

## Frequency rates and constraints on subject personal pronoun expression: Findings from the Puerto Rican highlands

JONATHAN HOLMQUIST

*Temple University*

### ABSTRACT

This study examines subject personal pronoun expression in the Spanish of the west-central highlands of Puerto Rico. Although rates of *s*-deletion are comparably high, rates of overt subject expression are shown to be much lower than rates reported for varieties of coastal Puerto Rican Spanish and U.S. mainland Puerto Rican Spanish. The linguistic constraints on overt versus null pronoun usage in the data are shown to coincide to a very large extent with constraints identified for other Puerto Rican dialects and also Castilian Spanish in central Spain, whereas of the social factors, only the distinction between farmers and nonfarmers is significant. The study suggests that, if rates of personal subject pronoun expression are an indication of dialectal variation, the rates presented here for this syntactic phenomenon represent the continuing effects of a conservative dialect in the interior of the island of Puerto Rico.

The island of Puerto Rico and Puerto Rican communities on the U.S. mainland have served as a workshop for the study of language and, especially, of language variation and change (Álvarez Nazario, 1974, 1977, 1982, 1991; Cameron, 1992, 1994, 2005; López Morales, 1983; Morales, 1986; Navarro Tomás, 1948; Poplack, 1980; Torres, 2002). This has particularly been the case for one of the most widely studied features of Spanish dialectology, the variable use of subject pronouns. Studies by Hochberg (1986) and Otheguy, Zentella, and Livert (2007) and a series of studies linking them share the finding of high frequencies of subject personal pronoun use in the samples of Puerto Rican Spanish they analyzed. The current study diverges from this finding by presenting data suggestive of a low subject pronoun frequency dialect in the interior of the island. It also examines linguistic and social factors in relation to subject personal pronoun usage, in the first case in relation to both high frequency Puerto Rican varieties and low frequency varieties, in particular varieties in central Spain. Lastly, this presentation will consider why a lower frequency, more conservative dialect with respect to the optional use of subject personal pronouns may be found in the interior of the island.

I am grateful to the American Philosophical Society and Temple University for grants supporting the fieldwork that made this presentation possible. I would like to thank the people of Castañer, Puerto Rico, for their friendship and support as well.

## CASTAÑER AND WEST-CENTRAL PUERTO RICO

The fieldwork for this study took place between 1993 and 1997 in the community of Castañer (population 3000) and its mountainous environs, on the border between the municipalities of Lares and Adjuntas in west-central Puerto Rico. The houses forming the nucleus of the town of Castañer were built in the mid-1930s on lands of Hacienda Castañer, in the midst of one of the richest coffee-producing regions on the island.

Navarro Tomás (1948) characterized the west-central area of the island as highly conservative. Phonologically, features of the region have included the preservation of tense tonic vowels similar to those of Castilian, the use of high final *-u* and *-i* (as in *lechi* ‘milk’ and *muchu* ‘much’), the preservation of the distinction between final *-r* and *-l*, and the use of Puerto Rican velar long *rr* (Navarro Tomás, 1948:48–50, 80–81, 95). My research (Holmquist, 2008:28–30) showed that morphosyntactically the west-central highlands have preserved archaic and nonstandard forms as well, particularly the use of the auxiliaries *ha* and *hamos* (first-person singular and plural, respectively, of present indicative *haber* ‘to have’) rather than *he* and *hemos*; the use of the first-person plural suffix *-nos* rather *-mos* as in *trabajábanos* ‘we used to work’; and the use of person/number concordance with existential *haber*, as in *hay/hayn/habemos* and *había/habían/habíanos* (third-person singular, third-person plural, first-person plural of present indicative and imperfect, respectively) versus the single forms *hay* and *había*. I have also found (Holmquist, 1998, 2003, 2005, 2008, 2011) that the preservation of these regional features, both phonological and morphosyntactic, depends largely on social factors such as age, occupation, and network ties.

Raised by this study is the question of whether or not a region characterized as originally conservative in areas of phonology and morphosyntax may still very generally preserve traditional usage in relation to an aspect of underlying syntax. Is it still characterized by a relatively low rate of personal subject pronoun expression closer to rates associated with the Spanish of central Spain and conservative areas such as Mexico in Latin America (Silva-Corvalán, 1977) than to much higher rates observed in linguistically innovative Latin American varieties, in this case, of Puerto Rican Spanish? Moreover, is this true in a region currently characterized by progressively higher rates of *s*-deletion across progressively younger generations and a high, 69%, overall rate of *s*-deletion among the same group of speakers sampled for this study (Holmquist, 2011)?

## THE SPEAKER AND SPEECH SAMPLES

The analysis here draws on a speaker sample of 60 persons. All but four are natives of Castañer or the surrounding mountain area; the four non-natives are from other areas of the western highlands. Twenty speakers are drawn from each of three generations, or age groups: 15 to 39 years, 40 to 64 years, and 65 years and older. Each age group is composed of 10 men and 10 women, and each of these

groups of 10 men and 10 women includes 5 speakers characterized as having more closed, or local, ties within the community and 5 having more global, or open, ties reaching beyond. The local versus global ties distinction has been made using the criteria of time spent in the community, membership in local organizations, ties to the agricultural economy in the form of participation in the annual coffee harvest, and education, with postsecondary degrees requiring study outside of the community, normally in coastal areas of the island.

Findings presented are based on recorded, semidirected conversations focusing on personal history, the family, the schools, local churches, local cooperatives, and the coffee, and on relations between Castañer and the larger municipal centers of Lares and Adjuntas, the interior of the island and the coast, and the island and the U.S. mainland. All of the conversations were directed by me; the speakers are friends and friends of friends of the local family that I stayed with on my visits to Castañer, including several extended visits prior to the beginning of the recordings. To include balanced representation, I selected the first 50 sites of subject personal pronoun expression or nonexpression in the conversation with each speaker for the analysis. The full data sample includes 3000 tokens.

