redundant or unnecessary. This explanation, however, doesn't seem to work well with our data since speakers did not tend to repeat negated clauses frequently. Another explanation put forth by Gridstead (1998) who notes that in Spanish whenever negative elements, such as negative quantifiers, are included in a clause, the use of preverbal subjects becomes ungrammatical. The inability to have both a negative quantifier and a preverbal subject personal pronoun could be a feasible explanation for the current data analysed since the instances of post-verbal subject pronouns was extremely low in the data collected. In fact, only about 4% of subjects in tokens with overt SPPs were in post-verbal position when I included both generations' data into the analysis, which highlights a strong preference for pre-verbal subject placement, and which would render the placement of overt subjects more rare in the presence of negative elements,

such as negative quantifiers. Unfortunately, due to the low number of post-verbal tokens, I was not able to conduct a more detailed analysis of the relation between negation and position of subjects in relation to the verb.

A study by Nagy et al. (2015) also found that negative clauses favoured the use of null subjects in second generation speakers of Russian in Toronto. However, they did not find the same pattern in first generation speakers of Russian. This, they explain, has been suggested to be due to a language contact with English since according to Harvie (1998) null subject use in negative expressions such as “don’t know” are becoming more prevalent in English. However, Nagy et al. (2015) conclude that there is no evidence of language contact effects in their data and the relationship between the observed intergenerational clause polarity changes and contact with English remains unconfirmed.

It is important to note as well that although the factor of polarity was identified as a conditioning factor through our logistic regression analysis, an analysis of the difference between the two conditions found that it was non-significant. That is, the use of overt SPPs with [+negative] and [-negative] clauses was found to be statistically the same according to the X^2 analysis of the two groups. This seemingly contradictory result suggests a weak association between this factor and the variable expression of Spanish SPPs in the data of the current study.

4.1.5 Interview modality

Interview modality was considered as a factor in the current analysis since the sociolinguistic interviews conducted in this investigation were carried out through in person contact, phone calls, or video calls. This proved to be significant specially since some of the interviews were collected during the year 2020, a year when the COVID-19 pandemic began. The pandemic proved to be a big obstacle in many investigations involving human subjects. However, the availability of technological tools such as Skype or Zoom, now makes it so that many research activities can continue to be conducted. Although distance interviews may not always be appropriate in linguistics research, as may be the case of investigations which focus more heavily

on phonetic analysis and thus require specialized equipment or a sound-proof environment, the current investigation did not require any of these more specialized tools. My analysis focused on morphosyntactic characteristics of participants' Spanish language use, as well as on their specific comments regarding their experiences as bilingual speakers in Canada. Therefore, the small sacrifice in sound quality that may have been made was greatly outweighed by the opportunity to continue data collection despite the pandemic restrictions. Even outside of the context of the pandemic, the ability to conduct sociolinguistic interviews at distance proved to be beneficial. The ability to offer participants a choice on the modality of the interview, made it possible for them to schedule time to participate even when balancing busy schedules since they could participate from their home, office, or preferred location. It is nevertheless important to consider, when beginning to vary the methodological approaches, whether any effects from the methodological choices could impact the variables being researched. In this case, since the variable expression of Spanish SPPs is conditioned by morphosyntactic as well as discourse-pragmatic factors, this was specially important.

In the end, the modality used for the interview was selected as an important conditioning factors in the variable expression of Spanish SPPs for second generation speakers only (section 3.1.3.3). Specifically, second generation speakers used overt subject pronouns at a statistically equal rate in video calls and in in-person interviews, but at a lower frequency during telephone interviews. This suggests that there is an effect of visual cues that play a part in the production of Spanish SPPs which leads to video-call interviews more closely resembling in-person interviews than phone interviews. It is, however, somewhat surprising that there is a decrease in use of overt SPPs since it is in this modality specifically that I would expect there to be an increased need for reiterating the subject and clarifying possible ambiguities verbally. One possible explanation for this pattern, in my opinion, may be in the comfort and type of discourse that was achieved with participants while using the different modalities. Studies such as that by Lastra and Butragueño (2015), have shown that the variable expression of Spanish SPPs can vary according to the discourse style. For instance, the authors found that speakers

showed a higher use of overt SPPs later on in the conversations when they had longer speech turns than at the beginning of the interviews conducted. This may well serve as an explanation for the observed pattern in our data. From an anecdotal point of view, I can report that in many cases I was better able to engage with participants and to reach a more comfortable conversation style when visual cues were available. In addition, the lack of visual presence of the interviewer, and the ability of the speaker to be interviewed from their home, can lead to a more disengaged conversation where participants are multitasking at home, and to greater difficulty building a comfortable rapport with someone who you haven't met. For instance, anecdotally, I experienced instances with at least two speakers where they seemed overly distracted and seemed to lose track of the conversation at times, and one instance where a speaker participated in the interview while driving (using hands-free technology) home from work. These less engaged conversations may have led to a different style of conversations, which may have affected the speakers' use of overt SPPs. Alternatively, it would be important to consider the relationship between length of speech turn and participants' rates of use of overt SPPs as well. In the case of many second generation speakers, regardless of whether they were in in-person, video, or phone interviews, their responses were shorter to most questions. This could be due to speakers' lack of confidence in their verbal abilities in Spanish, which many speakers did casually mention (sometimes even before the interview began to ensure they were in fact eligible to participate). In future investigations, modality as well as length of speech turn should be considered to differentiate between effects due to each of these factors.

It is important to highlight as well, however, that differences according to modality of the interview were not found to be significant for first generation speakers. In the case of this group, no participants participated in video-calls, and only three participated in distance interviews over the phone as most first generation participants preferred to meet in person. This may have influenced the results in that there may have been insufficient data from distance interviews from this group to allow differences to reach significance. However, in this group as well, we see a similar trend as that seen with second generation speakers, where use of overt SPPs is

less common in phone interviews than in in-person interviews.

4.1.6 Intergenerational analysis of the use of Spanish SPPs

1 No significant change in the rate of overt SPPs between generations

The current investigation determined through a multiple regression analysis that generation was not a significant factor in conditioning the variable use of Spanish SPPs (section 3.1.4). This agrees with results from investigations such as that by Nagy (2015), which also focus on examining the use of null vs. overt subject pronouns across generations of immigrant speakers of [+ null subject] languages in a situation of language contact with English. The author found that frequency of use of overt subject pronouns did not change significantly between first and second/third generations of speakers of Italian, Cantonese and Russian living in Toronto (Nagy 2015).

However, further statistical analysis in my study suggest that there are important differences between the two generations in regards to the frequency of use of overt vs. null Spanish SPPs when considering the averages for each group. Specifically, second generation speakers showed a higher frequency of use of overt SPPs (30.7%) as compared to the first generation of speakers (23.5%) and this difference was found to be significant after conducting a chi-square test. This result seems to align with findings of similar studies of English-Spanish language contact which show that speakers of Spanish who had been residing in the United States for longer periods of time tend to show higher frequencies of use of overt SPPs than newcomer Spanish speakers (Otheguy et al. 2007; Montrul 2004; Orozco 2004). Authors of these investigations argue that this increase is due to an erosion of discourse-pragmatic factors which condition the variable use of overt SPPs due to intense language contact with a [- null subject] language such as English.

2 Increased use of pre-verbal subjects

Despite the fact that the difference between generations in regards to the frequency of use of overt SPPs was determined to be not statistically significant, a significant difference was found between generations when considering the frequency of use of pre-verbal SPPs (section 3.1.4.1).

According to the Null Subject Parameter (NSP), null-subject languages, such as Spanish, tend to have a number of characteristics in common including the ability to omit subject pronouns and the ability to have flexible subject pronoun placement. Subject pronouns in Spanish when overtly produced can be placed either pre-verbally to follow a more traditional SVO order, or post-verbally in VS order. The placement of subject pronouns in Spanish varies according to discourse-pragmatic factors. For instance, post-verbal subjects tend to be associated with introduction of a new referent in the discourse, and with unaccusative verbs (Barrera-Tobón and Raña-Risso 2016; Silva-Corvalán 1994b). However, in bilingual communities such as Hispanic communities in the United States, an investigation by Barrera-Tobón and Raña-Risso (2016) found an increase in pre-verbal subject placement among second generation speakers. That is, second generation speakers show a more rigid SVO word order than Hispanic speakers in monolingual communities. For instance, Barrera-Tobón and Raña-Risso (2016) studied two generations of a population of Puerto Rican Spanish speakers in New York, and determined that there was a significant increase in the use of pre-verbal subjects in second generation speakers when compared against first generation speakers who were newcomers in New York. In the current study, when the placement of subject pronouns was analysed, I observed a similar finding. Second generation Colombian speakers living in London, Ontario use pre-verbal pronominal subjects with a significantly higher frequency than first generation speakers. This observed pattern seems to align with the study conducted by Barrera-Tobón and Raña-Risso (2016), as well as with the study conducted by Silvia-Corvalán (1994) who found that second and third generation Mexican Spanish speakers in Los Angeles showed a more rigid SVO order as opposed to first generation Mexican Spanish speakers. This increased preference for pre-

verbal subjects may reflect a language contact effect since [- null subject] languages such as English, have a more rigid SVO word order. According to Silva-Corvalán (1994a) the increase in use of pre-verbal subjects in English-Spanish bilingual speakers living in the United States shows a grammatical convergence with the use of subjects in English. Silva-Corvalán (1994a) explains that this may be due to a loss of Spanish pragmatic constraints for pre-verbal subject placement. That is, these bilingual speakers become less sensitive to the pragmatic conditions that in monolingual Spanish would be more likely to result in post-verbal subject placement. Since the results of this investigation align with these findings, it is suggested that the significant increase in use of pre-verbal subjects indicates evidence of language contact effects in Spanish speakers in Canada.

4.1.7 Complexification in the second generation group

The factors of verbal mood and clause polarity were found to be significant conditioning factors for second generation speakers but not for first generation speakers. This could signal that second generation speakers are increasingly relying on other factors to condition their use of overt SPPs. Note, however, that verbal mood and clause polarity, although not significant conditioning factors for the first generation group, follow similar tendencies in both generations. That is, both first and second generation groups show a higher frequency of use of overt SPPs when the verb is in the conditional and subjunctive moods, and a lower frequency of use for verbs in the indicative mood. Similarly in the case of clause polarity, both first and second generation speakers favour the use of overt SPPs in the case of [- negative] clauses. This indicates that although these factors did not reach significance within my analysis for first generation speakers, they are still relevant factors which contribute to the variable expression of Spanish SPPs. This in turn indicates an increased reliance on these factors within the grammar of second generation speakers rather than an innovative use of these factors by this generation.

Further, when we consider the results for the second generation analysis, the factor of switch reference is not one of the two top linguistic conditioning factors within the Rbrul model

as it is with first generation speakers. In fact, we see that switch reference, within the second generation group, drops in the constraint hierarchy below verbal mood. In addition, and as mentioned in my discussion of change of referent as a significant factor, there are quantitative differences in how second generation speakers behave in regards to this factor. Specifically, second generation speakers use more overt SPPs in contexts where there is referent continuity than first generation speakers. This evidence then could be not only showing an erosion of the discourse-pragmatic factors that Spanish relies on for constraining the variable expression of Spanish SPPs, but also an additional reliance on morphosyntactic cues of ambiguity by second generation speakers. Since according to Montrul (2004) and Sorace (2011) it is features at the interface that are more susceptible to erosion during L1 attrition, it is therefore in my opinion possible that second generation speakers are relying on more stable morphosyntactic cues in order to constrain their variable expression of Spanish SPPs and reduce ambiguity. This proposal has also been suggested by Shin (2014) previously. Shin (2014) explored differences between two generations of Hispanic speakers in New York and found that TMA (tense, mood, aspect) considerations significantly constrained the variable expression of third person singular Spanish SPPs in second generation speakers but not first generation speakers. She highlights that this use represents a complexification on the constraints in second generation speakers' grammar which condition the variation within a situation of language contact, and that this complexification is likely related to a need to reduce ambiguity in conversation. One possibility that she then puts forth is that second generation speakers are reacting to increased ambiguity due to simplification of the system elsewhere, which she suggests is likely a loss of sensitivity to the switch reference factor, by complexifying other aspects such as increased sensitivity to TMA cues.

In regards to the increased sensitivity to polarity of the clause, the explanation seems less clear, since it is not evident that negation would render a clause more ambiguous. One possible explanation for this additional observation of complexification may be related with the second generation group's increased preference for pre-verbal subjects. In both generations

there is a trend towards a lower use of overt SPPs with clauses which are negative. Further, negative elements are usually found pre-verbally in Spanish and some authors have claimed that these negative clauses tend to be related with lower overt SPP usage in Spanish because pre-verbal element negatives compete with subjects for the pre-verbal position in a given structure (Gridstead 1998). In these cases an overt SPP which accompanies a negative clause would be in a left-dislocated position serving a more emphatic purpose or it would be more likely to be placed post-verbally. As second generation speakers in my sample were shown to have an increased preference for pre-verbal subjects as compared to first generation speakers, it is then possible in my opinion, that the increase in significance for the polarity factor observed in second generation speakers is a result of the quantitative increase in instances where negative clauses competed with SPPs for the pre-verbal position in the phrase. Note, however, that clause polarity is the lowest ranked constraint in the hierarchy for second generation speakers and therefore there may only be a small tendency to have an increase in null subjects with negative clauses.

4.1.8 A case for cross-linguistic influence in second-generation Colombians in Canada

In my study, the main results when we consider the population of Colombian Spanish speakers living in London Ontario seem to signal to cross-linguistic influence from English on the use of Spanish SPPs by second generation speakers. Notably, when considering second generation speakers we see in the results of this investigation: a significant increase in preference for pre-verbal subjects; and a change in the hierarchy of factors conditioning the use of overt SPPs when compared with first generation speakers.

Through my analysis in this investigation I observed an increase between generations in the use of pre-verbal subjects as opposed to post-verbal subjects. Flexible placement of subject pronouns is a characteristic of Spanish, which is also thought to be associated with the Null Subject Parameter (NSP). As discussed in section 1.4, according to the NSP theory, languages

such as Spanish which allow for null subjects, also share other characteristics including flexibility in placement of subjects (pre- and post-verbally). English, however, is considered to be a [- null subject] language which allows for null subjects in fewer contexts (see section 1.4.2.2) and which shows a more rigid SVO order. Therefore, in a situation of English-Spanish language contact, a change between generations towards a more rigid SVO order is considered to indicate cross-linguistic influence from and convergence with English (Silva-Corvalán 1994b). This was a pattern observed by Silva-Corvalán (1994b), who also did not find an increase in overt SPP usage between generations, but who did see an increased preference for preverbal subjects among second and third generation Mexican Spanish speakers in Los Angeles. Therefore, this finding in my current investigation, which aligns with these results, may also point towards convergence with English grammatical system among second generation Colombian Spanish speakers in Canada.

In addition to a more rigid SVO order, second generation speakers in my sample also showed a complexification of the hierarchy of linguistic factors which condition the variable expression of Spanish SPPs. That is, while first generation speakers' use of Spanish SPPs is conditioned primarily by the pronoun type (grammatical person and number) and the change of referent factors, second generation speakers' variable expression of Spanish SPPs was also conditioned by verbal mood and clause polarity. As discussed in section (4.1.8) above, complexification of this kind among second generation speakers can indicate a need to compensate for an effect that has been reduced, eroded or lost. In this case, we see a reduced impact of switch reference on the constraining of second generation speakers' use of overt vs. null SPPs, which can lead to a need for the use of additional cues to disambiguate contexts. This goal can be achieved by turning attention to a feature that is more stable and less likely to be eroded such as the morphosyntactic feature verbal mood. This is also proposed by Shin (2014) as she also observed complexification of the hierarchy of factors conditioning overt SPP usage among second generation Spanish speakers in New York City, with TAM taking a stronger role than switch reference.

