Subject Pronoun Expression in a Variety of Spanish in Contact with Portuguese

Ana M. Carvalho and Michael Child
University of Arizona

1. Introduction

In both Spanish and Portuguese, the expression of a subject pronoun is variable: verbs may or may not be accompanied by a phonetically realized subject. The alternation between null and overt personal pronominal subjects has been the object of variation studies that seek to explore internal and external predictors underlying the output of this variable. These studies have shown that, in general, Spanish dialects tend to favor null pronouns (see Silva-Corvalán 2001 for a comprehensive summary) while in Brazilian Portuguese, overt expression of pronominal subjects is becoming more and more frequent (Kato & Negrão, 2000; Duarte 1993, 1995, 2003, among others). As a consequence, while Spanish continues to be a prototypically pro-drop language, Brazilian Portuguese has begun to be seen as a partially pro-drop system.

In this paper, we focus on subject pronoun expression in a variety of Spanish in contact with Portuguese in a bilingual town in northern Uruguay. By analyzing the speech of Spanish-Portuguese bilinguals, our first aim is to identify and describe the factors that condition the expression of the subject pronoun in this border variety of Spanish. In addition, we compare these results with distributional patterns found in monolingual varieties of both Brazilian Portuguese and Spanish by previous studies, in order to test to what extent the behavior of this variable in this bilingual context is affected by contact with Portuguese. Our ultimate goal in investigating cross-linguistic patterned variation is to explore to what extent bilinguals maximize parallels and minimize discrepancies between their languages in situations where similar languages are in contact.

2. Method

Data were collected in the town of Rivera, on the Uruguayan-Brazilian border. Rivera, a town of approximately 85,000, represents a case of societal bilingualism: while Spanish is the official language and the one mostly used in public domains, Portuguese has been maintained as a subordinate language since colonial times. Both qualitative (Behares 1984, Eizaincín 1975, 1978) and quantitative (Carvalho 2007) studies have shown that language choice, when predictable, is conditioned by the interlocutors' social status: while the middle classes show a clear preference to using Spanish, the working class continues to use Portuguese in in-group communication. Interactional factors such as topic of the conversation and formality of speech as language choice conditioners have just started to be studied, as well as conversational functions of code-switching (Carvalho, 2010).

Exact levels of bilingual proficiency among this population have not been systematically studied. Despite the lack of statistical data on the precise number of bilingual speakers along the northern Uruguayan border communities, there is a consensus that societal bilingualism prevails in urban and suburban centers, where most dwellers use both languages in their daily lives (Behares, 1984; Elizaincín, 1978, 1992, 2008; Carvalho, 2006a). However, in isolated rural communities it is possible to find passive bilinguals, native Portuguese speakers who understand Spanish but are not fluent. On the outskirts of the urban centers, where working-class members reside, among the elderly and the ones of rural origin one may find Portuguese-dominant bilinguals, with different proficiency levels in Spanish. Waltermire (2006) studies the influence of Portuguese in the Spanish phonology of Rivera bilinguals and finds that there is more Portuguese interference in the speech of the men, the elderly,

and the working-class in the town of Rivera, concluding that these social groups are the ones among which stronger influence of the Portuguese substratum is found, bringing evidence to the fact that they are dominant in Portuguese. Spanish, on the other hand, is preferred by the youth, the women, and the higher social status groups (Carvalho, 2007). Among these groups, it is possible to find Spanish dominant speakers with only passive competence in Portuguese in urban cohorts. Of these cases, one finds speakers who, due to social and/or geographic mobility, either stopped acquiring Portuguese (incomplete acquisition) or stopped using it (language attrition). In general, language socialization practices that favor Spanish use, common among the younger members of higher status groups who live downtown and are schooled in private institutions, are preferred due to social values that emphasize national characteristics over local ones, including the use of Spanish. The intimate relations between Portuguese and the local border cultural identity have been reported in the works of Behares (1984), Elizaincín (1978), and Carvalho (2007), and are the main factor responsible for the maintenance of Portuguese in these communities, despite the Spanish-only policies that were in place in Uruguay during the 20th century.