#### THE VARIABLE

The variable examined here is the use (or overt expression) in opposition to the lack of use (or nonexpression) of a personal subject pronoun with a human referent in association with a tensed verb whose subject can be either expressed or null. If alternation could not occur, the verb was excluded from the analysis.

The following example presents instances in which a speaker had the option of overtly expressing the subject personal pronoun or omitting it.

- (1) *Yo (a) vivía en Lares. (b) Nací en Lares y (c) vivía en Lares. Entonces, había necesidad de trabajar porque se había muerto mi papá y (d) teníamos una familia de seis personas. Y mi abuela y mi mamá, pues (e) vivíamos todos juntos. Entonces, pues, cuando (f) me gradué de octavo grado ... pues (g) me gradué con buenas notas. Entonces (h) ellas me ofrecieron para ir a estudiar pero (i) puse en balance la situación y (j) le dije a mi mamá: (k) yo prefiero trabajar.*

(Participant #19)

I (a) lived in Lares. (b) [I] was born and (c) [I] lived in Lares. Then, there was the need to work because my father had died and (d) [we] had a family of six people. And my grandmother and my mom, well (e), [we] lived all together. Then, well, when (f) [I] graduated from the eighth grade ... well (g) [I] graduated with good grades. So (h) they offered me (the chance) to go to study but (i) [I] weighed the situation and (j) [I] told my mom: (k) I prefer to work.

The speaker had the option of using all of the verbs, (a) through (k), with an overt subject personal pronoun, but she has done so only for (a), (h), and (k). The subject personal pronoun has been expressed in 3 of 11 opportunities.

Environments in which subject personal pronouns are obligatorily expressed or not expressed were excluded from this analysis. Exclusions of obligatorily expressed pronouns include:

- (2) Phrasal idioms, as in:

*Que yo sepa.* (Participant #10)

As far as I know.

- (3) Emphatic uses with *mismo/s* or *misma/s*, such as:

*Yo mismo quería ayudarlo.* (Participant #25)

I myself wanted to help him.

*Yo misma ... ha viajado mucho.* (Participant #43)

I myself ... have traveled a lot.

- (4) Equational sentences, as in:

*El problema es él.* (Participant #18)

The problem is him.

- (5) Instances of contrasting usage, such as:

*Ella era distante pero él rindió una labor muy buena.* (Participant #9)

She was removed but he produced very good work.

In the last case, although the use of *ella* is open to variation, the context presented by *rindió* is excluded from analysis because of the need to use *él* to preserve the contrast with the preceding subject of the same grammatical person and number.<sup>1</sup> Excluded as well from the analysis are nonspecific uses with *uno* and all nonpersonal uses with demonstrative pronouns, as for example in *Uno no se imagina ...* ‘One doesn’t imagine ...’ and *Eso duró tres años* ‘That lasted three years’, respectively.

Cases in which the use of a null subject pronoun was the only option were also excluded from the analysis. These include the following:

- (6) Nonspecific third-person plural subjects, as in:

*¿Le han hablado del café gourmet?*<sup>2</sup> (Participant #21)

Have they spoken to you about the gourmet coffee?

- (7) Nonpersonal subject pronouns, such as:

*Yo visitaba mucho la iglesia Episcopal porque (\*) nos quedaba más cerca.*

(Participant #5)

I visited the Episcopal church a lot because it was closer for us.

- (8) Subject-headed restrictive relative clauses, such as:

*Un señor que (\*) era vecino mío y compañero ...* (Participant #15)

A señor that was a neighbor and companion ...

## (9) Existential verbs:

(\*) *Hay un muchacho afuera.*

There is a boy outside. (Participant #59)

(\*) *Habían guaguas para transportarlos.* (Participant #5)

There were buses to transport them.

(10) Verbs with impersonal *se*, such as:

(\*) *Se empezó con la cooperativa.* (Participant #9)

It began with the cooperative.

## (11) Verbs referring to time or weather conditions:

(\*) *Hace cien años, hace setenta años ...* (Participant #21)

One hundred years ago, seventy years ago.

*En Camuy (\*) hace un calor ...* (Participant #35)

In Camuy it is hot.

## (12) Discourse markers occurring obligatorily without a subject pronoun, such as:

(\*) *Digo, esa es mi opinión.* (Participant #7)

I say, that is my opinion.

*Nadie se dedicaba a sembrar, (\*) digamos, yautía, ñame, batata.* (Participant #10)

Nobody was dedicated to planting, let's say, cocoyam, yam, sweet potato.

Discourse markers employing the second person, familiar or formal, are included in the analysis, primarily because they did exhibit variation-yielding forms such as: *tú sabes* 'you (sing. fam.) know' and *sabes* '[you (sing. fam.)] know', *Ud. sabe* 'you (sing. formal) know' and *sabe* '[you (sing. formal)] know,' and *¿entiende?* '[you (sing. formal)] understand?' and *¿Ud. entiende?* 'you (sing. formal) understand?'

#### OVERALL RATES AND RATES FOR SINGULAR VERSUS PLURAL

As indicated, existing studies have found overall rates of subject personal pronoun expression in data based on Puerto Rican speakers in other locations to be high in comparison with the rates of more conservative dialectal regions, such as in central Spain or Mexico. It should be noted, nevertheless, that, even among Puerto Rican speakers who have been found to produce relatively high rates of subject pronoun expression, the rates do not reach the majority of the cases, a finding that may reflect the status of the null subject as the older, underlying form.

