

# First Person Subject Pronoun Expression in the Spanish of Tucson

Kelly A. Lowther  
The University of Arizona

## Abstract

A well-known area of research on Spanish grammar is that of the variable expression of personal subject pronouns. This variable has been studied with respect to a wide range of both linguistic and social factors, and across various dialects including Peninsular Spanish, U.S., Latin American and Puerto Rican dialects. The present study examines the null and overt subject pronoun expression in the bilingual community of Tucson, Arizona, in relation to switch-reference, morphological ambiguity, gender, English and Spanish language proficiency, and social class.

The findings provide further evidence that switch reference is a significant factor that conditions subject pronoun variation in Spanish, as previous studies have reported (Cameron, 1992; Ávila-Jiménez, 1996; Bayley & Pease Álvarez, 1997; Flores-Ferrán, 2002; Silva Corvalán, 1994). This study also shows no correlation between morphological ambiguity nor English proficiency level and the rate of subject pronoun expression.

## Introduction

Spanish is classified as a pro-drop language, a language with rich verbal morphology that marks tense, mood, aspect, person, and number, leaving the speaker with the option to express or omit subject personal pronouns. This opposes non pro-drop languages, such as English, that lack this same inflectional richness and where, in almost all cases, subject pronouns must be overtly expressed. From a researcher's point of view, there emerges a fascinating area of investigation given the nature of the Spanish language. Specifically, what factors influence the usage of null and expressed subject pronouns? What are the linguistic and extralinguistic constraints on the election of pronominal overtness or covertness?

Of the research conducted thus far, the most common linguistic factors examined affecting pronoun expression include: morphological ambiguity of verbal form, form used in the previous mention of the verb's subject, and switch reference. Various extralinguistic factors, including age, gender, educational level and speech style, have

also been investigated. In addition, possible social motivators for the expression of personal pronouns in Spanish include contact with other dialects and languages.

Research to date has shown many similar effects in different dialects, although rates of pronoun expression vary. Cameron (1992) investigated the effects of several different linguistic factors on subject pronoun realization, including morphological ambiguity and switch reference in Madrid and San Juan dialects of Spanish. His findings revealed that same or switch reference is the main constraint on pronoun expression in both dialects, and furthermore, that the effects of morphological ambiguity interact with switch reference on pronoun expression. Other factors investigated, such as reference chains and verb class, revealed similar effects.

In her dissertation study of the Columbian speech community in Miami-Dade County in South Florida, Luz Hurtado (2002) examined the expressed and non-expressed use of second person, third person and impersonal pronouns in different dialectal varieties. Hurtado examined the ambiguity of verbal forms, and also considered the influence of factors such as contact with other Hispanic communities and contact with English. Her findings suggested that linguistic factors determine more the expression of subject pronouns than do social factors. No relation was found between pronoun realization and the level of contact with English or level of English proficiency.

Other studies examining the influence of other languages on Spanish subject pronoun expression have shown similar results. For example, Ávila-Jiménez's (1996) comparative study of Puerto Rican Spanish on the island and mainland U.S., found no correlation between the use of overt subject pronouns and exposure to English. Research by Silva-Corvalán (1990, 1994) examined subject expression and pronoun use in the bilingual Chicano speech community in Los Angeles, California, also with results indicating that most restrictions on pronoun expression agreed with those of monolingual Spanish communities. English contact did not influence speakers of this Chicano Spanish variety to use more overt forms. Interestingly, however, Silva-Corvalán's results did reveal a weakening of the coreference constraint in the bilingual environment, which she attributed to a loss of semantic-pragmatic options in secondary languages.

Likewise, in a study of subject pronoun usage in Mexican-descent children in Eastside, California, Bayley and Pease-Alvarez (1996) found that English dominant children do not exhibit greater pronoun expression than those children who learned Spanish in Mexico. Rather, the main constraints on pronoun expression are person/number and coreference. Their study did, however, reveal a weakening of the surface ambiguity constraint on pronoun expression in Eastside Spanish.