Alternatively, it is also important to note that the drop in significance for the factor of change of referent in second generation group could be due to the fact that the speakers in this group all began acquisition of English before the age of eight. According to Shin and Smith Cairns (2009) the development of syntax-discourse interface features, such as the sensitivity to referent continuity as a conditioning factor for the variable expression of Spanish SPPs, occurs later than the development of syntax-discourse interface features regarding lexical NPs. In fact, the authors highlight that sensitivity to referent continuity develops between the ages of seven and nine. In the context of my investigation, therefore, it is possible that second generation speakers, all of whom arrived in Canada prior to the age of nine, experienced cross-linguistic influence from English during this developmental period. Further, according to Müller (2019), simultaneous bilingual children who are receiving input from two different languages may receive conflicting information which impacts their development of adult-like structures resulting in a situation of cross-linguistic influence. The authors explain that some aspects of grammar are more susceptible to cross-linguistic influence. Notably, the aspects at the interface between two modules of grammar, such as is the case with the variable expression of Spanish SPPs which lie at the syntactic-pragmatics interface, are most vulnerable to cross-linguistic influence. Further, the authors note that cross-linguistic influence is only likely to occur if a bilingual child is exposed to one language which provides evidence which opens the possibility for two grammatical analyses, while being exposed to a language that lends support to only one of the two analyses. In this case the child received conflicting information and may take “short cuts” to facilitate language processing thus allowing a grammatical analysis from one language into the other. In the case of the current study, the use of variable expression of Spanish SPPs in contact with English seems therefore to provide the perfect conditions to see cross-linguistic influence among the second generation speakers. We see not only that the variable expression of SPPs lies at the syntax-pragmatics interface, but that the linguistic evidence from Spanish and English provides contradictory linguistic evidence to the developing bilingual child which can then lead to “short cuts” which result in errors such as a higher

use of overt SPPs than used by monolinguals through a misinterpretation or underdevelopment of sensitivity to pragmatic factors that condition this variable in monolingual Spanish but not monolingual English, such as the sensitivity to switch of reference. That is, ultimately, the conflicting information from Spanish, a [+ null subject] language, and English, a [- null subject] language, may have resulted in a less developed sensitivity towards switch reference as a conditioning factor for the variable expression of Spanish SPPs. This would then explain the lower ranking of referent continuity in the hierarchy of conditioning factors among second generation speakers, and would also explain the lack of effect among first generation speakers who arrived with a fully developed grammar and are therefore less susceptible to language contact effects.

These factors mentioned above suggest that there is erosion or underdevelopment of discourse-pragmatic features leading to cross-linguistic influence effects in second generation Spanish-English Colombian bilinguals in Canada. Further, I want to highlight that although generation was not selected as a factor which significantly conditions the variable expression of Spanish SPPs, there was an observable trend which showed that second generation speakers used overt SPPs with a higher frequency than first generation speakers, and the non-significance of the result could be a result of the particular sample of speakers used. The question needs to be raised about whether the second generation speakers in my data who showed higher frequencies of use of overt SPPs of between 30% and 40% may not in fact be more representative of second generation speakers in London Ontario than my data would suggest. I have some anecdotal evidence suggesting that the fact that in order to participate in this study speakers were required to speak and understand Spanish (in order to participate in the interview) discouraged some second generation speakers from participating as they perceived their level of Spanish as insufficient. In more than one case, despite my reassurance that using a mix of Spanish and English would be appropriate and that “perfect” Spanish was not necessary, speakers chose not to participate if they considered their language skills in Spanish to not be sufficient. This can mean that in many cases the second generation speakers who did volunteer to participate are

among the more advanced Spanish speakers in their generation in this population. Therefore, the generational difference in the real population of Colombian Spanish speakers in London Ontario may be more pronounced than my results indicate. This, however, can not be confirmed and will need to be addressed in future investigations by collecting data from a different sample of second generation participants.

Alternatively, it is important to mention that the difference observed between the two generations may be due to other differences in the speakers' use of structures that favour overt vs. null SPPs. Specifically, within the coded samples, second generation speakers were more likely to use singular SPPs than plural SPPs, and they were more likely to use structures with novel referents (either through a switch in referent or through a SPP in utterance initial position) than same referent structures as compared to first generation speakers. Since singular SPP constructions and novel referents are conditioning factors which favour the use of overt SPPs, the higher use of overt SPPs in second generation speakers may be reflecting these differences of use between the two generations rather than a true overall tendency of second generation speakers to use overt SPPs in contexts with true optionality.

It is also important to mention at this point that while some evidence in the current analysis, such as the preference for pre-verbal vs. post-verbal subjects among second generation speakers, could be used to support a theory of convergence towards a more English-like system, the fact that no effect was found among first-generation speakers in the sample, regardless of any measure of language contact intensity, leads to the conclusion that the observed cross-linguistic influence effects are more likely to be related to the speakers' bilingual first language acquisition process during which they were exposed to two competing grammars rather than to the overall exposure to English over an extended period of time.

In the United States, patterns observed in investigations which show an increase in use of overt SPPs across generations of Hispanics may nevertheless still be signaling an effect of influence from dialectal contact with Caribbean varieties of Spanish which are known to have some of the highest frequencies of use of overt SPPs (33-41%) in the Hispanic world

in addition to effects of contact with English. This is proposed since the extent to which the frequency of use of overt SPP production increases between generations in the United States seems to be more extreme than the increase observed in this investigation which did not reach a statistical level of significance. It is possible that despite the lower level of intensity of contact with English experienced in the United States, dialect leveling effects may also be contributing in an additive manner. The suggested possibility of an additive effect on United States' Hispanic populations from English and Caribbean varieties of Spanish has also been previously presented by (Otheguy et al. 2007).

One way in which this additive effect may be operating may be in terms of differences in norms of language prestige. That is, since in the United States, particularly in regions such as New York and Miami, it is Caribbean Hispanic populations that first arrived and established themselves, upon arrival to these communities, it is their language norms that enjoy greater prestige and to which newcomers assimilate as they incorporate themselves into the existing communities. These varieties of Spanish, although common in the United States thanks to influence from Puerto Rico, Cuba, and Dominican Republic, are more rare in Canada. Therefore, it is possible that the differences between communities in regards to the extent to which overt SPPs increase in frequency across generations (or as a result of length of residence) reflects differences in the prestige norms between communities in the United States and Canada. This could also explain why only some communities in the United States have documented an increase in overt SPP expression. For instance, although Silva-Corvalán (1994) suggests that Mexican bilingual speakers in her data are undergoing a process of erosion of discourse-pragmatic factors which lead to a more rigid SVO order, she found that the rate of use of overt SPPs remained statistically unchanged across three generations in Los Angeles. This pattern could be explained by the greater presence of Mexican populations in Los Angeles, who are speakers of a variety of Spanish with a lower frequency of use of overt SPPs, and who have been established in this region for longer than most other Hispanic populations the United States.

Further evidence of dialect contact effects based on prestige level has also been previously reported in regards to overt SPP usage in some monolingual varieties of Spanish. For instance, a study by Orozco (2015), showed a decrease in the use of overt Spanish SPPs across generations of Colombian speakers from Barranquilla, whose language variety is a Caribbean variety of Spanish. The author explains that this may be due to increased pressure to adopt a style that is closer to that of the Colombian prestige Spanish variety, which is that of Bogotá, a mainland Latin American Spanish variety. It is therefore possible that stronger differences across generations in some regions of the United States may also be due to a combination of language and prestige dialect contact effects.

In conclusion, results of this investigation when comparing use of Spanish SPPs across generations shows evidence of cross-linguistic influence in second generation bilingual speakers towards a grammar that is closer to that of English probably due to conflicting linguistic evidence during bilingual first language development. This is likely occurring, as explained by Silva-Corvalán (1994a), due to a loss of discourse-pragmatic constraints which leads to a more rigid SVO order. This same process can also lead to decreased sensitivity towards discourse-pragmatic factors which constrain the variable expression of Spanish SPPs (Shin and Smith Cairns 2009). This is in fact likely to be the case for second generation Colombian Spanish speakers living in London, Ontario. Although differences in use of overt SPPs between first and second generation speakers were not found to be statistically significant, this trend in combination with the increased preference for pre-verbal subjects, and the change in the hierarchy of factors conditioning the use of overt SPPs among second generation speakers, suggests an effect of contact with English through cross-linguistic influence.

4.2 Spanish language maintenance in Londombia

In this section, we discuss the results from the qualitative thematic analysis conducted in this investigation. Through this analysis I consider the habits, attitudes and opinions of

Colombian Spanish speakers living in London, Ontario to determine whether they align well with factors that are known to support language maintenance in heritage language communities in a situation of language contact.

4.2.1 Language use patterns

The results regarding the use of Spanish by Colombian speakers living in London, Ontario suggest that, with some limited exceptions, the use of Spanish is primarily restricted to the home and family environment (section 3.2.1.1). This is partly due to considerations of attitudes from members of the majority language community towards the use of Spanish at work or in public, and to a lack of sense of Hispanic community and few available resources. In the case of second generation speakers it is further reduced to instances when they communicate with first generation speakers. Although the home is the primary domain used by many communities to maintain a minority language, and therefore this observation is not atypical, according to Unsworth (2013), heritage language speakers need to have access to an appropriate quantity and quality of input in order to fully acquire productive proficiency in the heritage language. This requires communication in Spanish beyond the home environment, whether at school, socially, or professionally. This lack of opportunity to use Spanish in different contexts can then be a major contributor to language shift in the community within two or three generations.

It is important to highlight as well that in the CoSLO corpus, many of the second generation speakers who participated are the children of the first generation participants. This was largely due to the snowball sampling methodology employed. In addition, since the interviews were entirely conducted in Spanish, a high receptive and productive proficiency in Spanish was part of the inclusion criteria for participation. The participants are therefore speakers who were successful in acquiring (in the case of second generation speakers themselves) or in transmitting the language to the second generation of speakers who arrived as children or were born in Canada. As a result, it is difficult to comment on differences between families who are and those who are not successful. It is also not possible to determine how typical the language

transmission success achieved by these families is in comparison to other Colombian families in London, Ontario. Further, as mentioned in section (4.1.9) above, this also means that it is not possible to determine how typical the Spanish use of second generation speakers in my sample is when compared to the larger real population of second generation Colombian Spanish speakers in London, Ontario.

1 The home and family environment

Investigations of language contact and language maintenance have determined that among the main factors that contribute to the intergenerational maintenance of a heritage language is the extent to which speakers of the language continue to use their language on a regular basis. This relates to both quantity and variety of use. That is, considering how much speakers use their language and in which contexts. This is influenced by a number of different factors which I took into consideration in my analysis such as the linguistic background of family and friends, the opportunities to use the language in their professional life, and their perceptions of others' attitudes towards their language in the wider community.

When considering the home context, we see that speakers in both generations declare Spanish to be the primary language used in the home (3.2.1.1). However, differences in regards to this first general statement emerge when speakers are asked about further details on who in the household speaks Spanish or English to whom. While first generation speakers declare trying to impose the use of Spanish in the home, second generation speakers report using Spanish only when the parents (first-generation speakers) are present. Even then, when we further delve into specifics, both first and second generation speakers reported that a mixture between the two languages is more common, especially when younger siblings or other family members who are less proficient in Spanish are involved in the conversation. Further, when second generation speakers were asked about their use of Spanish in the home, they all expressed using English mostly with siblings in the majority of the time. The exceptions noted were if the parents were present, or if they were in a context where they wanted others to not be able to understand their

conversation.

The situation where the majority language, in this case English, is increasingly used with younger siblings in the home is not uncommon in the literature. According to Silva-Corvalán (1994a) younger siblings tend to show a lower proficiency in the heritage language, in this case Spanish, because as soon as they begin interacting with their school-aged siblings English becomes the principal language in which they interact. This then reduces the already limited input that these younger children receive since the majority language is now introduced in the home. In my data this was the case when siblings of similar age. At least two first generation speakers who were parents of two or more children expressed that their younger child had a lower or very limited proficiency in Spanish as compared to their first-born child. This trend, along with the reports of second generation speakers that they use Spanish only when the parents are present, points to a gradual disappearance of Spanish in the home context since it can be reasonably expected that when the children leave the family home (all second generation speakers were still living at home) use of Spanish in their own home may continue to decrease particularly if their future romantic partner is linguistically exogamous.

2 Social (friendship) contexts

Within the social environment, there were clear differences between the two generations (section 3.2.1.2). Mainly, first generation speakers who reported feeling more comfortable interacting with Hispanic or Colombian individuals noted that their closest friendships consequently were also first-generation Spanish speakers and that in their social interactions Spanish is the primary language spoken. One exception to this were situations where an individual who did not speak Spanish was present, but this was noted as a fairly rare occasion. On the contrary, most second generation speakers, with some exceptions, noted that the majority of their friends were not Hispanic and therefore English was used. In addition, even when interacting with other Hispanic friends, most of their friends were also second generation speakers and among them they communicate mostly in English, or a combination of English and Spanish.

Some exceptions to this were noted, however. For instance, some second generation speakers reported having met a few friends who had recently arrived in the country and did not speak English too well. In this case, second generation speakers make an effort to communicate in Spanish only, but some speakers report that this can be a conscious struggle. The other exceptions reported were cases where two or more Spanish heritage speakers were together in public and wished to communicate without being understood by others. Notice that this pattern seen within the social context converges with the pattern seen in the home context reiterating that second generation speakers tend to use Spanish only when in the presence of first generation Spanish speakers, whether that be their parents or friends who recently arrived in the country. It is also important to highlight that this pattern seen in the second generation speakers' social life may further indicate that it is unlikely that the use of Spanish will remain strong throughout their lives, especially as they form social and possible romantic relationships within their social circles, which as reported are mostly non-Hispanic or Hispanic but where English is the primary language used.

3 Professional contexts (School and/or work)

In considering the use of Spanish in the workplace and in public contexts, similar patterns emerged for both first and second generation speakers (3.2.1.3). Notably, both first and second generation speakers reported rarely or never using Spanish at school or in the workplace. This was due, in most cases, to lack of opportunities to do so. That is, the places of work conduct business primarily in English, and the schools attended by the second generation speakers offer instruction primarily and often exclusively in English only (with the exception of some Spanish courses taken by some participants). In the case of some first generation speakers, they occasionally use Spanish when a co-worker or someone at their place of work is also Hispanic. However, these instances remain informal in nature and are in most cases rare.

In contrast, in the case of second generation speakers, they tend to use Spanish only when informally serving as interpreters in their place of work or to provide service to Spanish speak-

ing clients who have limited abilities in English. It is unclear, however, whether second generation speakers bypass opportunities to use Spanish when the use of Spanish is not absolutely necessary, for instance when working alongside other Spanish-English bilingual workers. Unfortunately, this question was not addressed in the interview and should be considered in future investigations.

Another consideration that emerged was speakers' perceptions of others' attitudes towards their use of Spanish in the workplace. Although reports were definitely not unanimous, some speakers in both first and second generation groups reported that they had been told or felt that the use of Spanish was discouraged while at work. Notably, at least one first generation and one second generation speaker were overtly told by managers or co-workers to not use Spanish because they could not understand them.