The findings presented here for the Castañer sample will be based on 2882 sites of potential subject personal pronoun expression, after excluding 118 cases of nonspecific third-person plural subjects. The rate of subject personal pronoun expression in these data from the interior of Puerto Rico is relatively low at 28%. Table 1 presents findings for the overt expression of subject personal pronouns for six samples of Puerto Rican Spanish, three samples from Castile and Madrid,

TABLE 1. Overall rates of subject personal pronoun expression by study

	[+Pro]	Total
Castile (Rosengren, 1974)	21%	23,890
Madrid (Enríquez, 1984)	22%	22,357
Madrid (Cameron, 1992)	21%	2061
PR, San Juan (Morales, 1986)	46%	12,182
PR, San Juan (Cameron, 1992)	45%	2122
PR, San Juan (Ávila-Jiménez, 1996)	40%	4713
PR, Boston (Hochberg, 1986)	40%	3019
PR, New York City (Flores-Ferrán, 2004)	45%	15,617
PR, New York City (Otheguy et al., 2007)	35%	3805
Castañer, Puerto Rico	28%	2882

and the finding for Castañer. The findings of five of the six Puerto Rican samples are in a 40% to 46% range: Morales (1986), Cameron (1992), and Ávila-Jiménez (1996) in San Juan; Hochberg (1986) in Boston; and Flores-Ferrán (2004) in New York City.<sup>3</sup> The study of Otheguy et al. (2007) in New York City reports a rate of 35% for overt expression for Puerto Rican speakers.

In Table 1, the rates of overt expression for central Spain are all in the low 20% range. Rosengren's (1974) study was based on a corpus of written Castilian Spanish in contemporary plays and Enríquez's (1984) study was based on a spoken corpus gathered in Madrid, whereas Cameron's (1992) study in Madrid was based on a sample that he developed based on transcribed interviews found in *El habla de la ciudad de Madrid* (Esgueva & Cantarero, 1981). The overall rate of 28% in Castañer is more similar to the rates of the Castilian studies than to those of the Puerto Rican studies in San Juan and on the mainland.

When rates of personal pronoun expression are broken down for usage with singular and plural verb forms, the findings from Castañer, once again, are lower than those of the other Puerto Rican varieties, particularly for use with singular forms. Table 2 shows that subject pronouns are expressed more with singular verbs than with plural verbs in all of the varieties for which data are reported here.

Although somewhat higher, the frequency of use with singular verbs in Castañer, which is 32%, is closer to the frequencies reported for Castile, which are 23%, 25%, and 26%, than to those of the other sampled Puerto Rican varieties. With one exception, the 17% expression rate for plurals in Castañer is also lower than the plural rates for the other Puerto Rican varieties, although the differences are small. This indicates that the difference with other Puerto Rican varieties lies primarily in the lower Castañer rate with singular verbs. The difference of 15 percentage points separating overt expression for singular versus plural verbs is, nevertheless, much smaller in Castañer than for the other Puerto Rican samples, where the difference is in the 30-percentage-point range, and it is more in line with the differences between singulars and plurals reported for Castile.

TABLE 2. *Subject personal pronoun expression for singular and plural verb forms by study*

	[+Pro]	Total
Castile (Rosengren, 1974)		
Singular	23%	20,647
Plural	11%	3243
Madrid (Enríquez, 1984)		
Singular	25%	16,584
Plural	10%	5773
Madrid (Cameron, 1992)		
Singular	26%	1504
Plural	7%	549
PR, San Juan (Morales, 1986)		
Singular	52%	9865
Plural	19%	2317
PR, San Juan (Cameron, 1992)		
Singular	50%	1764
Plural	19%	358
PR, San Juan (Ávila-Jiménez, 1996)		
Singular	48%	3655
Plural	15%	1124
PR, New York City (Flores-Ferrán, 2004)		
Singular	51%	12,645
Plural	20%	2985
PR, Boston (Hochberg, 1986)		
Singular	46%	2206
Plural	21%	813
Castañer, Puerto Rico		
Singular	32%	1971
Plural	17%	911

## GRAMMATICAL AND FUNCTIONAL CONSTRAINTS

In his discussion of subject personal pronoun expression in his own data from San Juan and Madrid dialects, Cameron (1994) concluded that, although rates of use distinguish the Madrid and San Juan dialects, similar patterns of pronominal expression are found not only in relation to grammatical constraints but also in functional, discourse-oriented conditioning. Cameron's remarks referred specifically to the grammatical categories of ambiguous verbs and unambiguous verbs and to the functional constraint of switch versus same reference with respect to the subject of the preceding sentence. Although a wider range of independent variables, including semantic class of verb, clause type, and priming effects, has been examined in association with personal pronoun subject expression, here I examine the grammatical variables of verb category in relation to ambiguity and person-number of the pronoun in question, and the functional variables of switch versus same reference and position in the discourse of information clarifying the subject. I have selected these variables for purposes of ready comparison with other varieties of Puerto Rican Spanish, particularly in the cases of the first three. I will show that, although lower rates distinguish Castañer usage in the interior of

TABLE 3. *Subject personal pronoun expression by verb class (Castañer data, singular verbs only)*

	%	Weight	Total
<i>Ser</i> singular	50%	.77	80
Preterit singular	27%	.55	759
Two-way ambiguity	34%	.58	666
Three-way ambiguity	37%	.66	464

*Note:* Factor groups included: verb class, person-number, same versus switch reference, and position of clarification in discourse.

Puerto Rico from the usage of San Juan and mainland varieties, here too patterning is often very similar. In one notable instance, however, where patterning in Castañer is different, it will be shown to be similar to patterning in other conservative varieties in Madrid and Castile, in central Spain.