Perhaps the most extensive study of subject pronoun expression in a language contact environment is that of Flores-Ferrán (2004). Her investigation of the Puerto Rican Spanish dialect in the language contact environment of NYC, found that switch reference does act as a predictor of overt usage of subject pronouns. In relation to contact with English, she found

that NYC native-born Puerto Ricans consistently used more overt pronominal forms than both recent arrivals and established residents in NYC. This consistency supports an English contact hypothesis. However, Flores-Ferrán also found evidence to the contrary. In comparing frequencies of overt subject pronoun expression with speakers of Puerto Rico, she found no dissimilarities. Overall, there was a lack of correlation between language used in the home, age of arrival or total years in NYC with an increase in overt usage of subject pronouns.

In sum, research to date has demonstrated that subject pronoun variation is due to a wide range of factors. There has been considerable debate on the effects of the influence of English on subject pronoun expression. In general, studies have revealed that linguistic factors, such as switch reference and morphological ambiguity, are the strongest factors influencing the usage of null or expressed subject pronouns in Spanish.

Analysis in the current project will focus on the explicit expression or omission of the first person subject pronoun (*yo*) in the bilingual Sonoran Desert speech community of Tucson, Arizona. By 'speech community', here I refer to Labov's (1972) definition: "The speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms. These norms may be observed in overt types of evaluative behavior, and by the uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage" (120-121). Based on this definition, we can consider these speakers members of a speech community, participating in a set of 'shared norms.'

The study will examine the significance of effects of both linguistic and extralinguistic factors on variation of first person pronoun expression. Specifically, this study analyzes the effects of gender, social class, and English and Spanish proficiency level (as measured by participants' self-evaluation), in comparison with linguistic factors of switch reference and morphological ambiguity, which previous research studies have indicated to be the primary constraints on expressed pronoun variation.

Due to the fact that, as Silva-Corvalán (1994) highlights, in bilingual communities, the grammar of the less dominant language is exposed to a certain permeability because the changes it predicts are "less salient and could spread with relative ease" (Bayley and Pease-Alvarez, 1996: 86), I hypothesize that the higher the level of proficiency in English, the greater the rate of subject pronoun expression will be. Given the proximity of this bilingual community to the border of Mexico, it makes for a very interesting point of investigation. Since pronoun expression in English is obligatory, with English being the dominant language in this border community, I expect this variable to represent a change, specifically, a higher rate of pronoun expression in Spanish.

Also, much research to date has demonstrated linguistic factors to be the primary constraints on pronoun expression. Therefore, I hypothesize that switch reference will show as a significant factor that conditions Spanish subject pronoun variation.

## Methodology

The participants of this study were 18 native Spanish speakers (9 females/9males) of the bilingual community in Tucson, AZ. Participants ranged in age from 18-54, and represented lower to upper middle social classes. Each student in the course interviewed two participants and transcribed the interviews. In five of the interviews conducted, both participants were interviewed simultaneously. A total of 1007 tokens were identified for analysis, by selecting approximately the first 80 tokens from each interview in which the first person singular form of the verb was used. The collected data was then analyzed via VARBRUL multivariate analysis using GoldVarb2001, a quantitative statistical analysis program that both identifies the primary constraints on linguistic variation and allows for direct correlations to be made between linguistic variables (e.g. switch-reference, morphological ambiguity) and extra-linguistic factors (e.g. age, gender). Analysis was performed according to the codification factors detailed below.