The lack of use of a minority heritage language in the workplace is a common observation in studies of language maintenance (Pauwels 2016). Speakers from minority language communities, such as is the case of immigrant communities like the Hispanic population of London, Ontario, tend towards an increased use of the majority and dominant language as they enter the workforce in an attempt to secure better financial opportunities and professional success. This tends to be reverted later in life, for instance once speakers have retired, when their frequency of use of the heritage language may again increase. It is therefore not surprising that the speakers in the CoSLO corpus, who at the time of the interview were all between the ages of 18 and 66 and all still actively working and/or studying full-time, do not use Spanish while in a professional context.

4 Public contexts and majority community's perceptions

Perceptions of others' attitudes towards their use of Spanish was an important consideration as well for some speakers when in public (sections 3.2.1.3). However, speakers' behaviours even when they perceived negative attitudes towards their use of Spanish in public were different from person to person. This was most evident with first generation speakers. While some

speakers took the approach of not being concerned with others' attitudes, describing it more as "that's their problem, not mine", others considered that it was more respectful to speak in English when in public so that everyone understood what they were talking about. Second generation speakers, in contrast, mostly are unconcerned with others' attitudes towards their use of Spanish and their choice of language in public is more dependent on the language spoken by the individuals they were with. Alternatively, some second generation speakers reported speaking in Spanish to avoid being understood by others. This pattern in public contexts, similarly to those reported in the social or home contexts, highlights the strategic choice of language by second generation speakers in accordance to the need for privacy, or use of language for specific purposes. That is, second generation speakers will sometimes speak in English in the home to avoid being understood by parents, and will use Spanish socially and in public to avoid being understood by others around them. That being said, in some limited cases, at least two second generation speakers reported feeling "nervous" when speaking Spanish in public due to looks from others.

Hesitation to use Spanish in public due to fear of discrimination was reported by some speakers in both first and second generations. This hesitation comes partly from known recent events of discrimination towards Hispanic populations reported in the news, both in Canada and in the United States. These instances, which in some cases lead speakers of Spanish to attempt to speak in English even when not proficient in the language, can lead to a further reduction in the variety of contexts and situations where Spanish is spoken in the community by members of the Spanish heritage language community. This reduction in use of Spanish when in public, even in situations where individuals are with other Hispanics who have limited proficiency in English, can have the greatest impact on language maintenance when we consider the input received by younger second and third generation Hispanic children. Regardless of whether the decrease in use of Spanish in any domain is due to an attempt to be more respectful or for fear of others' perceptions, if parents are hesitant to speak in Spanish, for instance while grocery shopping, they increase instances in which they communicate to their children in English and

reduce both quantity and quality of Spanish language input for their children since topics relating to shopping and money, for instance, now become topics that are discussed in English. This leads not just to a reduced quantity of input, thus increasing the chances of incomplete acquisition in younger children, but can also lead to register reduction which can heavily impact the outcomes of language maintenance efforts (Unsworth 2013). In addition, according to Pauwels (2016), in situations where only one parent is a speaker of the heritage language, there is a need to strictly enforce the one-parent-one-language strategy. If the heritage language speaking parent speaks to the child in the heritage language only some of the time, the child will tend to communicate in the majority language with the parent more frequently and the opportunities to use and practice the heritage language in the home will diminish perhaps to the extent of being below the necessary level of language input for successful language acquisition. This may perhaps already be the situation being observed when considering some of the youngest members of the Hispanic community in London, Ontario. Participants interviewed in this study who reported that they spoke to younger children in their families (third generation Heritage speakers) in Spanish only some of the time, also mentioned that the children rarely communicate to them in Spanish, even when being addressed in Spanish. The overall consensus was that although these third generation children have some receptive proficiency in Spanish, they rarely ever make the effort to speak to first or second generation family members in Spanish. Pauwels (2016) explains that there is a strong and stable association between persons we communicate with and the language we use to communicate with them. If the heritage language is not established early on as the only language used to communicate with a given family member (whether a parent, uncle, aunt, or grandparent), the child will revert to the more dominant language, which tends to be the majority community language in a situation of language contact where the heritage language is a minority language only spoken by a small community and often not used outside the home.

4.2.2 Attitudes towards intergenerational language transmission

The attitudes of speakers towards their own language and its use by new generations is an essential consideration when we consider language maintenance efforts in a minority community (section 3.2.2). For instance, in the case of communities where the minority language is a non-standard variety or a variety that is stigmatized in the larger majority community, the language can begin to be rejected by the minority community itself as they seek to improve their own standing in the larger environment both in terms of prestige and financial or professional success. This is usually more so the case of smaller language varieties with no standardized grammars such as creole languages, African-American Vernacular English and Indigenous languages such as Maya in the Yucatan peninsula. However, immigrant communities such as Hispanics in Canada can be under comparable pressures depending on the status of the language in the region or area where it is used. In Canada, although Spanish itself is not stigmatized, and bilingualism and multiculturalism is openly encouraged by Canadian culture, newcomers from Hispanic countries (or any other countries where a foreign language other than French or English is used) are encouraged to quickly learn and increase their use of at least one of Canada's two official languages (English and French). This was reported, for instance, in an article by Guardado (2008) where he investigated the experiences of Hispanic families in Vancouver. Many of the families interviewed reported feeling pressure not just to learn the language themselves but to ensure their children acquired and used English at home to ease their assimilation into the new school environment. These kind of pressures are important to consider as they could potentially shift a community's priorities from language maintenance of the minority heritage language to assimilation into the new environment.

In the current investigation, when participants were asked about the importance they attributed to intergenerational transmission of the Spanish language, the primary theme that emerged was that maintaining Spanish across generations was a clear ideal and desire for most families. Among the reasons given to support this decision, most families mentioned reasons relating to the relevance of bilingualism for professional success, as well as cultural and family

considerations. It is important to note that no differences between the two generations were found in this respect.

Participants' comments show a very strong expressed desire to maintain the Spanish language. However, the likelihood of maintenance of the language can be decreased when bilingualism or professional success are the primary considerations. This is reflected for instance on the comments by several speakers who emphasize "learning another language" as their primary motivation to teach Spanish to their children while mentioning, for instance cognitive benefits of bilingualism. Many of speakers with these goals in mind also noted that learning French or Spanish, or both would be beneficial. In fact, at least one speaker believed that although learning Spanish would be nice, that this was more so a selfish consideration on their part since French would really be a more useful language to learn for future generations of Hispanics in Canada. In addition, at least three participants from the first generation group mentioned having placed their children in French immersion programs to foster official language bilingualism. In this way, bilingualism and professional success considerations as motivations to maintain the language can fail in encouraging Spanish use in Canada, while shifting the focus to more in demand languages in the national job market such as French (or even Mandarin in one case).

Other difficulties in pursuing maintenance of Spanish were also reported by participants. Notably, the abundant presence of English and lack of spaces to use Spanish are concerns for families who note that while maintaining the use of Spanish across several generations would be ideal, it seems to be an overly difficult task and unlikely to happen. Taking it one step further, one second generation speaker noted that they do want their future children to speak Spanish, but that it is not a main priority, especially if they decide to marry someone who is not a Spanish speaker.

These sources of added difficulty in achieving goals of language maintenance are important differences when we consider the maintenance of Spanish as a heritage language in Canada vs. in the United States. In the United States, the presence of Spanish is significant in many areas

such as Miami, New York, California, and Texas, among others. In fact, the United States is considered to this day the fifth country with the most Spanish speakers in the world, and services are often available in English and Spanish primarily. Therefore, motivations for bilingualism based on financial and professional success in the United States are very connected to the learning of Spanish. In Canada, however, French competes as a more in demand language offering better economic success, and the presence of Spanish is restricted to Latin convenience stores and Hispanic run services.

Literature on maintenance of heritage languages has previously expressed that families who are more successful in maintaining the use of the heritage language in the family across generations show a focus on culture and family, instead of bilingualism or financial success (Guardado 2008). Although it is true that more professional opportunities may be available for English-French bilinguals, maintenance of Spanish in Hispanic immigrant families encourage maintenance of ties with family members. This shift in focus would be more beneficial in aiding maintenance of Spanish in Canada, and is already present in a portion of the population interviewed. Among the participants interviewed in the current investigation, weakening of family ties was a source of concern for some first and second generation speakers. For instance, some first generation speakers were either already experiencing difficulties in communicating with their grandchildren and feeling a sense of disconnect from them as they were sometimes unable to fully understand the stories and sources of excitement of their grandchildren, or feared not being able to do so in the future. In addition, some second generation speakers seemed to also had reflected on these concerns and expressed wishing to maintain the use of Spanish in their homes in order to help facilitate family communication with their first generation parents and the transmission of cultural traditions that are strongly tied to the language such as the Christmas *novena de aguinaldos* and other typically Colombian traditions (section 3.2.3.2). This source of motivation, when adopted, is more likely to maintain the use of Spanish since no other languages can achieve the goals of cultivating cultural and family ties.

4.2.3 Attitudes towards endogamous/exogamous marriages

Following the considerations raised while discussing the attitudes to intergenerational transmission, I now also consider the connected factor of attitudes towards culturally and linguistically endogamous relationships and/or marriages. This topic is an essential factor to investigate since it informs us further about the expected changes in the home and family environment in future generations. According to Montrul (2011), one aspect that leads to language shift towards a majority language is a decrease in the number of contexts where the heritage language is used by the heritage speakers. In the case of second generation speakers in this investigation, all of them were still living at home where they were in daily contact with first generation speakers. This provides them with a space in their daily lives where they can use and practice their language. However, in the case of minority language communities, when speakers enter into linguistically exogamous marriages, this can lead to a further reduction of the contexts and opportunities where the heritage language is used. In addition, as the speakers likely communicate with their significant others in the majority language, any children in the future will have a further reduced quantity and quality of linguistic input and intergenerational transmission to the third generation of speakers is unlikely to occur.

When speaker participants in the current investigation were asked about the importance of marrying a Hispanic or Latin@ individual (either for themselves or for their children), most speakers from both generations expressed that this was not an important consideration (section 3.2.3.1). The Colombian speakers interviewed consider that other characteristics are more important when choosing a partner and language itself is not a primary consideration. This was common among first generation speakers who, although they often arrived to Canada already married to a Colombian spouse, considered that their children should instead take into account other virtues and personal characteristics such as honesty, loyalty, sense of humour, or intelligence when selecting a partner.

Some speakers also mentioned cultural compatibility as a more significant factor to take into account when selecting a romantic partner. That is, although language was not a primary

consideration, some speakers noted that marrying someone whose family values and culture were similar would make a marriage more successful. This in itself may in some cases encourage relationships with other Hispanics whose cultural background is more similar. However, in the case of at least two second generation speakers, they expressed that having a partner who is willing to engage in Hispanic cultural traditions is sufficient even if they do not speak the language. Specifically, one second generation speaker noted that in the past she dated someone who was too hesitant to try Colombian traditional foods, or to participate in cultural traditions, and this made it so that she could not continue the relationship. Although her current relationship is also with a non-Hispanic individual, she reported the relationship worked better since he was “open to other cultures”.

In contrast, one more theme that emerged in the data among some second generation speakers, and only one first generation speaker, was a desire to marry a Hispanic or Spanish-speaking individual to help facilitate communication with Hispanic family members (section 3.2.3.2). This pattern was seen when considering communication with older Hispanic family members (such as first-generation Hispanic parents), or younger future or present Hispanic family members. For instance, two female second generation siblings explained that they could see how frustrated their mother became when trying to communicate with their non-Hispanic sister-in-law. They each separately mentioned having observed this and reported this being a new motivator to ensure any future romantic partners were Hispanic or able to communicate in Spanish. On the other hand, some second generation speakers mentioned having a strong desire to ensure that any future children in the family were able to communicate with their grandparents, and that this was a motivator to ensure that any future partners were Hispanic since language transmission would be too difficult with a partner who did not speak Spanish. It is important to highlight that in both cases, this point of view was mostly mentioned by a few second generation speakers who, despite mentioning these motivations when considering their parents' ease of communication, they also expressed that linguistically endogamous relationships were not priorities and that they would not hesitate to enter an exogamous relationship. In addition,

we need to take into account the existing social interactions of these speakers. Even though a certain motivation exists to help parents and facilitate intergenerational transmission, most second generation speakers reported having few Spanish-speaking friends and communicating mostly in English with their few Spanish-English bilingual friends. If we consider that relationships are likely to arise from these social circles, and that their language use in social contexts are likely to remain unchanged, we can hypothesize that the majority of the second generation speakers interviewed in the CoSLO corpus are unlikely to carry on Spanish language use into their marriages and future home environments. This is supported by existing social network theory presented by Garcia (2003), who explains that a person's social network can affect the survival of the community language and the likelihood that language shift in the community will ensue. Therefore, the picture painted through analysis of second-generation speakers' social language behaviours, their social networks, and their attitudes towards exogamous relationships could translate into a more bleak outlook for the maintenance of Spanish language in Canada, since these speakers may soon lose the home and family environment contexts as a source of Spanish input, which remains the most important domain to foster heritage language maintenance (Fishman 1965; Wurm 2002). That being said, as Pauwels (2016) explains, linguistically exogamous relationships do not necessarily condemn a family to language shift towards use of the majority language. While it does make it more difficult to transfer the heritage language due to decreased linguistic input for the children of the household, and reduced input for the adult heritage speakers who are then at increased risk of language attrition as the environments in which they use Spanish continue to diminish, exogamous households can be successful in maintaining the heritage language and in raising simultaneous bilingual (or multilingual) children who are also more culturally diverse and who grow up with a wider understanding of the world, and cultural and linguistic diversity. This can be achieved, for example, through the one-parent-one-language approach and through frequent communication with Hispanic family members and friends.

4.2.4 Access to Spanish language resources (Heritage community impact)

Although it is generally accepted among language maintenance scholars that the home and family contexts are the most important domains to foster the intergenerational transmission and maintenance of a minority heritage language, the efforts from the larger heritage language community are also an important contributing factor (Fishman 1965; Wurm 2002; Valdés 2005; Pérez-Leroux et al. 2011; Montrul 2011; Pauwels 2016). One way in which minority language communities may come together in their efforts to maintain use of their language is through the creation of language schools. These institutions can vary from small sessions on weekends at the home of a community member to daily after school programs, or even more formal education through the foundation of bilingual schools. These programs vary widely from one community to the next in terms of their size, frequency, and the focus of the education program (from a focus on literacy alone to cultural/religious education among others). Although not all community-run schools are equally successful, communities which make the effort to offer these programs or provide other resources for the heritage community's youth are more likely to be successful in maintaining the heritage language than communities that don't. For instance, according to Montrul (2011), Hispanic families in Miami were more likely to be successful in maintaining the use of Spanish into the third generation when children were exposed to the language at home and attended a bilingual school where Spanish was also formally taught. The success of community programs doesn't just lie in the specific education offered in literacy in the heritage language, however, but also in creating social connections between community members and in providing a space where the heritage language can be used and encouraged, as well as in promoting pride in the cultural heritage and practices (Pauwels 2016).