Data for the present study were based on five months of field work in Rivera, where one of the authors carried out participant-observation and sociolinguistic interviews with bilinguals. All participants were born and raised in Rivera and selected on the basis of availability. All of them were raised bilingual in Portuguese and Spanish, used both languages in their daily lives, and were interviewed in both languages on different occasions. The conclusions presented here are based on transcriptions of 30-50 minutes of 12 one-hour interviews in Spanish only. Among these interviewees, there are 8 men and 4 women, 4 members of the working-class, 5 of the lower-middle class, and 3 of the mid-middle class. In addition, 4 represent the youngest group (15-28 years old), 3 fall in the middle (29-49 years-old) and 5 in the oldest (50-70 years old). Both socioeconomic status and age cohorts revealed themselves as important social variables in previous variation studies due to the community's social history: different socioeconomic groups have identified with the border culture in different ways, and the three generations have experienced different periods of urbanization (Carvalho 2003, 2004).

The interviews were conducted in the homes of the participants in a very spontaneous climate. The interviewer (a middle-class Brazilian female in her 30's) encouraged the participants to ask questions, introduce topics, produce lengthy narratives and maintain eye contact with her, in order to minimize the awareness of the recording. Except for the introductory questions and the parts of the interview designed to obtain formal style, the majority of the conversations focused on work, families, holidays, cooking, and individual histories.

A total of 1,706 tokens containing the variable subject pronoun expression were taken from these sociolinguistic interviews. Example 1 illustrates the realization of the 2nd person singular subject pronoun, while example 2 illustrates the absence of an overt 1st person singular subject pronoun (information inferable from the verbal morphology):

(1) Cuando **tú** insultás, **tú** te equivocas.
　　　When you insult, you make a mistake.

(2) **Ø** trabajé en la política
　　　Ø worked in politics

In accordance with the variationist methodology, we analyzed only forms that had a counterpart, meaning that only verbs that could be expressed with or without a subject pronoun. Thus, following Otheguy et al. (2007), we excluded the following cases: subject-headed relative clauses, null subjects that make reference to meteorological conditions, third person plural impersonal, explicit post-posed subjects, and hesitations/repetitions. All other cases were included.

3. Results

1,706 tokens were extracted from the transcriptions of the interviews and submitted to VARBRUL analysis using GoldVarb. Table 1 shows the overall subject personal pronoun rate:

Table 1: Expression of Subject Pronouns in Border Uruguayan Spanish

Expressed SP	602/1706	35.3%
Null SP	1104/1706	64.7%

Subject pronoun rates vary significantly among different dialects of Spanish. Among the dialects that have been studied, the one spoken in Madrid is seen as the one that favors null subject pronouns the most, while Caribbean dialects are known for the highest rate of pronoun expression (Cameron 1994). Otheguy et al.'s study (2007) shows that among Spanish speakers in New York, significantly different overall rates of expressed pronouns lead to a clear distinction between Caribbean dialects (including Puerto Rican, Cuban, Dominicans) and Mainland speakers (Ecuador, Colombia, and Mexico). Table 2 summarizes these tendencies.

Table 2: Cross-dialectal Comparison of Overall rates of Subject Pronoun Expression in Spanish

Madrid, Spain	21%
Buenos Aires, Argentina	36%
Rivera, Uruguay	**35%**
Santiago, Chile	38%
San Juan, Puerto Rico	45%

Sources: Madrid (Cameron 1994), Buenos Aires (Barrenechea & Alonso 1977 in Silva-Corvalán 2001), Santiago (Cifuentes 1980-1 in Silva-Corvalán 2001), San Juan (Cameron 1994).

As seen in table 2, as far as this variable is concerned, Spanish-Portuguese bilinguals in northern Uruguay seem to behave similarly to monolinguals of surrounding dialects of Spanish (Argentina and Chile[1]), and differently from the more radically different varieties (Madrid and San Juan). Thus, as far as this variable is concerned, Rivera Spanish participates in a homogeneous dialectal area in the Southern Cone, where the average overall rates of subject pronoun expression for all grammatical persons is around 37%. Hence, it is plausible to assume that pronoun expression in border Spanish seems to be immune to Brazilian Portuguese influence, at least in terms of overall rates of expression. Contrary to European Portuguese, which behaves like a prototypical null subject language, Brazilian Portuguese has undergone a change away from this pattern and towards overt subject pronouns in the last two centuries (Duarte 1993). Table 3 shows the overall rate of expressed pronouns in both varieties of Portuguese, in addition to Rivera Spanish.