## Codification Factors

The dependent variable in this study was the first person subject pronoun realization: whether the personal subject pronoun was expressed (R) or not expressed (N). Below is a list of the independent variable groups included for analysis and their multiple corresponding factors:

1. *Switch reference*: For this factor group, I adopted the definition provided by Cameron (1992, 314), where the relationship of same and switch reference is defined between two NP's where the second NP is the [+human] subject of a tensed verb that occurs after and nearest to another subject NP of a tensed verb. NP(1) is known as the trigger and NP(2) as the target, as illustrated in the following formula used by Cameron (1992: 314):

$$\frac{\text{NP + TensedV (X)}}{(1)} \dots \frac{\text{(Y) NP + TensedV (Z)}}{(2)}$$

Examples shall be used from the present corpus, provided with: name, gender, age, language proficiency (where *S* = Spanish dominant, *E* = English dominant, *B* = balanced bilingual, and *L* = Spanish monolingual) and socio-economic class (where *h* = high class, *m* = middle class, and *l* = lower class).

1.a. *Change in referent (C)*. For example:

(1) *Se regresaron todos. Yo me quedé aquí.* (Const., M, 21, S, m)

1.b. *Coreference with subject, or partial change in referent (P)*. For example:

(2) *Mi hermano y yo somos los que todavía no estamos casados y también tengo a mi mamá y mi papá.* (Les., F, 31, B, m)

1.c. *No change in referent, or same referent (Y)*. For example:

(3) Cuando *puedo* hablar el español, *yo hablo* el español. (Const., M, 21,S, m)

2. ***Morphological Ambiguity of the Verbal Form***; Morphological ambiguity was determined with respect only to verbal morphology. Morphologically ambiguous forms include the first and third person forms of the conditional, imperfect and subjunctive tenses. It should be noted that contextual or discourse factors were not taken into consideration and ambiguity was not evaluated in this respect.

2.a. *Ambiguous (A)*. (Subject pronoun may be any of the following: *yo, él, ella, Ud.*)

For example:

(4) Cuando le *atendía*. (Usage of imperfect tense; Gab., F, 30, B, m)

(5) Cuando ya *sea* profesor. (Usage of subjunctive tense; Gab., M, 26 E, m)

(6) Y lo *podría* hacer. (Usage of conditional tense; Gab., F, 30, B, m)

It should be noted that even in instances where the context or discourse indicated the subject, if the verb was ambiguous in its conjugated form, it was coded as ambiguous, as in the following example:

(7) Hay muchas personas que hablan el español, ¿no? Entonces yo hablo, pues, puro español cuando más *pueda*. (Cons., M, 21, S, m)

Although the contextual information provided in the discourse indicates that the subject of *pueda* is the first person pronoun (*yo*), this example is coded as ambiguous because the verbal form of the present subjunctive can have a subject of either *yo, él, ella or Ud.*

2.b. *Unambiguous (B)*. (Verbal morphology indicates that subject is first person pronoun)

For example:

(8) *Voy* al mercado.

### 3. ***Gender***

3.a. Male (M)

3.b. Female (F)

4. ***Language Proficiency***; Language proficiency was determined by self-evaluation on the part of the participants and according to the following bilingual continuum:

4.a. Dominant in Spanish (S)

4.b. Dominant in English (E)

- 4.c. Balanced Bilingual (B)
- 4.d. Spanish Monolingual (L)

5. *Socio-economic class*; Socio-economic class information was obtained from the participants by their respective interviewers.

- 5.a. Upper (h)
- 5.b. Middle (m)
- 5.c. Lower (l)

It should be noted that in coding cases of apparent unintentional repetition of the verb on the part of the speaker, the token was counted just once. These cases revealed themselves as the speaker needed more time to complete the thought, as illustrated in the following example:

(9) *No soy, no soy* muy amante a ninguno. (Les., F, 31, B, m)