A series of studies have examined subject personal pronoun expression in *-s*-deleting dialects with respect to potential ambiguity of verbs. In present indicative, present perfect, and future indicative forms, deletion of second-person singular *-s* creates potential two-way ambiguity with third-person singular forms. For example, the verb forms in (*él/ella*) *come* '(he/she) eats' and (*tú*) *come(s > Ø)* '(you) [sing. fam.] eat', (*él/ella*) *ha ido* '(he/she) has gone' and (*tú*) *ha(s > Ø)* *ido* '(you) [sing. fam.] have gone', and (*él/ella*) *trabajar*á '(he/she) will work' and (*tú*) *trabajar*á(*s > Ø*) '(you) [sing. fam.] will work' become indistinguishable. In a variety of forms that include the imperfect, the conditional, and both present and past subjunctives, upon deletion of second-person singular *-s*, potential three-way ambiguity results with first- and third-person forms. For example, in the imperfect the verb forms in (*yo*) *comía* '(I) used to eat', (*él/ella*) *comía* '(he/she) used to eat', and (*tú*) *comía(s > Ø)* '(you) [sing. fam.] used to eat' are indistinguishable. The distinctiveness of other verb classes including singular forms of *ser* 'to be' and singular preterits and all plural forms is not affected by final *-s*-deletion.

Table 3 shows findings given as both rates and Varbrul weights (Pintzuk, 1988) for singular verb classes in the data from Castañer. For Varbrul binomial analysis, progressively higher weights above .50 are progressively more favorable to the occurrence of a variant, progressively lower weights under .50 are progressively less favorable, and .50 is neutral.

In her pioneering study of subject personal pronouns in relation to verb classes in mainland Puerto Rican Spanish in Boston, Hochberg (1986) found rates of 48% for two-way ambiguity and 53% for three-way ambiguity, and in his subsequent study of San Juan speakers Cameron (1992) found rates of 50% and 55% for the same classes. In Table 3, the rates for two-way and three-way ambiguity are 34% and 37%, or markedly lower than those reported in the aforementioned studies. Nevertheless, the trend of factor weights is in the same direction. In other words, greater potential ambiguity favors a higher rate of usage. However, whereas



TABLE 4. *Subject personal pronoun expression by person-number (Castañer data)*

	%	Weight	Total
<i>Yo</i> (1SG)	34%	.61	1257
<i>Tú</i> (2SG)	47%	.67	40
<i>Ud.</i> (2SG)	32%	.40	71
<i>Él/ella</i> (3SG)	29%	.57	603
<i>Nosotros</i> (1PL)	17%	.29	496
<i>Ellos/as</i> (3PL)	17%	.33	415

*Note:* Factor groups included: person-number, same versus switch reference, and position of clarification in discourse. The verb category factor group has been excluded from this analysis because of interaction with the pronoun person and number factor group when both are run with same or switch reference and position of clarification in the discourse. The low Varbrul weight for *Ud.* reflects the low number of cases and the fact that most are found in situations of same reference.

Hochberg hypothesized the unambiguous verb forms promote lower rates of subject pronoun usage, Cameron (1992:198) found this to be true of preterit forms but not the singular forms of *ser*; in fact, he found the singular forms of *ser* to promote personal pronoun use most, a finding repeated here as well.<sup>4</sup>

Subject personal pronoun expression has also been described in correlation with the person and number of pronouns. Again, both rate differences and similarities with respect to several major constraints characterize the Castañer findings in relation to findings for other Puerto Rican varieties, as may be seen in Table 4.

The rates seen here for the *yo*, *tú*, and *él/ella* forms in the Castañer data are markedly lower than those reported by Cameron (1992:233) and Ávila-Jiménez (1996:97) for San Juan and Flores-Ferrán (2002:85) for New York City. The rates reported by these investigators are, for first-person singular, *yo*, 50%, 51%, and 52%, respectively, whereas the rate for the Castañer sample is 34%; for second-person singular familiar, *tú*, 60%, 61%, and 57%, whereas the rate here is 47%; and for third-person, *él/ella*, 39%, 38%, and 48%, respectively, whereas for the Castañer sample the rate is 29%.

Despite the difference in rates, there are similarities in the application of constraints connecting usage in Castañer and in these other Puerto Rican varieties. They include the higher rates for singular forms than for plurals, the pattern of favorability, *tú* > *yo* > *él/ella*, and the lower rates of expression for *Ud.* (second pers. sing., formal) than *tú*, reported by Avila-Jiménez, in San Juan, and Flores-Ferrán, in New York City.

Not coinciding with other varieties of Puerto Rican Spanish, however, is Castañer's usage in relation with the distinction between specific and nonspecific *tú*. For specific *tú* (as in *espero que tú vengas mañana* 'I hope you come tomorrow'), Cameron (1992:235), Ávila-Jiménez (1996:97), and Flores-Ferrán (2004:85) reported rates of 48%, 59%, and 53%, respectively. The rates they report for nonspecific *tú* (as in *es la carretera que tú sigues para llegar a Ponce* 'It is the highway that you take to get to Ponce') are even higher, 69%, 63%, and 60%, respectively. In the Castañer data, although the numbers are very

TABLE 5. *Subject personal pronoun expression by same and switch reference (Castañer data)*

	%	Weight	Total
Same	17%	.34	1618
Switch	42%	.70	1264

*Note:* Factor groups included: person-number, pronoun category, same versus switch reference, and position of clarification in discourse.

small, the trend appears to be reversed; the rate is 52% (17/33) for specific uses of *tú* and 29% (2/7) for nonspecific uses. This finding coincides with the direction of more favorability for specific versus nonspecific reported by Cameron (1994:324) for his Madrid sample as well as for Enríquez's Madrid sample. Although the difference is smaller and again, token counts are very low, the direction of favorability is the same for specific versus nonspecific uses of *Ud.* in the Castañer data, 33% (21/64) for specific *Ud.* and 29% (2/7) for nonspecific.<sup>5</sup>

With respect to subject expression in the context of the discourse, Flores-Ferrán (2004:61) observed that the constraint of switch versus same reference in relation to the subject of the preceding clause has been shown to condition subject personal pronoun expression across Spanish dialects. This is also the case in the findings from Castañer in west-central Puerto Rico, as is seen in Table 5.