Table 3: Cross-dialectal Comparison of Overall rates of Subject Pronoun Expression in Portuguese and Rivera Spanish

Brazilian Portuguese	56%
Rivera Spanish	**35%**
European Portuguese	22%

Portuguese rates extracted from Barbosa, Duarte & Kato 2005, based on written corpus of newspaper interviews.

As seen in table 3, the Spanish variety spoken by Portuguese-Spanish bilinguals in Rivera does not resemble the behavior documented in either variety of Portuguese: while Brazilian Portuguese overall rate of pronoun expression of 56% surpasses the one attested for Caribbean Spanish, European Portuguese behaves very similarly to what has been reported for Madrid Spanish, around 22%, both substantially different from the 37% found in our corpus of border Spanish.

The effects of contact on the expression of personal pronouns have been extensively studied in U.S. Spanish. Studies focus on the impact of English, a prototypical pro language, in the Spanish of

---

[1] As far as we know this variable has not been quantified in the speech of Montevideo, Uruguay, therefore we cannot make comparisons to this dialect.

bilinguals. While some find strong evidence that greater contact with English causes increase in the use of overt pronouns in Spanish, others find that the evidence is inconclusive (see Otheguy et al. 2007: 783-4 for a summary of the discussion).

Pronoun rates in Spanish in contact with Catalan, a more typologically similar language, are studied by De Prada Perez (2010). The author finds that despite very similar overall rates, Spanish and Catalan differ both in variable hierarchy and in constraint hierarchy. Given all the nuances involved in evaluating the permeability of linguistic systems, it would be premature to assess the effects of Portuguese in the pronoun rates of bilinguals in Rivera based on the present analysis and small corpus. Nonetheless, it is plausible to assume that within this corpus and in terms of overall rates, border Spanish does not seem to have approximated the patterns found in Brazilian Portuguese. On the contrary, its behavior resembles very closely what has been reported for monolingual varieties of surrounding dialects of Spanish.

3.1. Analysis of factors conditioning the expression of subject pronouns

Although comparisons of overall rates across dialects and languages may shed light on basic questions about the effect of contact, they say little about the effects of independent variables on the realization of subject pronoun expression. In addition, methods of data analysis may differ between studies, leading to the inability to compare results. Therefore, in this section, we aim at analyzing the linguistic and extra-linguistic factors that may condition the output of the variable in question, in order to explore the hierarchies of variables and constraints at work in this dialect. We consider this analysis to be an essential first step for further comparisons between border Spanish and monolingual varieties of both Spanish and Portuguese.

In the present study we chose to look at three sociolinguistic factors (gender, age, and socioeconomic class) as well as eight linguistic factors (parallelism, grammatical person, discourse connection, reflexive use of the verb, lexical content of the verb, tense/mood/aspect, clause type, and verb form ambiguity). Our choice of factor groups was guided by other studies, principally Otheguy et al. (2007).

3.1.1. Parallelism

Parallelism, or priming (see Travis 2007), can be defined as the tendency of linguistic forms to occur together within a stretch of discourse. This tendency has been found to be very useful in explaining variation in several variables in both Spanish and Portuguese and is classified by Labov as powerful, since there is evidence "that morphological and syntactic variation is controlled by a tendency to preserve parallel structures" (1994:550). As mentioned by Travis (2007: 107), although switch reference has been detected as a powerful conditioner of the expression of subject pronouns, there is still variation in contexts where there is continuity of reference. We assume, with Travis (see also Dumont 2006) that parallelism may affect the output of this variable, but mostly in contexts where co-referential tokens appear together. Therefore, in the analysis of the effects of parallelism, we coded for previous tokens in the discourse segment where only same-reference tokens appeared. We hypothesize that one expressed pronoun leads to another expressed pronoun consecutively, and that each time the parallel effect is stronger. Likewise, a null subject should condition a second null subject, etc. Thus, in the present study, we looked at the first, second and third occurrences within a stretch of discourse, differently from other studies where only the immediate previous token was considered (Travis 2007, Torres Cacoullos & Travis 2010a, 2010b).