In the community of Colombians living in London, Ontario, when speakers were asked about the availability of groups or resources to help children acquire and maintain the use of Spanish both first and second generation speakers expressed that they weren't aware of many, with the exception of one weekend Spanish school (section 3.2.4). While several speakers noted "having heard" about the Spanish school for children, which offers free classes for the

children of Hispanics in London on weekends, very few reported having encouraged their children to attend or having attended themselves when they were younger. First generation speakers reported that they felt it was unnecessary to bring their children to a school on weekends, that it was inaccessible due to it being held in a remote location which is not easily accessible via public transit (an important consideration for newcomer families who may not have a vehicle upon arriving in the country), or simply not being sure where it was. An important point to make at this point is to take into consideration that Hispanic families which newly arrive in the country are often more concerned with their own and their children's ability to integrate into the community. This, in Canada, heavily includes the need to learn the majority language since government, health, education and other essential services are not available in Spanish, and job opportunities are unavailable without a high proficiency of English. This, in addition to some misguided (although often well-intended) comments from teachers in the school encouraging the use of English in the home to facilitate the children's assimilation to the school and to aid in their continued education leads to pressure on the parents to focus on the learning of English rather than on seeking resources to maintain the use of Spanish in the home. Therefore, it is not uncommon to hear from many families that upon arrival in Canada, accessing Spanish classes for their children was not a main consideration, especially since they considered that they would be able to continue communication in the home in Spanish through their own efforts. One exception was the case of a first generation speaker who was trained as a speech pathologist in Colombia and who considered that maintaining formal education in Spanish essential. Interestingly, however, when her children were interviewed, although they did report attending the Spanish language school for a period of time when they were younger, they expressed not having liked the environment and having found that the class wasn't beneficial to them. In addition to the weekend school, there were also speakers who reported knowing of a church which offers mass in Spanish once a week, but none of our participants reported attending this service. It seems, therefore, that despite some resources being available, most speakers are not particularly interested in accessing them. This may be explained, as noted by

some speakers in the sample anecdotally, by differences between types of Hispanics. There is more that is needed to form a sense of community than a common language and culture. In the case of Colombians, there are sometimes strong pre-existing divisions according to class or other factors and many speakers reported preferring to stay somewhat separated from other Hispanics. Speakers in the sample instead considered that the home was the primary and in many cases the only resource used to teach and cultivate the use of Spanish in their families (section 3.2.4.3). Parents, in this case all first generation speakers, reported having consciously made strong efforts to maintain the language in their children through the use of prayers, children's movies, and books, as well as by encouraging the use of Spanish in family conversations in Canada and with family abroad. This was corroborated by second generation speakers who expressed recognizing their parents' efforts, even to the point of creating inter-family conflict when they were younger when imposing the use of Spanish or more commonly when strongly encouraging literacy development in Spanish at home.

4.2.5 Identification with cultural background (Second generation only)

In regards to minority languages, the connection between ties to and identity with the heritage culture has often been noted to be an important factor favoring language maintenance (Guardado 2002; Pérez-Leroux et al. 2011). In fact, according to Guardado (2002), "emotional ties with the L1 culture is one of the most important factors in L1 maintenance". Guardado goes on to find that in his study considering Hispanic families in Vancouver, second generation Hispanic children who had reported a strong sense of identity in connection to the Spanish language and to their heritage culture were also more successful in maintaining the language.

In the current investigation, when second generation speakers were asked about whether they more closely identified with Colombian or Canadian identity, two main patterns emerged. First, most second generation speakers explained immediately that this was a very difficult question that they have in the past reflected on. Many went on to explain that they did not feel fully Canadian when being with Canadians, and did not feel fully Colombian when being

with other first generation Colombians in Canada, or with Colombians while in Colombia (section 3.2.5.3). Many cited their relationship with language as an obstacle to their sense of identity, explaining that while they are aware that they display many Colombian characteristics, whether culturally or physically, they are more comfortable expressing themselves in English and this leads them to distance themselves from a Hispanic identity. Secondly, most speakers reported ultimately considering that they are more Canadian than Colombian (section 3.2.5.1). This was by no means an absolute for any speaker, and instead many expressed being, for example, 60% or 70% Canadian and 40% or 30% Colombian. This identity however is also not a stable concept and it's rather fluid according to who they are with. That is, while they are with Canadians they sense their lack of compatibility with them in certain aspects such as the home environment and family rules (for instance, in Colombian culture is not uncommon for an adult son or daughter to stay at home well into their 20's and in some cases until marriage), and when they are with Colombians they sense not really identifying with the cultural knowledge or language proficiency of these individuals. Interestingly, another theme that emerged reflects the speakers' considerations of others perspectives of them, which as mentioned above, contribute to the formation of their social identities (section 3.2.5.2). Many speakers expressed having been told by first generation Colombians, even those within their families, that they were "more Canadian than Colombian", and incorporated this into their assessments of their own identity. For instance one second-generation speaker who reported ultimately considering himself more Canadian than Colombian, followed his statement by explaining that although he is proud of his Colombian heritage and internally he feels a strong connection to this background, he know that others would consider him to be "more Canadian" and that his language skills in Spanish are not as strong, which leads him to conclude that he must be "more Canadian". Alternatively, the same concept operated as well in the opposite direction for speakers who had encountered labelling or discrimination, whether directed to themselves or other Hispanics in the community. In at least one case, one speaker explained how she had never felt different from other Canadian children until one day at work she was told by her manager that they

disliked when she spoke in Spanish with Hispanic customers. This, she explained, led her to the “realization” that she was “not fully Canadian”. This is again an example of how different experiences with members of the larger community can influence the formation of the social identity of an individual.

Therefore, based on our qualitative data, we can visualize the Colombian Canadian identity of second generation individuals as multi-faceted, fluid, and generally composed of four components: ties to family and the home culture, social and professional connections, language, and perceptions of themselves by others. In this manner, as individuals get older and have new experiences, their experiences may weaken or reinforce any of these four factors thus enhancing the Colombian or Canadian identity association accordingly. For instance, most speakers mentioned that they considered themselves mostly Canadian and the reasons given could be identified under three themes: I speak English everywhere and with everyone except at home; I don’t look/feel Colombian; and others tell me I’m not Colombian. In contrast, two speakers identified as more Colombian, and similarly they noted: I speak Spanish at home; I don’t think I look/feel Canadian; and I’ve been labeled as an outsider or I have experienced/witnessed discrimination against myself, my family or other Hispanics.

As speakers get older, experiences that change their language abilities via language attrition or that change their use of Spanish, for instance through exogamous relationships which alter the frequency of language use in the home, may contribute to a shift in their identity towards a more Canadian identity. This is anticipated by at least two speakers who mention wanting to use more Spanish at work or to marry a Hispanic individual to ensure their future professional or home environments allow them to maintain their ties to the language and the culture.

4.3 Individual speaker mixed-methods analysis

The last analysis conducted in this investigation is a mixed-method explanatory analysis of individual speaker data (section 3.3). Mixed-method approaches, although not traditionally

used to date in investigations of variable expression of Spanish SPPs, can help not just to better situate research analyses, but as I demonstrate in this section, to find explanations for patterns at the individual speaker level which may not be easily detectable through the separate use of quantitative or qualitative analyses.

When considering the quantitative data resulting from a multivariate analysis of speakers, I determined that most speakers who disfavoured the use of overt SPPs were from the first generation group (see Table 3.12 in section 3.3.1). In fact, only two second generation speakers showed use of overt SPPs in variable contexts which patterned with first generation speakers. In considering the individual answers of these two speakers to questions administered for the qualitative analysis, I can better understand these two speakers and determine in which ways, other than in their frequency of use of Spanish SPPs, they differ from other speakers in their generation group. Specifically, it is interesting to see that these two speakers were also the only two second generation speakers who reported feeling more Colombian than Canadian without much hesitation. This, considering that they had been in Canada since birth or before the age of three, separated them from others in their generation group who usually displayed greater insecurity in answering this question and who ultimately usually concluded that they considered themselves more Canadian than Colombian, due to a number of factors including language use, customs, and lived experiences. In addition, these two speakers are also some of the few speakers who identified a connection to family and culture as a primary motivator to maintain the use of the language, and one of these two speakers were among the very few second generation speakers who expressed wanting to find a partner in the future who is Hispanic or who is at least able to speak in Spanish, as this would help facilitate communication with the family. These attitudes displayed by these two speakers align well with the concept described by Guardado (2008) as *familismo*, which represents a close connection to the family of the heritage individual, and which has been shown to be a factor that is conducive to a higher chance of success in language maintenance efforts for heritage speakers (Guardado 2008; Perez-Leroux et al. 2011). It seems therefore, considering these two speakers, that their

attitudes towards the language, their sense of identity as Colombians, and their motivations for maintaining their use of Spanish may be influencing their use of Spanish at a level similar to that of a first-generation speaker. This contrasts with other studies on heritage languages in contact with English, such as that by Nagy et al. (2015), which did not find any relationship between heritage speakers' ethnic identities and their use of null vs. overt SPPs.

In following this analysis, I also identify two other second generation speakers who pattern similarly to first generation speakers in terms of the frequency of use of overt SPPs and who seem to pattern differently from other second generation speakers who showed higher frequencies of use, despite the fact that the quantitative results still placed them as favouring the use of overt SPPs. These two speakers did not identify as Colombian, but did also express a strong sense of connection to their family and cultural roots. In addition, these are also two of the very few speakers who also report wishing to find a Hispanic romantic partner in the future in order to facilitate communication with their family. The sibling pair explained that after having seen their mother struggle and experience strong frustration while trying to communicate with their non-Hispanic sister-in-law, they had decided to make finding a Hispanic partner a stronger priority. In addition to their connection to family group, these speakers also reported having been involved in multiple programs aimed at improving their use of Spanish. Specifically, they participated in a Hispanic religious church regularly, and took various opportunities to take Spanish language courses in high school and university even when they found themselves to be at a level too advanced for these courses. These experiences may have been instrumental in allowing these speakers to display a use of Spanish which closely patterns to that of first-generation speakers since they increased the quantity and quality (variety of sources) of input they received in Spanish. They would have also been more likely to meet and interact with first generation speakers who were responsible for the youth religious groups, or with other Hispanic youth whose families shared an increased interest in providing their children with more experiences in Spanish.

In contrast, when considering the second generation speakers with some of the highest rates

of use of overt SPPs, our analysis shows that these speakers report a greater disconnect from the Colombian identity as they never felt as fully belonging to the group. This weaker connection to the cultural background does not seem to come from a lack of connection to their family at a personal level, as neither speaker reports conflicts or any sense of alienation from family, rather it seems to have been based on their Spanish language abilities and on instances where they had been labeled as “more Canadian” than Colombian. This kind of labeling is often done by first generation family members who have no ill intention and simply express it jokingly to highlight differences in the second generation’s behaviours or their lesser knowledge of Colombian events, history, or pop culture. Nevertheless, these comments are taken into consideration by these speakers who internalize these views, especially during their adolescent years where they are still gathering information from their social environment and experiences to develop their independent social and personal identities separate from their parents Erikson (1950). In relation to language maintenance, due to the strong relationship drawn by individuals between language and cultural background, a weakened tie between their sense of identity and their heritage cultural background can then lead to decreased motivation for maintaining the language and thus a decreased chance of success in doing so. This is reflected in these speakers as well in the analysis of their social ties and lack of involvement in Hispanic community groups or educational opportunities. Both of these speakers mentioned having few Hispanic friends, communicating mostly in English with their Spanish-English bilingual friends, and having chosen to discontinue Spanish classes as children due to lack of interest in the content.

Therefore, these results suggest that input and frequency of use of Spanish, in combination with other social factors such as family connection, and identification with the heritage culture, are important considerations which may have measurable effects on the use of variable language features such as the variable expression of subject personal pronouns in Spanish.

This mixed-method analysis revealed non-linguistic factors that influence the variable expression of SPPs by identifying patterns regarding speakers’ identity, attitudes, and language use habits, which would have been impossible to detect using only a quantitative analysis. Al-

though rare at the moment, mixed-method analysis should be incorporated into more linguistic analysis research in the future to unveil hidden patterns discernible through more careful qualitative study of social patterns in the community.

Chapter 5

Conclusions

Through this investigation, I sought to add to the existing literature in linguistics research focused on the analysis of the variable expression of Spanish Subject Personal Pronouns (SPPs), by investigating the population of Colombian Spanish speakers living in London, Ontario. Although the variable expression of Spanish SPPs is a well studied phenomenon in both monolingual and bilingual varieties of Spanish, the study of this variable language aspect has not previously been addressed, to my knowledge, within a Canadian context.

The main research questions that I address in this investigation in order to fill the gap in the literature regarding this aspect of Spanish research in contact with English in Canada are: 1) What are the social and linguistic factors that condition the variable expression of Spanish SPPs in the Spanish of Colombians living in London, Ontario; 2) Are there any differences between the first and second generation immigrant speakers of Spanish in regards to their variable use of subject personal pronouns? If so, how do the two generations differ from each other?; and 3) How does the variable expression of SPPs in the Spanish of London, Ontario (within a Canadian context) compare to findings reported for Hispanic communities in the United States? In order to answer these questions, I conducted a quantitative research study using data from 20 sociolinguistic interviews that I collected, forming the CoSLO (Colombian Spanish in London, Ontario) corpus. Interviews in this corpus are composed of ten first generation and ten second

generation speakers from the region of Cundinamarca, Colombia, which allowed me to conduct comparisons across the two generations of speakers.

Through my analysis, I showed that the population of Colombian Spanish speakers in London, Ontario resembles other varieties of Spanish in previous investigations in regards to the main linguistic factors that condition the variable expression of Spanish SPPs. Notably, when considering data for both generations of speakers in the CoSLO corpus, the factors of pronoun type (grammatical person and number), and change of referent (or switch reference) were two of the most significant factors conditioning this variable aspect of Spanish. These results then corroborate previous research in regards to this variable language aspect which affirms that although different varieties of Spanish tend to vary according to their specific frequencies of use of overt SPPs, the factors conditioning this use remain stable from one variety to the next (Silva-Corvalán 1994b; Montrul 2004; Abreu 2009; Orozco 2015).

Although switch reference and pronoun type were significant factors in both generations, the two generations of speakers did differ, when considered separately, in regards to the hierarchy and number of factors that condition their variable expression of Spanish SPPs. Specifically, while first generation speakers' use of overt SPPs was conditioned by pronoun type, followed by change of referent, and length of residence, the use of overt SPPs in second generation speakers was conditioned by pronoun type, followed by verbal mood, interview modality, and polarity of the clause. The changes between the two generations seem to provide evidence of cross-linguistic influence and more specifically of lost sensitivity to discourse-pragmatic factors in the second generation of speakers.

The quantitative analysis shows that second generation speakers have a different hierarchical ranking in regards to the linguistic constraints which condition their use of overt SPPs, where switch reference becomes weaker and where verbal mood becomes a stronger factor. This result aligns with studies of Hispanic communities in the United States (Shin 2014), and may point to evidence of cross-linguistic influence which has resulted in weakened sensitivity of discourse-pragmatic features which condition the use of overt SPPs.