In Table 5, same reference, or coreferentiality, with the subject of the preceding clause (as in [*Yo/Ø*] *hablo bien el español, pero el inglés [yo/Ø] lo hablo muy mal* '[I/Ø] speak Spanish well, but English [I/Ø] speak poorly') inhibits overt expression of the subject in the second clause. A change of subject (as in [*Yo/Ø*] *estoy en la cocina y [ella/Ø] contesta el teléfono* '[I/Ø] am in the kitchen and ([she/Ø] answers the telephone') promotes the overt identification of the switched subject. Once again, although the direction of the effect is the same, rates of expression are much higher in San Juan and in the Puerto Rican Spanish of New York City, at 31% and 57% for same and switched in San Juan (Cameron, 1992, 1994) and at 38% and 54% in New York City (Flores-Ferrán, 2004). The Castañer rates of 17% and 42% are more similar to the 11% for same reference and 30% for switch in Madrid (Cameron, 1992).

Although the rates of expression are very different, the Varbrul weights underscore the parallel effects of this factor in San Juan and Madrid, as reported by Cameron (1992), and Castañer. The probability weights for the overt expression of the personal subject pronoun by same and switch reference reported by Cameron for San Juan are .34 and .64, and .34 and .65 for Madrid; in Castañer, as reported in Table 6, they are .34 and .70.

Cameron (1992) also examined the effects of change of reference in relation to broader discourse context, both in reference chains and in terms of distance from the affected, or target, subject. He found that the number and location of subject changes in a chain including the affected subject and the two preceding subjects

TABLE 6. *Subject personal pronoun express by position of clarifying information in conversation (Castañer data)*

	%	Weight	Total
Sentence	12%	.30	609
Turn	26%	.50	809
Prior discourse	36%	.62	1011

*Note:* Factor groups included: verb class, person-number, same versus switch reference, and position of clarification in discourse.

affect the use of overt pronouns in the target. More changes over more positions result in increased use of pronouns (Cameron, 1992:247). Distance of “settlements,” or clarifying information, from the target subject in terms of numbers of clauses also affects overt pronoun use (Cameron, 1992:249). These findings are true for both Cameron’s data drawn from San Juan speakers and his data drawn from Madrid speakers, although the frequencies in specific contexts are always lower in the second case.

Here, I will examine the broader conversational context in terms of the positioning of information clarifying the referent of a focus subject, whether in the discourse prior to a current turn, in the current turn prior to the sentence in which a subject pronoun is expressed or not expressed, or in the target sentence itself. The target sentence, here, includes compound and complex sentence structures. In example (13), the speaker (C) [Participant #1], talking to me, refers to members of the Castañer family, owners of Hacienda Castañer, for which the town of Castañer is named. Hacienda Castañer actually consisted of several large farms.

- (13) C: ... *Los Castañeres tenían tres fincas, pues, fincas de miles de cuerdas. ¿Ve?*  
(Participant #1)

... The Castañers had three farms, well, farms of thousands of acres. You see?

H: *¿Las fincas tenían café, china, guineo, igual que hoy?*

Did the farms have coffee, oranges, bananas, like today?

C: *No antes. Cuando se establecieron bastante fuertes ellos, pues sembraron mucho plátano.*

Not before. When (they) established themselves sufficiently well they, well (they) planted a lot of plantain.

The clarification, or referent, of the subject of the verb *establecieron* ‘established’ is Castañeres, found as the subject in C’s earlier turn in the discourse, not in the turn or sentence in which *establecieron* occurs. The subject of *establecieron*, *ellos* ‘they’ is expressed, in this case, after the verb. The referent of the verb

*sembraron* ‘planted’, however, is the subject pronoun *ellos*, found in the adverbial clause of the same sentence, and the subject of *sembraron* is not expressed. Clarification is also found within the sentence in example (14).

- (14) *Poco a poco al dueño le quitaban los hombres fuertes un pedazo, cantito de ese terreno que tenía ...*

Little by little from the owner the strong men took a large part, a chunk of the land that (he) had.

In this case, the clarifying referent of the unexpressed subject of *tenía*, which is *dueño*, is not a subject, but the object of the preposition *a*, within the same sentence. In example (15), the clarification of the subject of *dieron* lies not in the same sentence nor in the discourse preceding the turn, but earlier in the turn, in the form of the object of the preposition *de* ‘of’, or *Ramos*, of the preceding sentence.

- (15) C: ... *el papá mío se había casado con una hermana de los Ramos. Después, pues, le dieron a ella un terrenito a ojo.*

My father had married a sister of the Ramos. Later, well, (they) gave her a small plot of select land.

When clarifying information occurs in the turn or the discourse outside of the sentence, it may occur in more extended discourse or turns than is indicated in the examples given here.

Table 6 shows that position in the conversation, or discourse, construed in this way is indeed relevant in the examination of the recorded conversations from Castañer.<sup>6</sup> Although the frequencies are all low as is expected for the Castañer data, the same sort of incremental increase found by Cameron for reference chains and distance is found here as well for clarifying information in increasingly more removed discourse positions: in the sentence, the turn preceding the sentence, and discourse preceding the turn. The progression from 12% to 26% and 36% is confirmed by the progression of Varbrul weights, from .30 to .50 and .62.