It was necessary to conduct two Varbrul analyses because the first token in a parallel construction of same referents would coincide with tokens coded for 'switch reference' and 'first token', with regards to Discourse Connection, creating interactions among factor groups, which are problematic for Varbrul analysis[2]. We therefore conducted our first analysis (that provided all of the results in this study except for parallelism) and then conducted a separate analysis for parallelism, including only same-reference contexts.

---

[2] Thanks are due to Scott Schwenter, Robin Dodsworth, and Naomi Lapidus Shin for pointing out this potential interaction and for suggesting solutions. Needless to say, we alone are responsible for any errors.

In Table 4 below, E_ means that the verb in question was preceded by a verb that contained an expressed subject personal pronoun. Alternately, NN_ indicates that the verb is preceded by two verbs, both of which contained null subject personal pronouns and similarly EN_ signifies that the verb was preceded by two verbs: one with an expressed personal pronoun followed by one with a null subject pronoun. Table 4 below shows the VARBRUL results for the separate analysis conducted for parallelism in same reference contexts.

Table 4: VARBRUL Results for Factor Group "Parallelism"
Application Value = Overt Pronoun

_____

Parallelism	% expressed	% null	# of tokens	VARBRUL Weights
EE_	47.7%	52.3%	44	0.724
NE_	44.0%	56.0%	25	0.710
E_	38.8%	61.2%	152	0.680
EN_	33.3%	66.7%	54	0.614
1st Token	22.9%	77.1%	192	0.527
N_	14.3%	85.7%	189	0.401
NN_	7.7%	92.3%	130	0.241

_____

Log likelihood = -375.948	Significance = 0.024	Input = 0.199

The results indicate that a verb preceded by an expressed subject has a positive probability of being accompanied by another expressed subject (0.68), and that this probability increases when the verb is preceded by two verbs with expressed subjects (0.724). Likewise, a verb preceded by another verb with a null subject has a negative probability of being preceded by a verb with an expressed subject (0.40), a tendency that increases if the verb is preceded by two verbs with null subjects (0.241). What is not immediately clear is why a verb preceded by one null subject and one expressed subject (NE_) would have a higher probability of being expressed (0.710) than a verb being preceded by just one expressed subject (0.680). Interestingly, the first token in a stretch of discourse has an almost equal probability of being either null or expressed (0.527), all else being equal. Overall, it is clear that in the present study parallelism is a strong conditioner of subject personal pronoun expression.

3.1.2. Grammatical Person

First, second, and third person singular and plural verbs were coded to see what, if any, effect grammatical person had on subject pronoun expression. It should be noted that Rivera (and Uruguayan) Spanish have three different verb forms that were all coded for second person singular: tú, vos, and Ud. In addition, we coded for inanimate third person subject pronoun use, a usage that is relatively common in Brazilian Portuguese. However, we did not find any examples of inanimate third person subject pronouns in the tokens, contrary to Elizaincín's results (1995), which showed Portuguese-influenced usage of this variant in rural Uruguayan Spanish. Table 5 below shows the results of the VARBRUL analysis for grammatical person.

Table 5: VARBRUL Results for Factor Group "Grammatical Person"
Application Value = Overt Pronoun

Person	% expressed	% null	# of tokens	VARBRUL
2nd sing. non-spec	31.3%	68.7%	83	0.575
3rd singular	41.4%	58.6%	292	0.570
1st singular	38.0%	62.0%	931	0.543
2nd sing. spec	43.6%	56.4%	94	0.519
1st plural	25.2%	74.8%	147	0.352
3rd plural	14.5%	85.5%	159	0.237

Log likelihood = -972.237   Significance = 0.016   Input = 0.310

The VARBRUL results above indicate that third person plural is the grammatical person with the smallest probability of an explicit pronoun, followed by first person plural. All other persons present probabilities that are slightly positive. These results support the findings of other studies (see for example Otheguy et al. 2007 and Cameron 1994) that show that plural forms disfavor expressed subject personal pronouns.

3.1.3. Discourse Connection (Switch Reference)

One of the most important conditioners of subject personal pronoun expression in previous studies has been discourse connection, also called switch reference (see, for example, Silva-Corvalán 2001 and Cameron 1994). Because of the possibility of confusion when switching referents, a switch reference tends to favor subject personal pronoun expression. This tendency is confirmed in the present study. See Table 6 below.