According to Hulk and Müller (2000), during bilingual first language acquisition, bilingual children may be exposed to a wider range of syntactic possibilities than monolingual children, including multiple analyses for grammatical structures. Notably, in some cases where one of the two languages to which the child is exposed provides multiple possible analyses for one grammatical structure while the other language provides evidence of only one analysis for a grammatical structure which is similar at the surface level, the child may receive conflicting linguistic evidence. In this case, the child may choose to take a “short cut”, which alleviates the cognitive load of managing the two grammars, by using elements of one language in the other. That is, the child may more commonly apply to both languages the analysis which is strengthened by evidence from both languages. This strategy can lead, however, to errors in production in the language where two or more analyses are possible, depending on factors which do not operate in both languages. This situation may be at play in the case of the second generation Colombian speakers living in London, Ontario, as speakers favour the analysis for the use of overt SPPs which does not incorporate consideration of discourse-pragmatic factors such as switch-reference. According to Hulk and Müller (2000), this cross-linguistic influence is a process which is more common among children who are concurrently developing two grammars. This would then explain why no effects of language contact were observed in my sample among first-generation speakers who had a fully developed Spanish grammar and who were, therefore, not equally susceptible to cross-linguistic effects. This can be further corroborated when we consider the study by Shin and Smith Cairns (2009), who also found, similarly to the current study, that bilingual second generation speakers showed a higher use of overt SPPs and lower sensitivity to switch reference as a conditioning factor for variable Spanish SPP expression as compared to first generation speakers. The authors suggest that this is due to incomplete acquisition leading to underdeveloped sensitivity to syntax-pragmatic features such as aspects of the switch reference conditioning factor of the variable expression of Spanish SPPs. Importantly, Shin and Smith Cairns (2009) also note that the switch reference factor and sensitivity to the different aspects associated with this factor develop between the

ages of nine and fourteen before reaching adult-like usage. In fact, they showed that speakers in their sample between the ages of nine and fourteen tended to have higher rates of use of overt SPPs than adult first generation speakers and younger children in contexts where there was reference continuity. Specifically the authors propose that children in this age range have not fully developed sensitivity to the need to reduce redundancy in speech by paying attention to a syntax-pragmatic feature such as referent continuity. In the context of the current investigation, this becomes relevant when we consider that all second generation speakers in my sample arrived in Canada before the age of eight. While speakers vary on the level of Spanish development they had achieved at their time of arrival, if we consider the proposal by Shin and Smith Cairns (2009), then all speakers were still in the process of gathering linguistic evidence to develop their sensitivity to syntax-pragmatic factors such as switch reference. In this case then, it is possible that, as explained by Hulk and Müller (2000), for all second generation speakers in the sample the conflicting evidence leads to an underdevelopment of sensitivity to the syntax-pragmatic feature of switch reference as children opted for the syntactic analysis common to both languages, the more frequent use of overt SPPs, for which they would have received more linguistic evidence as they received input from English and Spanish.

Further, as syntax-pragmatic features become weakened due to the higher input from English which instead strengthens a different analysis, speakers may need to compensate for this loss using a different factor to avoid ambiguity since switch reference is a primary factor in helping ensure disambiguation of statements (section 4.1.2), and thus begin to rely more heavily on other available cues like verbal mood. Note as well, that according to Shin and Smith Cairns (2009), speakers develop attention to ambiguity but not redundancy at an earlier age, specifically before the age of nine. Therefore, it makes sense that speakers seek to develop strategies that help to reduce ambiguity but not redundancy in speech. Factors such as verbal mood in my study, or TAM (tense, aspect, mood) in the study by Shin (2014), are both morphosyntactic features which, according to Müller and Hulk (2001) and Sorace (2011), are more resistant to cross-linguistic influence and are therefore more stable. For this reason, the

observed changes in factors between the two generations serve as evidence of language contact effects which influence the variable expression of Spanish SPPs. There is further evidence as well when considering the analysis conducted on the intergenerational differences in regards to the placement of subjects (see section 3.1.5.1).

Analysis of placement of SPPs in my study showed an increased preference for pre-verbal subjects among second generation speakers. This reflects a more rigid SVO order in the second generation, as observed in similar studies Silva-Corvalán (1994a), and also indicates an effect of contact with English. Since English is a [- null subject] language which shows a more rigid SVO, the increased preference of speakers for the pre-verbal position shows deviation from characteristics associated from the NSP and which more closely resemble English grammar. This result then could serve as well as evidence of cross-linguistic influence or convergence between the two grammars of English-Spanish Colombian bilinguals living in London Ontario as they begin to approximate a grammar which deviates further from a [+ null subject] system and closer to a [- null subject] system. This strategy represents a strategy by bilinguals to lighten the cognitive load posed by the management and development of two grammars (Hulk and Muller 2000; Montrul 2004, 2011). That being said, it is more likely to reflect evidence of cross-linguistic influence during bilingual first language development since first generation speakers showed no clear sign of changes in their use of overt SPPs as a function of language contact intensity.

In addition, although my analysis which compared the frequencies of use of overt SPPs between the two generations revealed that there were no significant differences between the two generations, there was a sub-significant trend where second generation speakers still showed a higher use of overt SPPs as compared to first generation speakers. This difference, although not significant, may also provide evidence of erosion of discourse-pragmatic features due to language contact but this needs to be verified through future investigations. In my analysis, the differences between generations was not found to be significant due primarily to the fact that there were speakers in the sample whose average expression of SPPs deviated drastically

from the group means. However, it is important to highlight again among the limitations of the current study that this may have been due to an unintended sampling effect. The inclusion criteria for participants indicated that speakers needed to have a sufficient spoken and understanding proficiency in Spanish in order to be able to participate in the interviews. In my opinion, some second generation speakers may have felt intimidated by this requirement as they consider their level of Spanish insufficient. I base this opinion on anecdotal evidence where potential participants from the community expressed they did not believe their level of Spanish was “good enough” when contacted for recruitment. Despite reassuring them that their level of Spanish was appropriate, some speakers chose not to participate. This may indicate that the sample of second generation speakers in my corpus represents mostly speakers who were confident in their language abilities and who may be among the more proficient Spanish speakers in their generation. It is then possible that the higher frequency of use of overt SPPs, without omitting the outlier speaker, is more characteristic of the population, which would then show a significant increase in overt SPP usage.

A qualitative analysis of the data was also conducted to better inform the quantitative analysis on aspects related to the social environment, personal attitudes and language use habits of speakers. The results reveal that in fact, when we consider the quantitative and qualitative data as it related to individual speakers’ use of Spanish SPPs, important patterns emerge when adopting a mixed-methods explanatory approach to the analysis. Notably, I found that speakers in the second generation whose frequency of use of overt SPPs more closely aligned with that of first generation speakers, were speakers that had strong cultural ties to their Colombian heritage, who reported feeling an affinity to a Colombian identity, and who reported a desire to maintain family relationships as a strong motivator for learning and maintaining their use of the Spanish language. These attitudes seem, in these speakers, to align as well with a higher level of participation in Hispanic community groups, and to a higher desire to ensure their future family groups include the use of Spanish through seeking linguistically endogamous partners. This was further corroborated by quantitative results of the language questionnaire

which showed that speakers which aligned closer with first generation speaker SPP usage patterns, also reported a higher frequency of use of Spanish in their daily life across different contexts (Table 3.13). This contrasts with second generation speakers who most drastically differed from first generation speakers and showed some of the highest frequencies of use of overt Spanish SPPs. These speakers had a closer level of affinity to Canadian culture and identity, and had been uninterested in engaging in Hispanic community groups. Interestingly, this was also supported by quantitative data from the language questionnaire where these speakers reported using Spanish much less frequently than other speakers in their generation (Table 3.13). Therefore, this analysis allowed me to determine that factors such as identity, ties to a Hispanic cultural background, and motivation for learning and maintaining Spanish may in fact relate to measurable variable effects such as the frequency of use of overt expression of Spanish SPPs in bilingual speakers. Future studies investigating variable expression of Spanish SPPs should also consider adopting a mixed-methods approach to further investigate the significance of the relationship identified through this study.

In addition to considerations strictly relating to the variable use of Spanish SPPs, I also used data from the collected sociolinguistic interviews to address whether the attitudes and language use habits of Colombian Spanish speakers in Canada align with known factors that are conducive towards the maintenance of a minority heritage language in a majority context.

In order to answer this question, a number of factors that are known to influence the maintenance of a heritage language were considered including: use of Spanish language in different contexts, attitudes towards intergenerational transmission of the Spanish language, attitudes towards endogamous/exogamous relationships, availability of community or language-learning resources, and ties to Hispanic culture and identity. All of these factors have been identified as significant contributors to the maintenance of a heritage language (Guardado 2008; Pérez-Leroux et al. 2011; Pauwels 2016). Specific questions were designed and included in the interview in order to address each of the factors. After conducting a thematic analysis of these answers, I was able to identify patterns which emerge of factors that condition the ongoing

process of language shift in this community of Hispanic speakers. For instance, in regards to language use, second generation speakers report using Spanish only when in the presence of first generation speakers, despite efforts of the parents of second generation speakers to impose the use of Spanish in the home. In addition, while first generation speakers report using mostly Spanish for most social interaction, the majority of second generation speakers report communicating in English mostly in social situations even when interacting with other heritage Hispanic speakers. This pattern also connects to the question of speakers' attitudes towards endogamous/exogamous marriages.

Most speakers, both in first and second generation groups, report that finding a Hispanic partner was not a primary consideration and other qualities were more important in a partner such as honesty, a hard-working disposition, or a good sense of humor. That being said, some speakers did note that more than finding someone who spoke Spanish, finding someone whose culture was compatible with theirs or who was open to engage in the Colombian culture or traditions was an important consideration to ensure a lasting relationship. This is an important factor since as second generation speakers progress through their adult life and begin to establish their own homes, their use of Spanish in the home will become influenced by the language used with their romantic partners. Considering that language is not an important consideration for most Colombian participants, and that most second generation speakers use mostly English in their social interactions, even when communicating with other Hispanic bilinguals, it is unlikely that the use of Spanish in the home will be maintained for these speakers when they move out from their parents' home. That being said, it is important to highlight that entering into an exogamous marriage does not necessarily mean that a heritage speaker will be unsuccessful in maintaining their language, but that it will be an additional challenge to be overcome.

The attitudes of speakers towards intergenerational transmission will play an important role as well in determining the success of Spanish maintenance in Canada beyond this second generation. Although most speakers confirmed that maintaining the use of Spanish was important to them, many followed this statement by highlighting the difficulty associated with this goal.

For instance, not only is the abundant presence of English and the increased importance of French over Spanish for professional success an obstacle, but so is the lack of availability of Spanish-learning resources or sense of Hispanic community. Most speakers reported not being aware of or having never accessed available resources in the community. In addition, speakers highlighted mainly the importance of French for professional success in Canada, and the strong influence of English in all aspects of life, including the pressure felt by speakers upon arriving to Canada. In fact, several second generation speakers expressed that although they would like to pass on the language to their children, they thought it was unlikely. This was also already being seen in descriptions of third generation family members who were described as having receptive bilingualism abilities at best and rarely if ever were pushed to communicate in Spanish. It is possible that this is partly due as well to the reasons given for the importance of maintaining the language. For many first and second generation speakers, having their children (or future children) learn Spanish is an important consideration in order to develop bilingualism and improve economic opportunities. This goal, however, is more closely tied in Canada to the learning of French than Spanish, and it is likely that as a primary motivator it won't sufficiently encourage Hispanic speakers to prioritize Spanish learning in their homes. Instead, it is likely that individuals who reported motivation to maintain Spanish related to the goal of maintaining cultural and family ties will be more successful in their efforts in the future as this goal cannot be achieved through the acquisition of a different language.

As mentioned previously, only two second generation speakers identified themselves as feeling more Colombian than Canadian. Instead, most second generation speakers considered that their personal identity and behaviours were more closely aligned with being Canadian. Nevertheless, it is important to highlight that most of this sub-population, the children of first generation Hispanic immigrants, declared a sense of confused, divided, and/or dual identity. Most speakers hesitated, laughed, or went on to explain that the question regarding their national identity was one they had reflected on previously and had not really been able to arrive at a conclusive answer. Some speakers explained this as the feeling of not really belonging in any

group, or “floating” in between. These feelings seemed to be tied to experiences of labeling or discrimination from either of their two nationality identity groups. The sense of identity or ties to one’s heritage culture is an essential aspect which can influence language maintenance as it impacts the motivations for maintaining the language. For instance, speakers who feel disconnected from their heritage culture, or who are labeled as outsiders, might begin to internalize an idea of not belonging, and may not see a strong motivation to enforce that identity through further language development. It is therefore my recommendation that in households where heritage language speakers are being raised, parents work to develop an appreciation for the heritage culture and a strong sense of family and community unity.

Although there are many ways to help children feel more connected to their heritage culture, some ways in which this may be achieved by families at home is through nurturing a love for the culture by sharing with them knowledge about various aspects of the heritage country, community and culture. For instance, parents may spend time with children reading in Spanish to develop literacy or informally educating them about the history and cultural traditions of the heritage country and of their family.

Parents may also spend time with the children showing them different foods and different kinds of music in the heritage language, whether it be traditional or popular music. In fact, popular culture can facilitate language maintenance as explained in the study by Guardado (2008). The author explains that speakers who had successfully maintained the use of Spanish in Vancouver were usually individuals who not only had a strong sense of L1 identity, but who enjoyed listening to popular Hispanic music and who had an admiration for Hispanic singers and actors. One of his participants indicated that the recognition that Hispanic popular artists receive currently in North America has been instrumental in encouraging language maintenance as it strengthens a desire in Hispanic children to identify and be part of the culture and the community. Although the study by Guardado (2008) dates back to over a decade ago, I believe this remains relevant today as Hispanic singers, such as *reggaeton* singers, have continued to rise in popularity in recent years. This is reflected by the increasing prevalence

of collaborations between American and Canadian singers with Hispanic artists in numerous popular songs such as *Despacito* (2017) by Luis Fonsi and Daddy Yankee (ft. Justin Bieber), *Mi Gente* (2018) by J. Balvin and Willy William (ft. Beyoncé), and *Mía* (2018) by Bad Bunny (ft. Drake), to name but a few.

Another important way to help children develop a connection with their heritage culture is by including and encouraging participation of the children in cultural traditions in the home and beyond. In the case of the Colombian community, children should be active participants in traditions which could include *la Noche de Velitas* ‘the Night of Candles’¹, *la Novena de Aguinaldos* ‘Christmas Novena’, and other practices such as dancing and singing common traditional songs at family celebrations. In these cases, even if the use of the heritage language is not strictly enforced, the participation in these events will help the bilingual child to develop a sense of connection to the heritage culture as well as a sense of belonging to the community as these items become a more common aspect of their daily lives and thus a part of their identity formation process.

Individuals in the larger community can also contribute to the effort to maintain the heritage language locally by becoming more engaged with the heritage language community. For instance, by simply attending local events or by volunteering time to contribute to the community using any personal skills and/or talents one may have, individuals can help to strengthen the sense of community that is locally felt. More importantly, these simple actions can create more spaces and contexts in which the heritage language can be used, and where individuals may meet and interact with other members from the heritage language community, potentially creating friendships and other relationships. Further, both parents and members of the larger heritage community should strive to ensure that children are not labeled as being different from other members of the heritage community, even when such comments may be done in pass-

¹*La Noche de Velitas* is a holiday observed in Colombia on December 7 every year which celebrates the immaculate conception of Jesus. On this night, individuals light candles and offer a prayer to the Virgin Mary for each candle lit. In Colombia this is done outside and streets are lined with candles and lanterns on this night. This is followed by sharing of food and drinks and a general celebration where neighbours and families come together. This celebration marks the unofficial start of the Christmas season.

ing without ill intentions, and that they are helped in developing a sense of belonging and of national identity connected to their heritage.

Given the qualitative data analysis conducted in consideration of different factors which have been known to contribute towards the maintenance of heritage languages, it is my conclusion that in the community of Colombian Spanish speakers in London, Ontario, the use of Spanish within families is unlikely to be maintained through the third generation of speakers, and instead language shift in as little as two generations is likely to occur. Evidence from the analyses conducted in this investigation indicate that there are already notable language contact effects in second generation speakers which suggest that cross-linguistic influence during language development was present in this generation. This is evident through the intergenerational differences such as the increase in use of pre-verbal subjects, the general tendency to use overt SPPs more frequently, and the complexification of the factors conditioning the use of overt SPPs. However, it is important to highlight that when considering individual speaker variation, I also identify evidence that factors such as frequency of use of the language, driven at times by positive attitudes towards the heritage language, can help in language maintenance efforts and attenuate effects of language contact.