#### SOCIAL FACTORS

Although meaningful and statistically significant patterns are clearly present for linguistic constraints on the appearance of subject personal pronouns in the Castañer data, in other Puerto Rican varieties, and beyond, social constraints may be much more difficult to identify. In fact, Silva-Corvalán (2001:131) stated “*no ha sido fácil asignar valor social y/o estilístico a los casos de variación sintáctica estudiados en diversas comunidades. Más bien, la mayor parte de la variación sintáctica parece estar condicionada por factores sintácticos,*

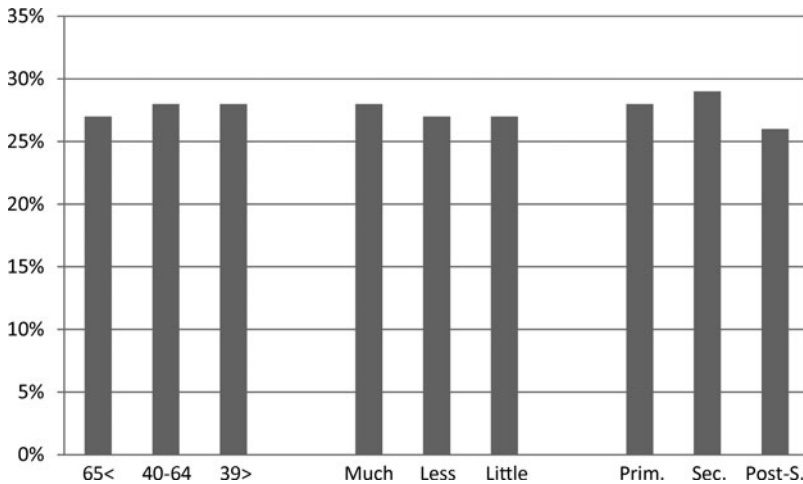


FIGURE 1. Subject personal pronoun expression by age, time in Castañer, and education.

*semánticos y pragmáticos ...* ‘it has not been easy to assign social and/or stylistic value to the cases of syntactic variation studied in diverse communities. Rather, the majority of syntactic variation appears to be conditioned by pragmatic, semantic, and syntactic factors ...’. This includes factors such as those reviewed herein. Flores-Ferrán (2002:642) indicated that findings of investigations of sociolinguistic stratification vary according to regional dialect and have not been consistent. For example, Cameron (1992) reported no systematic social effects on the overt and null categories in his study in San Juan, whereas Ávila-Jiménez (1996), in a study of speakers drawn largely from the same variety, found effects for age and occupation level. Morales (1986) did not find support for the hypothesis of English influence on subject expression among Spanish/English bilinguals, whereas Flores-Ferrán (2002) showed a mild effect, but contradictory findings are later documented in Flores-Ferrán (2004).

Rates of subject expression by social factors are shown in two figures for ease of exposition. Figure 1 shows age, time in Castañer, and level of education.

The findings for all of the subgroups hover around the level of 28% (793/2882), which is the overall rate of expression for the full sample, and Varbrul rejects each of these factor groups as a possible contributor to the probability of overt pronoun use. Growing up before, during, or after the period of the formation of the town (with rates of 27% [259/949], 28% [266/965], and 28% [268/968], respectively); time spent in the community as determined by a formula combining time of residence with time of absence (with rates of 28% [416/1483], 27% [142/531], and 27% [235/868] for “much,” “less,” and “little,” respectively); and education level (with rates of 28% [254/906], 29% [257/644], and 26% [282/793] for primary, secondary, and postsecondary, with postsecondary requiring periods of study on the coast) do not have significant effects.

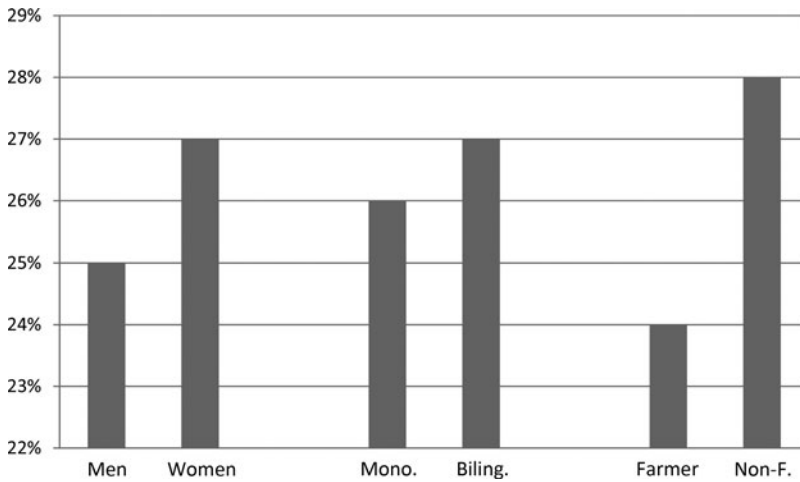


FIGURE 2. Subject personal pronoun expression by gender, language status, and farmer status.

Findings for speaker gender along with two additional factors are presented in Figure 2. The additional factors are the distinction between monolingual Spanish speakers and bilingual Spanish and English speakers as determined by self-reporting and confirmed by observation, and also the distinction between farmers and nonfarmers among the group of participants.

Once again, the findings for overt pronoun use are at or near the 28% level. Men have used pronouns in 26% (342/1305) of the eligible sentences and women in 29% (451/1126), and monolinguals in 27% (628/2303) and bilinguals in 28% (165/579), and Varbrul rejects both factor groups. The distinction between farmers and nonfarmers, however, yields a different result. Although the distinction is not great, with farmers expressing the personal pronoun in 25% (268/1062) of eligible sentences and nonfarmers in 29% (525/1820), it is statistically significant in the multivariate analysis. The farmer group, which combines hacienda owners, large and small farm owners, and farm laborers and includes both men and women in all categories, used the pronoun less often. In other words, the generally low rate of personal subject pronoun expression is even lower among farmers.