Table 6: VARBRUL Results for Factor Group "Discourse Connection"
Application Value = Overt Pronoun

Disc. Con.	% expressed	% null	# of tokens	VARBRUL
Dif ref. (switch ref.)	49.5%	50.5%	582	0.651
First token	36.7%	63.3%	338	0.558
Same ref./dif. TMA	26.5%	73.5%	249	0.377
Same ref./same TMA	23.1%	76.9%	537	0.357

Log likelihood = -972.237   Significance = 0.016   Input = 0.310

These results are similar to the above mentioned studies in that verbs that refer to the same referent as a previous verb, regardless of whether or not the tense, mood, and aspect is the same, exhibit the smallest probability of an expressed subject personal pronoun, although switching the tense/mood/aspect does increase the probability of expressed subjects (0.357 and 0.377). In contrast, a verb that refers to a different referent than the previous verb (switch reference) has the highest probability of an expressed pronoun (0.651). The first occurrence of a verb with a possible subject personal pronoun (usually in conjunction with first or second person referents) has a slightly positive probability of containing an expressed pronoun.

3.1.4. Lexical Content of the Verb

The lexical content of the verb was analyzed to see if certain lexical categories of verbs favored the expression of subject personal pronouns. Lexical categories were based off of Travis (2007). See Table 7 below.

Table 7: VARBRUL Results for Factor Group "Lexical Content of the Verb"
Application Value = Overt Pronoun

_____

Lexical content	% expressed	% null	# of tokens	VARBRUL Weights
Psych/cog.	41.9%	58.1%	365	0.593
Copulative	39.5%	60.5%	210	0.566
Speech	42.1%	57.9%	121	0.513
Motion	30.4%	69.6%	207	0.497
Other	31.4%	68.6%	803	0.439

_____

Log likelihood = -972.237   Significance = 0.016   Input = 0.310

Although the actual probability rates differ from Travis' 2007 study, the ranking of the variables is nearly identical, with the exception that in Travis' study "speech" and "motion" are switched. Similar to that study, psychological and copulative verbs favor subject pronoun expression and "other" verbs disfavor subject pronoun expression (see Travis 2007). Verbs that deal with motion and speech do not seem to strongly favor or disfavor subject personal pronoun expression in the present study, although in Travis' study motion verbs disfavored pronoun expression.

3.1.5. Reflexive Use of the Verb

Reflexive verb use has been shown to condition subject personal pronoun expression (Otheguy et al. 2007) with reflexive verbs disfavoring expressed subjects. This is probably due to the added referential information contained within the reflexive pronoun that might make the use of the subject personal pronoun seem redundant. This tendency is verified in the present study. See Table 8 below.

Table 8: Percentages for Factor Group "Reflexive Use of the Verb"
Application Value = Overt Pronoun

_____

Reflexive Use	% expressed	% null	# of tokens	VARBRUL
Non-reflexive verb	36.1%	63.9%	1572	0.510
Reflexive verb	25.4%	74.6%	134	0.381

_____

Log likelihood = -972.237   Significance = 0.016   Input = 0.310

As might be expected, verbs that contain a reflexive pronoun disfavor the use of a subject personal pronoun whereas verbs that are not accompanied by a reflexive pronoun neither favor nor disfavor an expressed personal pronoun.

3.1.6. Tense/Mood/Aspect

In addition to lexical content, we analyzed the tense/mood/aspect of the verb to determine if this factor conditioned the variable usage of subject personal pronouns. VARBRUL discarded this factor group, but frequencies are shown in Table 9.

Table 9: Percentages for Factor Group "Tense/Mood/Aspect"

TMA	% expressed	% null	# of tokens
subjunctive	46.2%	53.8%	39
conditional	45.5%	54.5%	11
imperfect	38.7%	61.3%	256
present	37.9%	62.1%	856
preterit	30.1%	69.9%	425
imperative	30.0%	70.0%	20
perfective	24.0%	76.0%	25
future	23.1%	76.9%	52
progressive	18.2%	81.8%	22

3.1.7. Clause Type

Although we analyzed the type of clause where the verb appears, VARBRUL discarded this factor group as not being significant. The percentages of each type of clause appear below in Table 10.