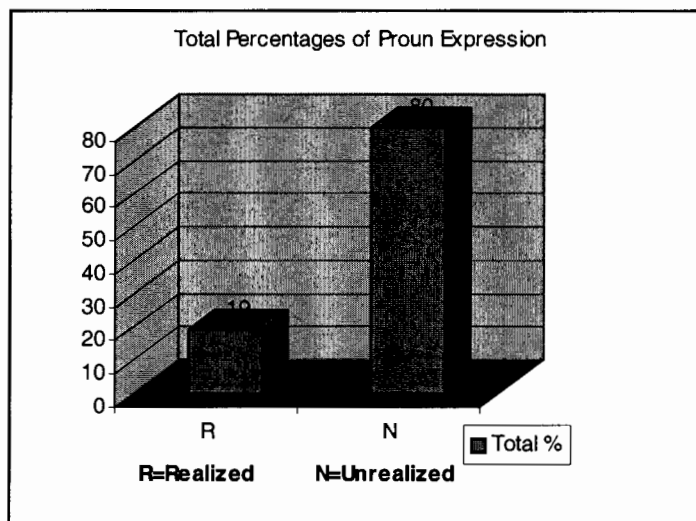
However, in cases demonstrating a listing of events, actions, etc., the verbs were counted as separate tokens, as in the following example:

(10) *Fui* a España, *fui* a Portugal. (Jos., M, 28, S, m)

This completes the listing of codification factors used for analysis in the present study and explanation of how each was measured. I now turn to presentation and discussion of the results.

## Results and Discussion

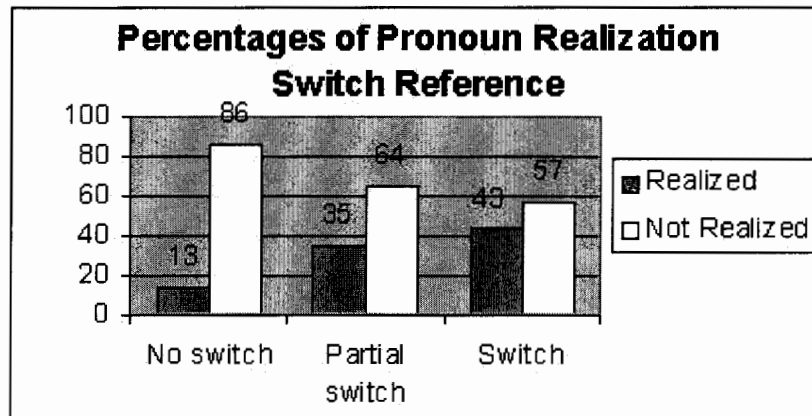
Not surprisingly, the results indicated an overall tendency to not express first person subject pronouns in the bilingual community of Tucson, Arizona. Of a total number of 1007 tokens, 194 pronouns were realized (19%) and 813 were not realized (80%), as illustrated in Graph 1 below.



Graph 1. *Total Percentages of Pronoun Expression*

Data was analyzed using VARBRUL, Goldvarb2001. Results of multivariate analysis revealed that switch reference is a significant factor that conditions variation of first person subject pronouns in Tucson Spanish. All other factor groups, including gender, socio-economic class, language proficiency and morphological ambiguity, were eliminated by VARBRUL as having an insignificant influence on pronoun variation.

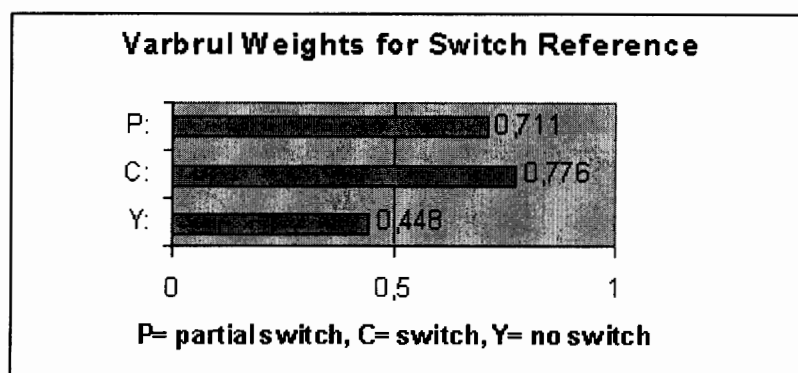
If we examine the results from switch reference, we can see that the percentages of overt pronoun usage increase with a change in referent. This trend can be easily viewed in Graph 2.