5.0.1 Future research

Some of the limitations of this study should be addressed through further research in order to confirm the results identified through the analyses conducted. For instance, future research should consider collecting data from a different group of first and second generation Hispanic speakers in London, Ontario to identify whether the non-significant change in the rate of variable expression of Spanish SPPs is truly representative of the population of Colombians in the city or whether it is an effect of the specific sample of speakers who participated in the study. In order to avoid running into similar complications whereby second generation speakers who are less confident in their level of Spanish self-exclude, it will be important to ensure that recruitment methods are adjusted. Further, future investigations should consider focusing on an

analysis of second person singular pronouns more closely to gain a better understanding of the differences in use of Spanish SPPs according to specificity. In order to achieve this, however, it would be important to consider using different strategies that allow for better collection of data of Spanish SPPs across all grammatical persons. This may require deviating from the more traditional sociolinguistic interview methodology since this method in the current study resulted in very low token counts for both second person singular and second person plural forms. Finally, in the future, I suggest an investigation of Spanish-French contact in Canada in regards to the variable expression of Spanish SPPs. Since French is also a [- null subject] language, an analysis of the effects of this contact situation within Canada could further confirm that the effects observed in this investigation come as a result of language contact with a [-null subject] language.

References

- Official Languages of Canada Act. 1969. (*r.s.c., 1985, c. 31 (4th supp).*). Retrieved from the Justice of Laws website of the Government of Canada : <https://laws-lois.justice.gc.ca/eng/acts/o-3.01/fulltext.html>.
- Abreu, Laurel. 2009. Spanish subject personal pronoun use by monolinguals, bilinguals and second language learners. *Ph.D. Dissertation. ProQuest Dissertations Publishing. University of Florida.* .
- Abreu, Laurel. 2012. *Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish*, In Selected proceedings of the 14th Hispanic Linguistics Symposium, Eds. Kimberly Geeslin and Manuel Díaz-Campos. Cascadilla Proceedings Project, Somerville, MA, 1–8.
- Adams, Marianne. 1987. From Old French to the theory of pro-drop. *Natural Language and Linguistic Theory* 5(1): 1–32.
- Avila-Jiménez, Barbara. 1995. A sociolinguistic analysis of a change in progress: Pronominal overtiness in Puerto Rican Spanish. *Cornwell Working Papers in Linguistics* 13: 25–47.
- Bad Bunny, ft. Drake. 2018. *Mia*, In X100pre. Producers, DJ Luian and Mambo Kingz. Puerto Rico: Rimas Entertainment.
- Bailey, Guy, Tom Wikle, Jan Tillery, and Lori Sand. 1991. The apparent time construct. *Language Variation and Change* 13(3): 241–264.

- Barrera-Tobón, Carolina and Rocío Raña-Risso. 2016. *A corpus-based sociolinguistic study of contact-induced changes in subject placement in the Spanish of New York City.*, In *Spanish Language and Sociolinguistic Analysis: Issues in Hispanic and Lusophone Linguistics* 8, eds Sandro Sessarego and Fernando Tejedo-Herrero, 323-342. Amsterdam: John Benjamins.
- Bartens, Angela. 2001. The rocky road to education in creole. *Estudios de sociolingüística* 2(2): 27–56.
- Barusch, Amanda, Christina Gringeri, and Molly George. 2011. Rigor in qualitative social work research: Strategies used in published articles. *Social Work Research* 35(1): 11–19.
- Bentivoglio, Paola. 1987. *Los sujetos pronominales de primera persona en el habla de caracas.* Central University of Venezuela. Caracas, Venezuela.
- Bonnici, Lisa M. and Robert Bayley. 2010. Recent research on Latinos in the USA and Canada, part 2: Spanish varieties. *Language and Linguistics Compass* 4(2): 121–134.
- Braun, Virginia and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3: 77–101.
- Brezina, Vaclav. 2018. *Statistics in corpus linguistics.* Cambridge University Press.
- Bullock, Barbara and Almeida Jacqueline Toribio. 2009. Reconsidering Dominican Spanish: Data from the rural Cibao. *Revista Internacional de Lingüística Iberoamericana* 14: 49–73.
- Bylund, Emmanuel. 2009. Maturation constraints and first language attrition. *Language learning* 59(3): 687–715.
- Camacho, Jose. 2008. Syntactic variation: the case of Spanish and Portuguese subjects. *Studies in Hispanic and Lusophone Linguistics* 1(2): 415–433.
- Camacho, José. 2013. *Null subjects.* Cambridge Studies in Linguistics.

- Camacho, José. 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*, eds. Mary A. Kato and Francisco Ordonez. Oxford Scholarship Online. Retrieved 14 May. 2021, from <https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780190465889.001.0001/acprof-9780190465889-chapter-2>.
- Cameron, Richard. 1993. Ambiguous agreement, functional compensation, and nonspecific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language variation and change* 5(3): 305–334.
- Chomsky, Noam. 1982. *Lectures on government and binding*. Dordrecht, Foris.
- Cruz Enríquez, Maura. 2014. El habla de la comunidad cubana en Montreal. *Universidad de Montreal: Memoria de Maestría* .
- Cuza, Alejandro. 2017. *Cuban Spanish dialectology: Variation, contact, and change*. Georgetown University Press, Washington D.C.
- D’Alessandro, Roberta. 2015. *Null subjects*, In *Contemporary Linguistic Parameters*, eds. Antonio Fábregas, Jaume Mateu, and Michael Putnam, 201-226. Bloomsbury Press.
- Dubinski, Kate. 2017. Man in southwestern Ontario charged after family attacked with bat amid shouts of ‘ISIS’. *CBC News*. <https://www.cbc.ca/news/canada/london/family-attacked-bat-yelling-isis-st-thomas-1.4439437> .
- Erikson, Erik H. 1950. *Childhood and society*,. New York: W.W. Norton and Company.
- Fang, Tingting. 2017. How to maintain a minority language through education. *Chinese Studies* 6: 1–11.
- Fishman, Joshua. 1965. Who speaks what language to whom and when? *La linguistique* 2: 67–88.

- Flores-Ferrán, Nydia. 2002. *A sociolinguistic perspective on the use of subject personal pronouns in Spanish narratives of Puerto Ricans in New York City*. Munich: Lincom-Europa.
- Flores-Ferrán, Nydia. 2004. Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? *Language Variation and Change* 16: 49–73.
- Flórez, Luis. 1961. El atlas lingüístico-etnográfico de Colombia (ALEC). Nota informativa. *Thesaurus: Boletín del Instituto Caro y Cuervo* 16: 77–125.
- Fonsi, Luis, Daddy Yankee, and ft. Justin Bieber. 2017. *Despacito (remix)*, In Vida. Producers, Mauricio Rengifo and Andrés Torres. Bogotá, Colombia: Universal Latin / Republic / RBMG / Def Jam / Schoolboy.
- French Language Services Act. 1986. (*r.s.o., 1986, c. f.32*). Retrieved from the e-Laws website of the Government of Ontario: <https://www.ontario.ca/laws/statute/90f32>.
- Gaffield, Chad. 1987. *Language, schooling, and cultural conflict: the origins of the French-language controversy in Ontario*. McGill-Queen's University Press.
- Garcia, Mary Ellen. 2003. Recent research on language maintenance. *Annual Review of Applied Linguistics* 23: 22–43.
- Garrido, Marisol. 2007. Language attitude in Colombian Spanish: Cachacos vs. Costeños. *Lenguaje y literatura* 2(2): 246–257.
- Gridstead, John Allen Ray. 1998. Subjects, sentential negation and imperatives in child Spanish and Catalan. *Doctoral dissertation, University of California, ProQuest Dissertations Publishing*.
- Guardado, Martín. 2002. Loss and maintenance of first language skills: Case studies of Hispanic families in Vancouver. *Canadian Modern Language Review* 58(3): 341–363.

- Guardado, Martín. 2008. Language identity, and cultural awareness in Spanish-speaking families. *Canadian Ethnic Studies Journal* 40(3): 171–181.
- Guardado, Martín. 2010. Heritage language development: Preserving a mythic past or envisioning the future of Canadian identity? *Journal of Language, Identity, and Education*, 9(5): 329–346.
- Gutierrez, Manuel J. 2003. Simplification and innovation in U.S. Spanish. *Multilingua* 22: 169–184.
- Haegeman, Liliane. 1999. *Adult null subjects in non pro-drop languages*, In The acquisition of syntax, eds. Marc-Ariel Friedeman, and Luigi Rizzi. London: Addison, Wesley, and Longman.
- Harrington, Sophie and Ana Teresa Pérez-Leroux. 2016. Subjunctive and subject pronoun realization: A study of “no creo que”. *Borealis-An International Journal of Hispanic Linguistics* 5(1): 87–106.
- Harvie, Dawn. 1998. Null subject in English: Wonder if it exists? *Cahiers linguistiques d'Ottawa* 26: 15–25.
- Haugen, Einar. 1953. *The Norwegian language in America, Vol 1*. Philadelphia: University of Pennsylvania Press.
- Heap, David. 1990. Questions with preverbal subject in Dominican Spanish: Morphological variation and syntactic restructuring. *Journal of the Atlantic Provinces Linguistic Association/Revue de l'Association de Linguistique des Provinces Atlantique* 14: 59–67.
- Henríquez Ureña, Pedro. 1939. Ello. *Revista de filología Hispánica* 1(3): 13–38.
- Hernandez, Angelica and Yvan Rose. 2019. Introducing Phon to the sociophonetics community. *Paper presented at the annual meeting of the Canadian Linguistics Association, Vancouver, BC* .

- Hinzelin, Marc-Olivier and Georg A. Kaiser. 2007. *El pronombre ello en el léxico del español dominicano*, In *Language contact and language change in the Caribbean and beyond*, eds. Wiltrud Mihatsch and Monika Sokol. Frankfurt: Lang.
- Hochberg, Judith G. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62(3): 609–621.
- Hoffman, Michol. 2001. Salvadoran Spanish /-s/ aspiration and deletion in a bilingual context. *University of Pennsylvania Working papers in linguistics* 7(3): 115–181.
- Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36: 533–564.
- Huang, Yan. 1994. *The syntax and pragmatics of anaphora: a study with special reference to chinese*. Cambridge: Cambridge University Press .
- Hulk, Aafke and Natascha Müller. 2000. Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and cognition* 3(3): 227–244.
- Hurtado, Luz Marcela. 2005. *El uso de tú, usted y uno en el español de los colombianos y colombo-americanos*, In *Contactos y contextos lingüísticos*, eds. Luis A. Ortiz López and Manel Lacorte, 187–200. Frankfurt, Madrid: Vervuert Verlagsgesellschaft.
- J. Balvin, Willy William, and ft. Beyoncé. 2017. *Mi gente*, In Released as single. Producers, Willy William and Beyoncé. Universal Music Latin Entertainment, Republic Records, Columbia Records, and Parkwood Entertainment.
- Jarvis, Scott. 2009. *Lexical transfer*, In *The bilingual mental lexicon: Interdisciplinary approaches*, ed. Aneta Pavlenko. Great Britain: Cronwell Press Ltd.
- Johnson, Daniel Ezra. 2009. Getting off the Goldvarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and linguistics compass* 3: 359–383.

- Kaltsa, Maria, Ianthi Maria Tsimpli, and Jason Rothman. 2015. Exploring the source of differences and similarities in L1 attrition and heritage speaker competence: Evidence from pronominal resolution. *Lingua* 164(B): 266–288.
- Kuzelewska, Elzbieta. 2016. Language policy in Switzerland. *Studies in logic, grammar and rhetoric* 45(58): 125–140.
- Labov, William. 1972. *Sociolinguistic patterns*. University of Pennsylvania Press,.
- Labov, William. 1994. *Principles of linguistic change. Volume 1: Internal factors*. Oxford: Blackwell Publishers.
- Lane, Pia. 2010. “we did what we thought was best for our children” : A nexus analysis of language shift in a Kven community. *International Journal of the Sociology of Language* 202: 63–78.
- Lastra, Yolanda and Pedro Martín Butragueño. 2015. *Subject pronoun expression in oral Mexican Spanish*, In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*, eds. Ana M. Carrvalho, and Rafael Orozco, and Naomi Lapidus Shin. Washington, D.C.: Georgetown University Press.
- Lavandera, Beatriz R. 1978. Where does the sociolinguistic variable stop? *Language in society* 7: 171–182.
- Linton, April. 2004. A critical mass model of bilingualism among U.S.-born Hispanics. *Social Forces* 83(1): 279–314.
- Lipski, John M. 1989. Spanish pro-drop meets the bilingual speaker. *Paper presented at the conference of New Ways of Analyzing Variation in English* : 1–26.
- Lipski, John M. 2002. Subject pronoun usage among Spanish dialects. *Pennsylvania State University* : 1–27.

- Lozano, Rosina. 2018. *An American language: The history of Spanish in the United States*. University of California Press.
- Manzoor, Amir. 2016. *Designs of mixed methods research*, In *Mixed Methods Research for Improved Scientific Study*, eds. Mette Lise Baran, and Janeth Elisabeth Jones, 79-105. IGI Global.
- Mayo, Laia and Robin Clark. 2010. Pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics* 42(3): 781–799.
- McWhorter, John. 1995. The scarcity of Spanish-based creoles explained. *Language in Society* 24(2): 213–244.
- Montes Giraldo, José Joaquín. 1982. El español de Colombia. propuesta de clasificación dialectal. *Thesaurus* XXXVII(1): 23–92.
- Montrul, Silvina. 2004. Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*. 7(2): 125–142.
- Montrul, Silvina. 2011. Morphological errors in Spanish second language speakers and heritage speakers. *Studies in Second Language Acquisition* 33.
- Montrul, Silvina. 2012. Is the heritage language like a second language? *Eurosla* 12.
- Montrul, Silvina and Tania Ionin. 2012. Dominant language transfer in Spanish heritage speakers and L2 learners in the interpretation of definite articles. *Modern Language Journal* 96(1): 70–94.
- Mougeon, Raymond, Édouard Beniak, Cora Brent-Palmer, Monique Belanger, and Walter Cichocki. 1982. *Le français parlé en situation minoritaire*, vol. 1. Québec City: International Center for Research on Bilingualism.
- Muysken, Pieter. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16(4): 709–730.