Table 10: Percentages for Factor Group "Clause Type"

Clause type	% expressed	% null	# of tokens
Subordinate (other)	39.1%	60.9%	258
Main	36.6%	63.4%	837
Subordinate (relative)	33.3%	63.7%	57
Coordinate	31.7%	68.3%	555

3.1.8. Verb form Ambiguity

The final linguistic factor we assessed was verb form ambiguity, expecting this to condition subject personal pronoun expression because of the need to disambiguate certain verb forms, such as the first and third person imperfect conditional verb forms (ex. *trabajaba*--yo, él, ella, Ud., etc.). The VARBRUL analysis showed that verb form ambiguity was also not a significant conditioner of subject personal pronoun use and discarded this factor group. The percentages expressed and null subjects for ambiguous and unambiguous verb forms appear below in Table 11.

Table 11: Percentages for Factor Group "Verb Form Ambiguity"

Verb form ambiguity	% expressed	% null	# of tokens
Ambiguous verb form	41.2%	58.8%	245
Unambiguous verb form	34.3%	65.7%	1462

3.1.9. Extra-Linguistic Factors

In addition to the linguistic factors mentioned above, we coded for three sociolinguistic factors: gender, age, and socioeconomic status. In previous studies, these extra-linguistic factors did not seem to influence the expression of subject personal pronouns (see Silva-Corvalán 2001). We assumed that this would be the case with the present study. VARBRUL did discard socioeconomic status, however both age and gender turned out to be significant. See tables 12, 13, and 14 below:

Table 12: VARBRUL Results for Factor Group "Gender"
Application Value = Overt Pronoun
_____

Gender	% expressed	% null	# of tokens	VARBRUL Weights
Female	41.2%	58.8%	833	0.551
Male	29.7%	70.3%	873	0.452

_____
Log likelihood = -972.237   Significance = 0.016   Input = 0.310

One can see that males disfavored expressed subjects whereas females slightly favored expressed subjects. However, because of the small number of subjects, we are not confident that this factor conditions the expression of subject personal pronouns in this variety of Spanish. More subjects need to be added to see if this trend continues.

Table 13: VARBRUL Results for Factor Group "Age"
Application Value = Overt Pronoun
_____

Age	% expressed	% null	# of tokens	VARBRUL Weights
30-49	44.3%	55.7%	305	0.600
50-70	39.7%	60.3 %	1003	0.566
16-29	17.3%	82.7%	398	0.273

_____
Log likelihood = -972.237   Significance = 0.016   Input = 0.310

Subjects in the 16-29 age group favored null subjects whereas subjects in the 30-49 age group favored expressed subjects. Subjects in the 50-70 age group slightly favored expressed subjects. Again, we are not confidant in these results for the reasons mentioned for "gender" above and because age divisions are by nature somewhat arbitrary. We plan on adding more subjects and tokens in a future study to see if this trend continues.

Table 14: Percentages for Factor Group "Socioeconomic Status"
_____

Socioeconomic status	% expressed	% null	# of tokens
Working class	42.5%	57.5%	650
Low-middle class	36.4%	63.6%	615
High-middle class	23.1%	76.9%	442

_____

As was mentioned above, VARBRUL discarded "socioeconomic status" as not being significant. Age and gender, however, may indeed be factors related to the different levels of the influence of the

Portuguese substratum (Carvalho 2007, Waltermire 2006), however, more subjects need to be included in the analysis so that it is possible to attest if this tendency continues.

3.2. Rank of factors and discussion

As was shown above, the linguistic factor groups that condition the expression of subject personal pronouns in Uruguayan border Spanish are, in order of their VARBRUL ranking, parallelism, grammatical person, discourse connection, lexical content of the verb, and reflexive use of the verb. Compare this ranking with the factor group ranking in Otheguy et al.'s (2007) study of Ecuadorian, Columbian, and Mexican newcomers in the U.S.A in Table 15 below.