We may interpret the lack of distinctions among groupings for age, education, time in Castañer, gender, and bilingualism as suggesting a very high level of uniformity in the use of personal subject pronouns in this speaker sample from the community of Castañer, in west-central Puerto Rico. Nevertheless, the distinction that is statistically significant, showing that farmers employ subject personal pronouns even less often than the rest of the population do, may serve to link low usage directly to the land and the economic activity that traditionally have supported this dialectal region.

## CONCLUSION

Analysis of frequency rates in the use of overt as opposed to null subject personal pronouns in sampled speech from the community of Castañer in the western highlands of Puerto Rico shows that the rates are markedly lower than rates reported for other varieties of Puerto Rican Spanish, whether in the San Juan metropolitan area or on the U.S. mainland. A review of results of multivariate analysis also indicates, however, that grammatical and functional constraints in these data show the same direction of effect as in other varieties. Furthermore, the low rate of overt pronoun expression has not been found to be affected by social conditioning in any major way.

To interpret these findings, it may be possible to conclude that if, as Cameron (1994:319) suggested, synchronic differences in rates of variation across dialects reflect diachronic change within a language, and the Spanish dialects of Madrid and San Juan represent two stages, earlier and later, characterized by lower and higher rates of usage, a similar relationship may exist between the dialect of Castañer in the interior highlands and other dialects of Puerto Rican Spanish. The dialect of Castañer represents a stage at which the rate is lower, whereas the dialects of San Juan and also the U.S. mainland reviewed here represent a stage at which the rate of personal pronoun expression is higher.

This raises that question of why the speech in the rural interior, especially of western Puerto Rico, may be conservative with respect to subject pronoun expression. Is it geography? Why is the speech in this area more like the speech of central Spain with respect to this feature? Both Navarro Tomás (1948) and Álvarez Nazario (1991) suggested that the Spanish that was brought to the island and to the rest of the Americas by the original colonizers was Castilian in nature. Navarro Tomás (1948:28) also suggested that this was the language used for communication among speakers of other regional varieties from the Iberian Peninsula. Writing of the consonant system, Álvarez Nazario (1991:65) said that

*... el español que se trae a las Antillas y a la Tierra Firme próxima al mar Caribe durante la primera mitad del siglo XVI se ajustaría en general a los patrones de pronunciación de ... "una lengua común de tipo castellano," "con clara articulación de los finales, con ll, y por supuesto con la -d- intervocálica mantenida," pero a la luz ya portadora de ciertas transformaciones que habrían de difundirse muy pronto en ambas orillas del Atlántico ...*

... the Spanish that is brought to the Antilles and to the mainland close to the Caribbean during the first half of the 16th century would conform in general to the pronunciation patterns of ... "a common language of Castilian type," "with clear articulation of final sounds, with ll, and of course with intervocalic -d- preserved," but already clearly barer of certain transformations that very soon would spread on both sides of the Atlantic ...

One may imagine, then, that accompanying the grammatical and phonological features of Castilian, or of Castilian showing some of the early transformative influences from southern Spain, was the system of subject pronoun usage underlying that of the Spanish of central Spain today. This would have been the system that underlay the pronoun usage of early colonizers' descendants, who inhabited the *vegas* 'meadows' of coastal areas of the island. Writing of the interior of the island in the mid to late 1800s, historian Fernando Picó (1985:21) suggested that it was inhabited, above all, by descendants of the original colonizers, "*los criollos de nuestras costas ... desplazados por el crecimiento de las haciendas azucareras*" 'the creoles of our costs ... displaced by the growth of the sugar haciendas'. One might propose then that, displaced by sugar plantations and new waves of immigrants with heavy representation from innovative dialectal regions of southern Spain and the Canary Islands, small farmers and descendants of original settlers, like the indigenous *taíno* population before them, moved to the interior highlands and carried with them their more traditional form of Spanish. As Navarro Tomás (1948:114) suggested, "*En los moderados y suaves giros del habla jíbara sobreviven ... huellas de dejos y cadencias de tradición peninsular*" 'In the gentle and modulated turns of rural mountain speech ... traces and cadences of peninsular tradition survive'. In spite of a currently high 69% rate of *s*-deletion (Holmquist, 2011), today these cadences and traces may incorporate elements of archaic vowel and consonant usage that occur variably in the region, and the low frequencies and patterns of subject pronoun expression described here that appear, even today, to be much more uniformly present. This may mean as well that it was in coastal areas, where new arrivals came into contact with original settlers and speakers of other dialects and languages in plantation society, that the more heavy reliance on personal pronouns to mark and identify subjects evolved.

## NOTES

1. Not all researchers agree with this exclusion. Some argue that these contexts do allow for variation, for example, Matos Amaral and Schwenker (2005:125) and Otheguy et al. (2007:775–776).
2. A nonspecific *ellos*, or personal subject pronoun, has been considered by some investigators to be possible in some nonspecific third-person plural contexts. For example, if a speaker had said *En la costa, ellos pasan más calor* 'On the coast, they experience more heat', the hearer might envision a group of people from the coast and perhaps the referent would be less nonspecific, but it could be interpreted as nonspecific because it would not be referring to *x*, *y*, and *z*. Uses of this type, however, were not found in the data, and this is considered here to be an obligatorily null context.
3. For Ávila-Jiménez's (1996) speaker sample, the largest representation is drawn from the San Juan metropolitan area. The sample also includes representation from other areas of the north coast and very small numbers of speakers from other points on the island. The overall findings that I have presented for Morales (1986), Cameron (1992), and Ávila-Jiménez (1996) in Table 1 are based on findings they report for individual pronoun categories broken down by person and number. The same is true for the findings broken down by singular and plural for these investigators in Table 2. In the cases of Cameron (1992) and Ávila-Jiménez (1996), findings reported for small numbers of impersonal *uno* have been excluded.
4. It should be noted that Silva-Corvalán (2001:160–162) proposed an alternative explanation for the progression of less to increasingly more favorability seen in the progression from nonambiguous preterit forms to forms offering the possibility of two-way ambiguity, often present forms, and forms offering three-way ambiguity, including the imperfects. She found the progression to be present as well in Los Angeles Mexican-American Spanish, a non-*s*-deleting dialect, and Cameron (1994) found the pattern to be one of those shared by Madrid dialects, traditionally also non-*s*-deleting. Silva-Corvalán



suggested that the preterits highlight verbal action and de-emphasize the actor, therefore, prompting less pronoun use. She suggested that presents represent action that is factual but less dynamic and promotes more attention to subject, seen in more frequent use of pronouns. She also suggested that the action of imperfects, subjunctives, and conditionals is typically background action, less dynamic, and most descriptive of subject, which is expressed most frequently for this category.