- Muysken, Pieter and Rene Appel. 1987. *Language contact and bilingualism*. Amsterdam University Press.
- Müller, Natascha. 2019. Null subjects and null objects in monolingual and bilingual children: a commentary. *Probus* 31(2): 279–298.
- Müller, Natascha and Aafke Hulk. 2001. Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and cognition* 4(1): 1–21.
- Nagy, Naomi. 2015. A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua* 164(B): 309–327.
- Nagy, Naomi, Nina Aghdasi, Derek Denis, and Alexandra Motut. 2015. Null subjects in heritage languages: Contact effects in a cross-linguistic context. *Selected papers from NWAV 39*, 17(2): 135–144.
- Oakland, Ross. 2003. Adios city of Anglos: Welcome to Londombia; London, Ontario draws Latin vibe 'the new wave is Colombian'. *Toronto Star*.
- Ordóñez, Francisco and Esthela Treviño. 1999. Left dislocated subjects and the pro-drop parameter: A case study of Spanish. *Lingua* 107: 39–68.
- Orozco, Rafael. 2004. *A sociolinguistic study of Colombian spanish in Colombia and New York City*. Doctoral Dissertation, New York University.
- Orozco, Rafael. 2015. *Subject pronoun expression in Spanish: A cross-dialectal perspective*, In *Pronominal Variation in Colombian Costeño Spanish*, eds. Ana M.Carvalho, Rafael Orozco, and Naomi Lapidus Shin, 171-181. Washington, DC: Georgetown University Press.
- Orozco, Rafael and Gregory R. Guy. 2008. *El uso variable de los pronombres sujetos: ¿Qué pasa en la costa caribe colombiana?*, In *Selected Proceedings of the 4th Workshop on Span-*

- ish Sociolinguistics, eds. Maurice Westmoreland and Juan Antonio Thomas. 3. Cascadilla Proceedings Project. Somerville, MA, 70–80.
- Otheguy, Ricardo and Nancy Stern. 2011. On so-called Spanglish. *International Journal of Bilingualism* 15(1): 770–802.
- Otheguy, Ricardo and Ana C. Zentella. 2012. *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford University Press.
- Otheguy, Ricardo, Ana C. Zentella, and David Livert. 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 83(4): 770–802.
- Paradis, Johanne and Samuel Navarro. 2003. Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of input? *Journal of Child Language* 30(2): 371–393.
- Pato, Enrique. 2013-2017. *El español de Montréal - COLEM (Corpus Oral de la Lengua Española en Montréal)*. Montréal: Université de Montréal.
- Pauwels, Ana. 2016. *Language maintenance and shift*. Cambridge University Press.
- Penny, Ralph. 2004. *Variation and change in Spanish*. UK: Cambridge University Press.
- Peskova, Andrea. 2013. Experimenting with pro-drop in Spanish. *SKY Journal of Linguistics* 26: 117–149.
- Plyley, Christopher and Angelica Hernandez. 2019. Linguistic barriers to STEM education of creole speakers in St. Croix. *Paper presented at AACU: Transforming STEM Higher Education, Chicago*.
- Poplack, Shana. 1980. Deletion and disambiguation in Puerto Rican Spanish. *Linguistics Society of America* 56: 371–385.

- Poplack, Shana. 1988. Conséquences linguistiques du contact de langues: Un modèle d'analyse variationniste. *Language et Societé* 43: 23–48.
- Potowski, Kim. 2014a. *Ethnolinguistic identities and ideologies among Mexicans, PRicans, and Mexiricans in Chicago*, In *A Sociolinguistic of Diaspora: Latino practices, identities and ideologies*, eds. Rosina Márquez Reiter, and Luisa Martín Rojo. New York: Routledge.
- Potowski, Kim. 2014b. *Handbook of Spanish second language acquisition*, ed. kimberly l. geeslin, In *Heritage learners of Spanish*. Hoboken: John Wiley Sons, Inc.
- Pérez-Leroux, Ana T., Alejandro Cuza, and Danielle Thomas. 2011. *From parental attitudes to input conditions: A look at Spanish-English bilingual development in Toronto.*, In *Bilingual youth: Spanish in English-speaking societies*, eds. Kim Potowski and Jason Rothman. Amsterdam: John Benjamins Publishing Company.
- Ranson, Diana L. 1991. Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Change and Variation* 3(2): 133–152.
- Rickford, John R. 1998. *The creole origins of African American Vernacular English: Evidence from copula absence*, In *African American English*, ed. Salikoko S. Mufwene, John R. Rickford, Guy Bailey and John Baugh. London: Routledge.
- Roberts, Ian and Anders Holmberg. 2010. *Introduction: parameters in minimalist theory*, In *Parametric Variation: Null subjects in minimalist theory*, eds. T. Biberauer, A. Holmberg, I. Roberts, and M. Sheehan. Cambridge: Cambridge University Press.
- Rose, Yvan, Brian MacWhinney, Rodrigue Byrne, Gregory Hedlund, Keith Maddocks, Philip O'Brien, and Todd Wareham. 2006. *Introducing Phon: A software solution for the study of phonological acquisition*. Proceedings of the 30th Annual Boston University Conference on Language Development. Somerville, MA: Cascadilla Press.
- RStudio Team. 2020. *Rstudio: Integrated development for R*. Boston, MA: RStudio.

- Sankoff, David and Shana Poplack. 1981. A formal grammar for code-switching. *Papers in linguistics: International journal of human communication* 14: 3–46.
- Sankoff, Gillian. 2001. *Linguistic outcomes of language contact*, In Handbook of Sociolinguistics, eds. Peter Trudgill J, Jack Chambers, and Natalie Schilling-Estes. Oxford: Basil Blackwell.
- Schechter, Sandra R. and Robert Bayley. 1997. Language socialization practices and cultural identity: Case studies of Mexican-descent families in California and Texas. *TESOL Quarterly* 31(3): 513–541.
- Sessarego and Gutierrez-Rexach. 2017. Revisiting the null subject parameter : New insights from afro-peruvian spanish. *Isogloss* 3(1): 43–68.
- Shin, Naomi Lapidus. 2014. Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change* 26(3): 303–330.
- Shin, Naomi Lapidus and Rafael Orozco. 2013. Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. *Language in Society* 42: 429–452.
- Shin, Naomi Lapidus and Helen Smith Cairns. 2009. *Subject pronouns in child Spanish and continuity of reference.*, In Selected Proceedings of the 11th Hispanic Linguistics Symposium, eds. Joseph Collentine, Maryellen García, Barbara Lafford, and Francisco Marcos Marín. 155-164. Somerville, MA: Cascadilla Proceedings Project.
- Shuy, Roger W. 2001. *In person vs. telephone interviewing*, In Handbook of interview research, eds. Jaber F. Gubrium, and James A. Holstein. SAGE Publications, Inc.
- Siegel, Jeff. 2007. Creoles and minority dialects in education: An update. *Language and education* 21(1): 66–86.

- Silva-Corvalán, Carmen. 1994a. The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change* 6: 255–272.
- Silva-Corvalán, Carmen. 1994b. *Language contact and change: Spanish in Los Angeles*. Oxford University Press.
- Silva-Corvalán, Carmen. 2008. The limits of convergence in language contact. *Journal of Language Contact* 2(1): 213–224.
- Sorace, Antonella. 2011. Pinning down the concept of “interface” in bilinguals. *Linguistic approaches to bilingualism* 1(1): 1–33.
- Statistics Canada. 2017. Census profile. 2016 census. *Statistics Canada Catalogue. No. 98-316-X2016001. Ottawa. Released October 25, 2017.* .
- Talmy, Steven. 2010. Qualitative interviews in applied linguistics: From research instrument to social practice. *Annual Review of Applied Linguistics* 30(1): 128–148.
- Tennant, Jeff, David Heap, Angelica Hernandez, and Jorge Rosés Labrada. 2015. Aspiration vs. deletion of /-s/ in contemporary Eastern Cuban Spanish: Differing constraints. *Presented at New Ways of Analyzing Variation* 44 .
- Thomason, Sarah G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G. and Terrence Kaufman. 2001. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Toribio, Almeida Jacqueline. 2000. Setting parametric limits on dialectal variation in Spanish. *Lingua* 110: 315–341.
- Torres, Lourdes. 2018. Latinx? *Latino Studies* 16: 283–285.

- Torres Cacoullos, Renna and Catherine E. Travis. 2010. Testing convergence via code-switching: priming and the structure of variable subject expression. *International Journal of Bilingualism* 15: 241–267.
- Torres Cacoullos, Renna and Catherine E. Travis. 2015. *Foundations for the study of subject pronoun expression in Spanish in contact with English: Assessing interlinguistic (dis)similarity via intralinguistic variability*, In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*, eds, Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin, 83-102. Washington, DC: Georgetown University Press.
- Travis, Catherine E. 2007. Genre effects on subject expression on Spanish: Priming in narrative and conversation. *Language Variation and Change* 19: 101–135.
- Treffers-Daller and Raymond Mougeon. 2005. The role of transfer in language variation and change: Evidence from contact varieties of French. *Bilingualism: Language and Cognition* 8(2): 93–98.
- Trudgill, Peter. 1974. *Sociolinguistics: An introduction to language and society*. Penguin, UK.
- Truman, Lauren. 2019. *Language attitudes of bilingual Maya women living in Yucatan.*, In *Contact, community and connections: Current approaches to Spanish in multilingual populations*, eds. Gregory L. Thompson and Scott M. Alvord. Malaga, Spain: Vernon Press, 35–58.
- Turan, Ümit Deniz. 1996. Null vs. overt subjects in Turkish discourse: A centering analysis. *Doctoral dissertation, Institute for Research in Cognitive Science Technical Reports Series. University of Pennsylvania* 95: 1–174.
- Unsworth, Sharon. 2013. Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics* 33: 21–50.

- Valdman, Albert. 1989. *The use of creole as a school medium and decreolization in Haiti*, In Literacy in school and society : Topics in language and linguistics. Boston, MA: Springer.
- Valdés, Guadalupe. 2005. Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *Modern Language Journal* 89: 410–426.
- Wei, Li. 2000. *Dimensions of bilingualism*, In *The Bilingualism Reader*. New York, NY: Routledge.
- Weinreich, Uriel. 1968. *Languages in contact: Findings and problems*. The Hague: Mouton.
- Wigglesworth, Gillian, Rosey Bilington, and Deborah Loakes. 2013. Creole speakers and standard language education. *Language and Linguistics Compass* 7(7): 388–397.
- Wurm, Steven. 2002. *Strategies for language maintenance and revival*, In *Language endangerment and language maintenance: an active approach*, eds. David Bradley, and Maya Bradley. Routledge. New York.

Appendix A

Western Research Ethics Board Approval

This Appendix includes the approval notices from the Western Research Ethics board, as well as the letters with the letters showing approval for continuation of research procedures. The first letter included, indicates that all of the initial study procedures were approved. The second letter represents the approval to the amendment submitted later to authorize the distance interviews and the participants' necessary verbal consent in these cases. The last two letters show the approval notices for the applications submitted to renew ethics approval to continue research.



Date: 5 December 2018

To: David Heap

Project ID: 112793

Study Title: London Ontario Colombian Spanish Corpus.

Application Type: NMREB Initial Application

Review Type: Delegated

Full Board Reporting Date: January 11 2019

Date Approval Issued: 05/Dec/2018

REB Approval Expiry Date: 05/Dec/2019

Dear David Heap

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the above mentioned study, as of the date noted above. NMREB approval for this study remains valid until the expiry date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

This research study is to be conducted by the investigator noted above. All other required institutional approvals must also be obtained prior to the conduct of the study.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
English - Interview guide - LOCS	Interview Guide	23/Nov/2018	2
Facebook Ad_Spanish and English_LOCS	Recruitment Materials	08/Nov/2018	1
Letter of Information and Consent - English - LOCS	Written Consent/Assent	23/Nov/2018	2
Letter of Information and Consent - Spanish - LOCS	Written Consent/Assent	23/Nov/2018	2
Newspaper Ad_Spanish and English_LOCS	Recruitment Materials	08/Nov/2018	1
questionnaire language use - English -LOCS	Paper Survey	08/Nov/2018	
questionnaire language use - Spanish - LOCS	Paper Survey	08/Nov/2018	1
Recruitment poster - English and Spanish - LOCS	Recruitment Materials	23/Oct/2018	1
Script for in-person recruitment - Spanish and English - LOCS	Oral Script	23/Nov/2018	2
Spanish - Interview guide - LOCS	Interview Guide	23/Nov/2018	2
Translation Attestation - Heap - REB112793 - London Ontario Colombian Spanish (LOCS) 20181127	Additional Consent Documents	27/Nov/2018	1
Video Excerpt from "Betty la Fea"	Other Data Collection Instruments	23/Oct/2018	1

No deviations from, or changes to the protocol should be initiated without prior written approval from the NMREB, except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Kelly Patterson, Research Ethics Officer on behalf of Dr. Randal Graham, NMREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).



Western Research

Date: 8 April 2019

To: David Heap

Project ID: 112793

Study Title: London Ontario Colombian Spanish Corpus.

Application Type: NMREB Amendment Form

Review Type: Delegated

Full Board Reporting Date: May 3 2019

Date Approval Issued: 08/Apr/2019

REB Approval Expiry Date: 05/Dec/2019

Dear David Heap,

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the amendment, as of the date noted above.

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Language Questionnaire - Online - English - LOCS	Online Survey	04/Apr/2019	1
Language Questionnaire - Online - Spanish - LOCS	Online Survey	04/Apr/2019	1
Letter of Information and Consent - English - LOCS - Phone&Online (Apr 4)	Verbal Consent/Assent	04/Apr/2019	2
Letter of Information and Consent - Spanish - LOCS - Phone&Online (Apr 4)	Verbal Consent/Assent	04/Apr/2019	2

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Kelly Patterson, Research Ethics Officer on behalf of Dr. Randal Graham, NMREB Chair

Note: *This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).*



Western Research

Date: 6 December 2019

To: David Heap

Project ID: 112793

Study Title: London Ontario Colombian Spanish Corpus.

Application Type: Continuing Ethics Review (CER) Form

Review Type: Delegated

Meeting Date: 10/Jan/2020

Date Approval Issued: 06/Dec/2019

REB Approval Expiry Date: 05/Dec/2020

Dear David Heap,

The Western University Non-Medical Research Ethics Board has reviewed this application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Daniel Wyzynski, Research Ethics Coordinator, on behalf of Prof. Randal Graham, NMREB Chair

Note: *This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).*



Western Research

Date: 2 December 2020

To: David Heap

Project ID: 112793

Study Title: London Ontario Colombian Spanish Corpus.

Application Type: Continuing Ethics Review (CER) Form

Review Type: Delegated

Date Approval Issued: 02/Dec/2020

REB Approval Expiry Date: 05/Dec/2021

Dear David Heap,

The Western University Non-Medical Research Ethics Board has reviewed this application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

The Office of Human Research Ethics

Note: *This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).*

Appendix B

Interview Guides

This appendix includes the guide for the semidirected sociolinguistic interviews conducted for the current project investigation. Following the Spanish guide a translated document in English is also provided.

Questions in these guides were designed to elicit a natural, informal conversation which might include story-telling (family history), emotional responses (soccer events), and descriptions (friendship qualities that are important). There is a second section of the interview, titled "Life in Canada". In this section the questions are designed to gain further information about the participants' experiences as immigrants in Canada. This section includes two versions. One for first generation immigrants, and one for second generation immigrants. The experiences and concerns of the two generations are very different and these questions were designed with these distinctions in mind. Finally, the third section of the interview, titled *Idioma* (Language) was introduced to collect information on participants' attitudes towards their own language, its use and maintenance in Canada, and their own bilingualism.

Londombia – Guía para la entrevista semidirigida

Información personal:

1. ¿Cómo se llama?
2. ¿Cuántos años tiene?
3. ¿En dónde vive?
4. ¿Hace cuántos años que vive en London?
5. ¿De qué parte de Colombia es usted? ¿Y su familia?
6. ¿A qué se dedica?

Familia

7. Cuénteme un poco acerca de su familia.
 - a. ¿Como se llaman sus padres?
 - i. ¿Puede describirlos un poco?
 - ii. ¿Cómo se conocieron sus padres?
 - b. ¿Tiene hijos?
 - i. ¿Cuántos hijos tiene?
 - ii. ¿Y cómo son ellos?
 - iii. ¿Cómo eran cuando chiquitos?
 - c. ¿Tiene hermanos?
 - i. ¿Cuántos hermanos tiene?
 - ii. ¿Puede describirlos un poco?
8. ¿Su familia está aquí en Canadá también?
9. ¿Tiene todavía mucha familia en Colombia?