Graph 2. Percentages of Pronoun Realization for Switch Reference

When there is no switch in referent, pronoun realization is at 13%, while a partial switch in referent (or coreference) increases to 35%, and finally in the case of a switch in referent, 43%.

VARBRUL weights for switch reference indicate that this factor strongly conditions variation of first person pronoun expression, as illustrated in Graph 3 (change in referent:  $p_i=.776$ ; partial change, or coreference:  $p_i=.711$ ). When there is no change in referent, realization of pronoun expression is inhibited ( $p_i=.448$ ). These results confirm the hypothesis that switch reference significantly influences subject pronoun realization.



Graph 3. VARBRUL Weights for Switch Reference

Language proficiency level was found not to be a significant factor in pronoun expression, with VARBRUL results shown in Table 1. These findings thereby fail to support an English contact hypothesis. These results are consistent with previous research findings that compare the influence of English on subject pronoun expression (Cameron, 1992; Flores, 2002; Flores-Ferrán, 2004; Silva-Corvalán 1990, 1994; Ávila-Jiménez, 1996; Hurtado, 2002; Bayley & Pease Alvarez, 1996).

Language Proficiency Level	Varbrul Weights
Spanish Dominant	513
English Dominant	460
Balanced Bilingual	560
Spanish Monolingual	416

Table 1. *VARBRUL Weights for Language Proficiency Level*

Finally, as Table 2 demonstrates, morphological ambiguity of verb form was not a significant factor influencing pronoun expression.

Morphological Ambiguity of Verbal Form	Varbrul Weights
Ambiguous	578
Unambiguous	489

Table 2. *VARBRUL Weights for Morphological Ambiguity of Verbal Form*

This result is both consistent and inconsistent with previous findings. For example, Ávila-Jiménez (1996) found no correlation between verb ambiguity and pronominal expression. Similar results were also found by Enríquez (1984) in Madrid, Miró Vera and Angel de Pineda (1982) in Sevilla, and Barrenechea and Alonso (1977) in Buenos Aires (Cameron, 1992). Other findings, however, have shown morphological ambiguity to determine overt pronoun expression, (Silva-Corvalán, 1982, 1990; Cameron, 1992, Hochberg 1986).

The research findings on morphological ambiguity are contradictory. Not all research on Spanish dialects shows the same effects for morphological ambiguity on pronoun expression. [With respect to this issue, Cameron (1992) concludes that, “inflectional richness is a graded feature not simply of languages, but of dialects of the same language” (p. 310). Indeed, it may be the case that the dialect of Spanish in Tucson differs in this respect from some other dialects. These results may also suggest, however, that verbal morphology is not always necessary since other contextual clues in the discourse often clearly indicate the subject, perhaps influencing the speaker to omit the subject pronoun.



## **Conclusion**

In conclusion, the results from the present study indicate that linguistic factors condition first person subject pronoun variation more than extralinguistic factors in the Spanish of Tucson, Arizona. These findings are in accordance with previous research results on subject pronoun expression (Cameron, 1992; Flores, 2002; Flores-Ferrán, 2004; Silva-Corvalán 1990, 1994; Ávila-Jiménez, 1996; Hurtado, 2002; Bayley & Pease Alvarez, 1996). Switch reference emerges as a significant factor influencing subject pronoun expression, as studies of other Spanish dialects have shown. Gender and socio-economic class in this study were shown to be insignificant.

Interestingly, there was no evidence of the influence of English on an increased usage of overt Spanish subject pronouns. The results fail to support an English-contact hypothesis, as shown by previous research findings (Cameron, 1992; Flores-Ferrán, 2004; Silva-Corvalán 1990, 1994; Ávila-Jiménez, 1996; Hurtado, 2002; Bayley & Pease Alvarez, 1996).