Table 15: VARBRUL ranking of significant factor groups as compared with Otheguy et al. 2007
_____

Uruguayan Border Spanish		Ecuador, Columbia, Mexico Newcomers	
Rank	Variable	Rank	Variable
1.	Parallelism	1.	Person
2.	Person	2.	Discourse Connection
3.	Discourse Connection	3.	Set Phrase
4.	Lexical Content of Verb	4.	Tense/Mood/Aspect
5.	Reflexive Use of the Verb	5.	Clause
		6.	Lexical Content of the Verb
		7.	Reflexive Use of the Verb

The order found for Uruguayan border Spanish is very comparable to the one found by Otheguy et al. (2007) for Spanish-speaking immigrants in New York. In Uruguayan border Spanish, parallelism is the strongest conditioner of subject personal pronoun usage. In other words, within a stretch of discourse, speakers tend to continue to either use expressed or null subjects when referring to the same referent, depending on what they initially chose. Interestingly this indicates that much of the variable expression of subject pronouns may actually be based more on which variant preceded the verb than on other factors such as verb type, clause, etc. Travis (2007) notes that "this has profound implications for our view of grammar, as it indicates that the grammar of discourse is developed on-line, as a response to and deriving from what precedes" (pg. 132). Interestingly, in Dumont's (2006) study of full noun phrases as subjects, parallelism (or "previous mention") was also the strongest conditioner of both personal pronoun expression and full NP's as subjects.

With regards to the lexical content of the verb, compound forms (future, perfective, progressive) and imperative forms favor null subjects whereas subjunctive forms favor expressed subjects. One possible reason for the latter is that the subjunctive mood is usually used in switch reference environments, which also tend to favor expressed subjects.

In relation to grammatical person, the present study further supports other studies that show that plural forms favor null subjects and that the first person singular favors expressed subject pronouns (see Silva-Corvalán 2001 and Otheguy et al. 2007). It should be noted, however, that studies based on sociolinguistic interviews may understate the tendency to use expressed subjects with 1st person singular forms as compared with casual conversations due to the somewhat unusual situation of having an interviewer and an interviewee. In this situation, and especially in the beginning stages of the interview, it would seem superfluous for the interviewees to answer questions directed at them with expressed subject personal pronouns.

Finally, the reflexive use of the verb disfavors subject personal pronoun usage, presumably because the reflexive pronoun provides contextual information that would make the subject personal pronoun redundant. In addition, verb form ambiguity does not appear to condition the variable expression of subject personal pronouns, contrary to a functional hypothesis that predicts the need of expressed subject pronouns in cases where verb form is ambiguous. It appears that contextual knowledge clarifies much of the ambiguity inherent in the morphological forms.

4. Conclusion

In closing, this study has detected parallelism, grammatical person, discourse connection, lexical content of the verb, and the reflexive use of the verb to be the linguistic conditioners of subject pronoun expression in Uruguayan border Spanish. Both age and gender were classified as important extra-linguistic conditioners; however, the corpus is too small for any conclusive interpretation with regards to these sociolinguistic factors. As far as the effects of Portuguese on border Spanish, overall rates extracted from this corpus suggest that bilingual speakers seem to be using subject personal pronouns at the same rate of monolingual speakers of surrounding dialects of Spanish, which leads us to discard, a priori, a contact hypothesis. These preliminary results yield a basis of comparison to be used by further studies. In order to appropriately assess the extent of the influence of the Portuguese contact, it is necessary, first, to expand our corpus. In addition, it is crucial that we code the same subjects' interviews in Portuguese to then compare the distribution of this variable in Uruguayan Portuguese with current results of Uruguayan border Spanish.

References

Barbosa, Pilar; Duarte, Maria Eugênica; Kato, Mary. 2005. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4:2, 11-52.

Behares, Luis E. Diglosia en la sociedad escolar de la frontera uruguaya con Brasil. Matriz social del bilinguismo. *Cuadernos de Estudios Lingüísticos*. 6, 228-234, 1984.

Cameron, Richard. 1994. Ambiguous agreement, functional compensation, and nonspecific tú in the Spanish of San Juan, Puerto Rico and Madrid, Spain. *Language Variation and Change*. 5, 305-334.

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

Carvalho, Ana Maria. 2010. Contribuições da sociolingüística ao ensino do português em comunidades bilíngües do norte do Uruguai. *Revista Pro-Posições*. 21:3. 45-65.

Carvalho, Ana Maria. 2007. Diagnóstico sociolingüístico de comunidades escolares fronterizas en el norte de Uruguay. In: Brian, Nicolás; Brovetto, Claudia; Geymonat, Javier (Eds.) Montevideo: Administración Nacional de Educación Pública. República Oriental del Uruguay, 44-96.

Carvalho, Ana Maria. 2006. Políticas lingüísticas do século passado nos dias de hoje. O dilema da educação bilíngüe no Uruguai. *Language Problems and Language Planning*, 30,2, 149-171.