Amor y amistad

10. ¿Está casado(a)?
 - a. ¿Cómo es su pareja?
 - i. ¿Qué lo enamoró de esta persona?
 - ii. ¿Cómo conoció a su pareja?
11. ¿Qué cualidades busca usted en una pareja?
 - a. ¿Es importante para usted que su pareja sea hispana(o), o eso no importa?
 - b. ¿Cree que hay dificultades en casarse con alguien que no es hispano/latino? ¿Cuáles?
12. ¿Tiene muchos amigos?
 - a. ¿Cómo son sus amigos?
 - b. ¿Qué es importante buscar en un amigo?
 - c. ¿Sus amigos son hispanos?
 - i. ¿Habla con ellos en español?
 1. ¿Aproximadamente que porcentaje del tiempo?
 - d. ¿De dónde son sus amigos?
 - e. ¿Qué le gusta hacer cuando esta con sus amigos?

Hobbies

13. ¿Qué le gusta hacer en su tiempo libre o en los fines de semana?

14. ¿Qué tipo de comida es el que más le gusta?
15. ¿Qué tipo de música prefiere?
 - a. ¿Quién es su cantante o artista favorito? ¿Por qué?
 - b. ¿Le gusta el reggaetón?
16. ¿Le gustan los deportes?
 - a. ¿Qué deporte es su favorito?
 - b. ¿Practica algún deporte?
17. ¿Le gusta el fútbol?
 - a. ¿Vio el mundial de este año?
 - i. ¿Con quién lo vio?
 - ii. ¿Cómo eran las reuniones para ver el mundial? Descríbamelas un poco.
 - iii. ¿Qué pensó de la selección de este año?
 - b. ¿Qué pensó del partido contra Inglaterra en el que eliminaron a Colombia?
 - i. ¿Qué pensó del árbitro?
 1. ¿Su familia y amigos están de acuerdo con usted?
 - ii. ¿Piensa que James debió haber jugado?
 1. ¿Su familia y amigos están de acuerdo con usted?
 - c. ¿Alguna vez ha visto a Colombia jugar en el mundial en persona?
 - i. ¿Le gustaría ir a un mundial a ver a Colombia jugar en un futuro?
18. ¿Qué tipo de televisión le gusta ver?
 - a. ¿Cuál es su programa de televisión favorito?
 - i. ¿De qué se trata?
 - b. ¿Le gusta ver televisión en español?
19. ¿Le gusta ver novelas?
 - a. ¿Está viendo alguna ahora?
 - i. ¿De qué se trata la novela?
 - b. ¿Cuál ha sido la mejor novela colombiana que ha visto en su opinión?
 - i. ¿De qué se trataba la novela?

Life in Canada (Questions for 1st generation speakers)

20. ¿En qué año llegó a Canadá?
 - a. ¿Cómo fueron los primeros días en Canadá?
 - b. ¿Fue duro el ajustarse a la vida en Canadá?
 - i. ¿Qué dificultades encontró?
 - c. ¿A quién en su familia le dio más duro el ajustarse a vivir en Canadá?
 - d. A algunas personas les gusta el invierno, y a muchas otras les parece de los más difícil de vivir en Canadá. ¿Qué piensa usted?
21. ¿Le gusta vivir en Canadá?
 - a. ¿Hay aspectos de la vida aquí que usted cambiaría?
22. ¿Qué fiestas celebra aquí en familia? ¿Cuál(es) son las más importantes?
 - a. ¿Qué se hace en esas fiestas?
 - b. ¿Qué tipo de comida se encuentra en esas celebraciones?
 - i. ¿Qué tipo de comida se encontraba en las celebraciones en Colombia?
 - c. ¿Ha cambiado la manera como celebra fiestas?

23. ¿Extraña a Colombia?
a. ¿Vuelve a menudo? ¿Por qué o por qué no?
24. ¿Planea regresar a Colombia algún día?

Life in Canada/Connection with Colombia (Questions for 2nd generation speakers)

25. ¿Quién fue la primera persona en su familia en llegar a Canadá?
a. ¿Sabe cómo fue esa experiencia de llegar a Canadá para ellos?
26. ¿Ha ido a Colombia?
a. ¿Le gustaría ir? ¿Por qué o por qué no?
27. ¿Qué cree que son las mayores diferencias entre vivir en Canadá y vivir en Colombia?
a. ¿Hay aspectos de la vida aquí que usted cambiaría?
28. ¿Tiene una relación cercana con familia y amigos en Colombia?
a. ¿Con qué frecuencia habla usted con ellos en Colombia?
29. ¿Se siente más colombiano o canadiense? ¿Por qué?
30. ¿Qué es lo que más se le dificulta de ser colombo-canadiense?
31. ¿Qué es lo que más le gusta de ser colombo-canadiense?

Idioma

32. ¿Usa más el español o el inglés en la casa?
a. ¿Con sus padres?
b. ¿Con sus hermanos?
c. ¿Con sus hijos?
33. ¿Cree que es importante que los hijos de hispanos en Canadá aprendan a hablar bien el español?
a. ¿Por qué o por qué no?
b. ¿Es importante que lo sepan leer y escribir?
c. ¿Hay oportunidades para que los niños hispanos aprendan y practiquen el español?
34. Algunas personas dicen que prefieren hablar en inglés en lugares públicos. ¿Usted qué opina?
35. ¿Utiliza usted español en el trabajo?
a. ¿Le ha servido el español en el trabajo?
36. ¿Habla usted otros idiomas?

Londombia – Guide for the semi-structured interview

Personal Information

1. What is your name?
2. How old are you?
3. Where do you live?
4. For how long have you lived in London, Ontario?
5. What part of Colombia are you from? And your family?
6. What is your occupation?

Family

7. Tell me a little bit about your family.
 - a. What are the names of your parents?
 - i. Can you describe them a bit?
 - ii. How did your parents meet?
 - b. Do you have any children?
 - i. How many children do you have?
 - ii. What are they like?
 - iii. What were they like when they were little?
 - c. Do you have any siblings?
 - i. How many siblings do you have?
 - ii. Can you describe them a bit?
8. Is your family here in Canada too?
9. Do you still have a lot of family in Colombia?

Love and Friendship

10. Are you married?
 - a. What is your partner like?
 - i. What made you fall in love with them?
 - ii. How did you meet them?
11. What qualities do you look for in a partner?
 - a. Is it important for you that your partner be Hispanic, or that doesn't matter?
 - b. Do you think that there are difficulties in marrying someone that is not Hispanic like you? Which ones?
12. Do you have a lot of friends?
 - a. What are your friends like?
 - b. What qualities are important in a friend?
 - c. Are your friends Hispanic?
 - i. Do you speak with them in Spanish?
 1. Approximately what percentage of the time?
 - d. Where are your friends from?
 - e. What do you like to do when you spend time with your friends?

Hobbies

13. What do you like to do in your free time?
14. What is your favourite type of food?
15. What is your favourite type of music?
 - a. Who is your favourite singer? Why?
 - b. Do you like reggaeton?
16. Do you like sports?
 - a. What is your favourite sport?
 - b. Do you play any sports?
17. Do you like soccer?
 - a. Did you watch the world cup this year?
 - i. Who did you watch it with?
 - ii. What were the get togethers to watch the world cup like? Describe them a bit.
 - iii. What did you think of the Colombian team this year?
 - b. What did you think about the game against England where Colombia was eliminated?
 - i. What did you think about the referee in that game?
 1. Do your family and friends agree with you on that?
 - ii. Do you think James should have played in the game?
 1. Do your family and friends agree with you on that?
 - c. Have you ever seen Colombia play in a world cup in person?
 - i. Would you like to go to a world cup to watch Colombia play someday?
18. What kind of TV shows do you like to watch?
 - a. What is your favourite TV show?
 - i. What is it about?
 - b. Do you like to watch TV in Spanish?
19. Do you like to watch novelas?
 - a. Are you watching one now?
 - i. What is the novela about?
 - b. What is, in your opinion, the best Colombian novela that you have watched?
 - i. What was that novela about?

Life in Canada (Questions for 1st generation speakers)

20. In what year did you arrive to Canada?
 - a. What were those first days in Canada like?
 - b. Was it hard to adjust to life in Canada?
 - i. What difficulties did you find?
 - c. Who do you think, in your family, had the hardest time adjusting to life in Canada?
 - d. Some people enjoy the winter, but for many others it is the hardest part of living in Canada. What do you think?
21. Do you like living in Canada
 - a. Are there aspects of life in Canada that you would change if you could?
22. What festivities do you celebrate with your family?
 - a. Which ones are the most important?
 - b. What do you do in these celebrations?

- c. What type of food do you usually find in these celebrations?
 - i. What type of food did you find in these celebrations in Colombia?
 - d. Have you changed the way you celebrate these festivities?
23. Do you miss Colombia?
- a. Do you go back often? Why or why not?
24. Do you plan on returning to Colombia someday?

Life in Canada/Connection with Colombia (Questions for 2nd generation speakers)

25. Who was the first person in your family to come to Canada?
- a. Do you know what their experiences were like?
26. Have you been to Colombia?
- a. Would you like to go? Why or why not?
27. What do you think are some of the major differences between living in Canada and living in Colombia?
- a. Are there any aspects of life here that you would change if you could?
28. Do you keep a close rapport with family and friends in Colombia?
- a. How often do you speak with them?
29. Do you feel more Colombian or Canadian? Why?
30. What is the hardest part of being Colombian-Canadian?
31. What is the best part of being Colombian-Canadian?

Language

32. Do you use more Spanish or English at home?
- a. With your parents?
 - b. With your siblings?
 - c. With your children?
33. Do you think it's important that the children of Hispanic families in Canada learn Spanish?
- a. Why or why not?
 - b. Is it important that they know how to read it and write it as well?
 - c. Are there opportunities for children to learn and practice Spanish here?
34. Some people say that they prefer to speak in English in public places. What is your opinion?
35. Do you use Spanish at work?
- a. Has knowing Spanish helped you in your work?
36. Do you speak any other languages?


Appendix C

Language Questionnaires


This Appendix, includes the language questionnaire that was administered to all participants. The same questionnaire was given to participants regardless of whether they participated in the interviews in person, or at distance through phone call or through Skype. The only difference in the case of the speakers who chose to participate at distance was that the quiz was administered via a SurveyMonkey.com link.

Questionnaires were available in English and in Spanish (both included below), and participants were given the choice to complete the questionnaire in either language. The majority of participants chose to complete the questionnaire in Spanish.

Language questionnaire


1. Please select the options that describe your language abilities in Spanish and English most accurately: 

	A lot of difficulty	Some difficulty	A little difficulty	No difficulty
Understanding Spanish I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking Spanish I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading Spanish I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Writing Spanish I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Understanding English I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Speaking English I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reading English I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Writing English I have...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>


2. Please select the answer that most accurately describes your use of Spanish on a daily basis. 

	Always (90-100%)	Frequently (60-89%)	About half the time (40-59%)	Sometimes (10-39%)	Never (0-9%)	Not applicable
I speak Spanish to my partner at home.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My partner at home speaks Spanish to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish to my children.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My children speak Spanish to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My children speak Spanish with my partner.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish with my parents (I live with my parents).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish with my parents (I don't live with my parents).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish with someone else at home.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I use Spanish during family get-togethers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I use Spanish at school/at work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish with my hispanic friends.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish with my bilingual (english-spanish bilingual) friends.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish with my family abroad.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I speak Spanish with my friends abroad.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comments

3. Please state in which country (or countries) are your family and friends who live abroad. 


DONE

Powered by
 SurveyMonkey

Cuestionario de idioma


1. Por favor seleccione las opciones que mejor describan sus habilidades en Español y en Inglés: 

	Mucha dificultad	Algo de dificultad	Muy poca dificultad	Ninguna dificultad
Entendiendo español tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hablando español tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Leyendo en español tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Escribiendo en español tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Entendiendo inglés tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hablando inglés tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Leyendo inglés tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Escribiendo inglés tengo...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>


2. Por favor seleccione la opción más adecuada para Usted en cada situación: 

	Siempre (90-100%)	Frecuentemente (60-89%)	Aproximadamente la mitad del tiempo (40-59%)	Aveces (10-39%)	Nunca (0-9%)	No me aplica
Yo le hablo en español a mi pareja en la casa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mi pareja me habla en español en la casa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo le hablo a mis hijo.a.s en español en la casa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mis hijo.a.s me hablan en español en la casa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mis hijos le hablan en español con mi pareja en la casa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo con mis padres en español en la casa. (vivo con mis padres)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo con mis padres en español. (no vivo con mis padres)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo español con alguien mas en mi casa.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo en español durante reuniones familiares.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo español en el trabajo/college/universidad.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo en español con mis amigos hispanos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo en español con mis amigos bilingües (español-inglés).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo en español con mi familia en el exterior.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Yo hablo en español con mis amigos en el exterior.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comments

3. Por favor mencione en qué país(es) viven su familia y amigos que están en el exterior? 

FINALIZAR

Powered by
 SurveyMonkey

Curriculum Vitae

Name: Angélica Hernández

Post-Secondary Education and Degrees: The University of Western Ontario
Ph.D. French Studies (Specialization in Linguistics)
2016 - 2021

The University of Western Ontario
M.A. French Studies (Specialization in Linguistics)
2014 - 2015

The University of Western Ontario
Diplôme de Français Pratique
2013 - 2014

The University of Western Ontario
B.Sc. Honors Specialization in Biology
2009 - 2013

Honours and Awards: Ontario Graduate Scholarship
2017 - 2018

Canadian Linguistics Association - Best Student Poster Presentation
2017

Ontario Graduate Scholarship
2016 - 2017

Western University Teaching Honour Roll
2016

Related Work Experience: Bilingual Proceedings Editor
Canadian Linguistics Association
2020 - 2021

Online Teaching Support Intern
The University of Western Ontario
May 2020 - December 2020

Research Assistant - Course Development
The University of Western Ontario
January 2019 - August 2019

Co-Mentor - Emerging Caribbean Scientists
The University of the Virgin Islands
Summer 2018

Adjunct Professor - Spanish and French
The University of the Virgin Islands
January 2016 - May 2016

Teaching Assistant - Instructor of Record (French 1002)
The University of Western Ontario
2014 - 2015

Research Assistant - Linguistics
The University of Western Ontario
2013 - 2015

Publications:

Plyley, Christopher and **Angélica Hernández**. (Submitted). Language Barriers in mathematics for Creole-speaking students in St. Croix, USVI. *Journal of Language, Identity and Education*.

Hernández, Angelica, Sarah Babcock, and Andrea Boyer. (Accepted). "She's too young to be here": Perceptions of a Young Female University Professor. Case Study. *Teaching Innovations Projects*

Hernández, Angelica, and Yvan Rose. (Accepted). Introducing Phon to the Sociophonetics Community. *Proceedings from the 2019 Annual Meeting of the Canadian Linguistics Association*. Vancouver, BC.

Hernández Constantín, Angélica. 2017. Hubieron usos de formas plurales del verbo impersonal haber en el Caribe. *Proceedings of Conferencia Científica Internacional 8va edición*, Symposium 3, article 5. Holguín, Cuba.

Hernández Constantín, Angélica. 2016. La pluralisation du verbe impersonnel haber dans l'espagnol des Caraïbes et de l'Amérique Latine Continentale. *Western Papers in Linguistics*. 2 :1-27.