This study's findings on morphological ambiguity contribute to the pool of conflicting findings on this factor. The results suggest that perhaps speakers rely on contextual clues more so than verbal morphology. My suggestion for future study would be to examine ambiguity also in terms of context, a perspective on ambiguity that the current study fails to provide.

In looking back at the data, an important factor that may have influenced the frequency of subject pronoun expression is turn in speech. The current study did not include this as a factor group. This factor may prove to be particularly important in conversations with more than one other person, and in the case of this study's interviews with more than one participant. In such situations, the speaker may use the pronoun as a means of 'taking the floor'. Turn in speech would be a factor group that I would include in the future.

Also interesting for future investigation would be an examination of emphatic or contrastive subject pronoun usage, as done by Davidson (1995). Subject pronouns overtly used for emphasis or contrast were common in the present study's data, and this type of analysis might have served beneficial to better understand the role of subject pronoun expression in Spanish.

In sum, results from the present study provide valuable evidence in support of the role of linguistic factors in subject pronoun expression in Spanish. These findings are consistent with of the majority of previous studies on pronoun expression. Moreover, they also confirm previous study findings that contact with English does not increase overt subject pronoun expression, pointing toward a no contact hypothesis.

## **Cited Works**

Ávila-Jiménez, B. (1996) Subject pronoun expression in Puerto Rican Spanish: A sociolinguistic, morphological, and discourse analysis. Doctoral dissertation, Cornell University.

Barrenechea, A. M. and A. Alonso (1977) Los pronombres personales sujetos en el español hablado en Buenos Aires. In J. L. Blanch (ed.), *Estudios sobre el español hablado en las principales*

ciudades de América. Ciudad Universitaria: Universidad Nacional Autónoma de México, 333-49.

Bayley, R. and L. Pease Alvarez (1996) Null and expressed pronoun variation in Mexican-descendent children's Spanish. In Jennifer Arnold, Renée Blake, Brad Davison, Scott Schwenter, & Julie Solomon, eds. *Sociolinguistic Variation Data Theory and Analysis*, NWAV 23 at Stanford, 84-99.

Cameron, R. (1992) Ambiguous agreement, functional compensation, and nonspecific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language variation and change*. 5, 305-34.

Davidson, B. (1995) 'Pragmatic weight' and Spanish subject pronouns: The pragmatic and discourse uses of 'tú' and 'yo' in spoken Madrid Spanish. *Journal of Pragmatics*. 26, 542-65.

Enríquez, E. (1984) *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científica, Instituto Miguel de Cervantes.

Flores-Ferrán, N. (2004) Spanish subject personal pronoun use in New York City Puerto Ricans: Can we rest the case of English contact? *Language Variation and Change*. 16, 49-73.

Flores, N. (2002) Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City: a variationist study. *Dissertation Abstracts International*, A: The Humanities and Social Sciences 62. 12, June 4143-A.

Hochberg, J. (1986) Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language*. 62(3), 609-21.

Hurtado, L. M. (2002) Variable subject expression in the Spanish of Miami-Dade Columbian and Columbian-American residents. *Dissertation Abstracts International*, A: The Humanities and Social Sciences. 62, 9, March 3027-A.

Labov, W. (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.  
Miró Vera, R. and M. Angel de Pineda (1982) Determinación sociolingüística de la presencia/ausencia del pronombre personal sujeto. In M. Teresa Palet Plaja (ed.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*. Seville: Servicio de Publicaciones de la Universidad de Sevilla, 37-44.

Silva-Corvalán, C. (1982) Subject expression and placement in Mexican-American Spanish. In J. Amastae & Elías-Olivares (eds.), *Spanish in the United States: Sociolinguistic aspects*. New York: Cambridge University Press, 93-120.

Silva-Corvalán, C. (1990) Current issues in studies of language contact. *Hispania*. 73, 162-176.

Silva-Corvalán, C. (1994) *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon, 154-169.