Carvalho, Ana Maria. 2004. 'I speak like the guys on TV': Palatalization and the urbanization of Uruguayan Portuguese. *Language Variation and Change*, 16,2, 127-141.

Carvalho, Ana Maria. 2003. Rumo a uma definição do português uruguaio. *Revista Internacional de Lingüística Iberoamericana.* 2, 135-159.

De Prada Perez, Ana. Variation in Subject Expression in Western Romance. 2010. In Colina, S., Olarrea, A. & Carvalho, A. M. (eds.). *Romance Linguistics 2009. Selected Papers from the 39th Linguistic Symposium on Romance Languages (LSRL)*: Amsterdam: John Benjamins, 267-284.

Duarte, Maria Eugênia. 2003. A evolução na representação do sujeito pronominal em dois tempos. In Paiva, M. C. & Duarte, M. E. (eds.). *Mudança Linguística em Tempo Real*. Rio de Janeiro: FABERJ. 115-128.

Duarte, Maria Eugênia L. 1995. *A perda do princípio 'evite o pronome' no português brasileiro*. Campinas: Unicamp. Tese de Doutorado.

Duarte, Maria Eugênia L. 1993. Do pronome nulo ao pronome pleno: a tragetória do sujeito no português do Brasil. In: Roberts I; Kato M (orgs.) *Português Brasileiro: Uma viagem diacrônica*. Campinas: Unicamp, 107-128.

Dumont, Jenny. 2006. Full NPs as Subjects. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 286-296. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1386

Elizaincín, Adolfo. 2008. Oito considerações sobre o contato linguístico. In Espiga, J. & Elizaincín, A. (Orgs.). *Español y português: um (velho) Novo Mundo de fronteiras e contatos*. Pelotas: Educat, 405-424.

Elizaincín, Adolfo. 1992. *Dialectos en contacto. Español y Portugués en España y América*. Montevidéu: Arca.

Elizaincín, Adolfo. 1995. Personal Pronouns for Inanimate Entities in Uruguayan Spanish in Contact With Portuguese. In Silva-Corvalán, C. (Org.). *Spanish in four continents*: studies in language contact and bilingualism. Washington, DC: Georgetown University Press. 117-132.

Elizaincín, Adolfo. 1978. Bilingüismo y problemas educativos en la zona fronteriza Uruguayo-Brasileña. *Lingüística y Filología de América Latina*. Lima: Universidad de San Marcos, 301-310.

Elizaincín, Adolfo. 1975. El bilingüismo de la frontera uruguayo-brasileña, *Letras de Hoje*, 20, 65-75.

Kato, Mary; Negrão, Esmeralda. 2000. *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt: Vervuert-Iberoamericana.

Labov, William. 1994. *Principles of Linguistic Change: Volume 1: Internal Factors*. Oxford: Basil Blackwell

Otheguy, Ricardo; Zentella, Ana Celia; and Livert, David. 2007. Language and dialect contact in Spanish in New York: Towards the formation of a speech community. *Language* 83: 4, 770-802.

Torres Cacoullos, Rena; Travis, Catherine. 2010a. Variable yo expression in New Mexico: English influence? In Susana Rivera-Mills & Daniel J. Villa (eds.), Spanish in the U.S. Southwest: A language in transition, 185-206. Madrid: Iberoamericana.

Torres Cacoullos, Rena; Travis, Catherine. 2010b. Testing convergence via code-switching: priming and the structure of variable subject expression. International Journal of Bilingualism. Online. ijb.sage.com.

Travis, Catherine E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19, 101-135.

Silva-Corvalán, Carmen. 2001. *Sociolingüística y pragmática del español*. Georgetown: Georgetown University Press.

Waltermire, Mark. 2006. *Social and linguistic correlates of Spanish and Portuguese bilingualism on the Uruguayan-Brazilian border*. Ph.D. Dissertation. University of New Mexico, 2006.

Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics

edited by Jim Michnowicz and Robin Dodsworth

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Carvalho, Ana M. and Michael Child. 2011. Subject Pronoun Expression in a Variety of Spanish in Contact with Portuguese. In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, ed. Jim Michnowicz and Robin Dodsworth, 14-25. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2502.