5. One element of the findings seen here for specific versus nonspecific *tú* does not coincide with a very specific prediction made by Cameron (1996:98). Limiting his observation to *tú*, Cameron predicted that dialects with a relatively high overall, or combined, rate of 35% for the expression of specific and nonspecific *tú* will have a higher rate for nonspecific *tú* than for specific *tú*. Here, there is a relatively high rate, 38% (42/111), of expression for the combination of specific and nonspecific *tú* and *Ud.*, that is, second-person singular usage; nevertheless, the rate for the specific application is higher in both cases. Caveats that Cameron gave for his prediction include that there should be at least 100 second-person *tú* subjects included in the sample. In Table 4, there are 111 cases for the combination of specific and nonspecific *tú* and *Ud.* Cameron also indicated that, for purposes of his prediction, the *tú* forms should not include discourse markers, a category that has been included here, providing a circumstance that may have influenced this outcome.

6. For the position of clarifying information factor group, a fourth category was included in the analysis. It is context and includes instances for which clarifying information was not found in the conversation itself. The percentage rate for overt expression with clarification depending on context is 32% and the Varbrul weight is .50.

## REFERENCES

- Álvarez Nazario, Manuel. (1974). *El elemento afronegroide en el español de Puerto Rico: Contribución al estudio del negro en América*. San Juan: Instituto de Cultura Puertorriqueña.
- \_\_\_\_\_. (1977). *El influjo indígena en el español de Puerto Rico*. Río Piedras: Editorial Universitaria, Universidad de Puerto Rico.
- \_\_\_\_\_. (1982). *Orígenes y desarrollo del español en Puerto Rico (Siglos XVI y XVIII)*. Río Piedras: Editorial de la Universidad de Puerto Rico.
- \_\_\_\_\_. (1991). *Historia de la lengua española en Puerto Rico: Su pasado y su presente en el marco de la realidad social*. San Juan: Academia Puertorriqueña de la Lengua Española.
- Ávila-Jiménez, Barbara. (1996). Subject pronoun expression in Puerto Rican Spanish: A sociolinguistic, morphological, and discourse analysis. Ph.D. dissertation, Cornell University.
- Cameron, Richard. (1992). Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation. Ph.D. dissertation, University of Pennsylvania.
- \_\_\_\_\_. (1994). Ambiguous agreement, functional compensation, and non-specific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5:305–334.
- \_\_\_\_\_. (1996). A community based test of a linguistic hypothesis. *Language in Society* 31:61–111.
- \_\_\_\_\_. (2005). Aging and gendering. *Language in Society* 34:23–61.
- Enríquez, Emilia. (1984). *El pronombre sujeto personal en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
- Esgueva, M., & Cantarero, M. (eds.). 1981. *El habla de la ciudad de Madrid: Materiales para su estudio*. Madrid: Consejo Superior de Investigaciones Científicas, Instituto Miguel de Cervantes.
- Flores-Ferrán, Nydia. (2002). *Subject personal pronouns in Spanish narratives of Puerto Ricans in New York City*. Munich, LINCOM EUROPA.
- \_\_\_\_\_. (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.
- Hochberg, Judith. (1986). Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62(3):609–621.
- Holmquist, Jonathan. (1998). High lands—high vowels: A sample of men’s speech in rural Puerto Rico. In C. Paradis (ed.), *Papers in sociolinguistics: NWAV 26 a l’Université Laval*. Quebec: Université Laval. 73–79.
- \_\_\_\_\_. (2003). Coffee farmers, social integration and five phonological features: Regional socio-dialectology in west-central Puerto Rico. In L. Sayahi (ed.), *Selected proceedings of the First Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 70–76.
- \_\_\_\_\_. (2005). Social stratification in women’s speech in rural Puerto Rico: A study of five phonological features. In L. Sayahi (ed.), *Selected proceedings of the Second Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 109–119.

- \_\_\_\_\_. (2008). Gender in context: Features and factors in men's and women's speech in rural Puerto Rico. In M. Westmoreland & Juan Tomas Testa (eds.), *Selected proceedings of the Fourth Workshop on Spanish Sociolinguistics*. Somerville, MA: Cascadilla Proceedings Project. 17–35.
- \_\_\_\_\_. (2011). Gender and variation: Word-final /s/ in men's and women's speech in Puerto Rico's western highlands. In M. Diaz-Campos (ed.), *Handbook of Hispanic sociolinguistics*. London: Blackwell. 230–243.
- López Morales, Humberto. (1983). *Estratificación social del español de San Juan de Puerto Rico*. México, D.F.: Universidad Nacional Autónoma de México.
- Matos Amaral, Patricia, & Schwenter, Scott A. (2005). Contrast and the (non-)occurrence of subject pronouns. In D. Eddington (ed.), *Selected proceedings of the 7th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project. 116–127.
- Morales, Amparo. (1986). La expresión del sujeto pronominal en el español de Puerto Rico. In A. Morales (ed.), *Gramáticas en contacto: Análisis sintácticos sobre el español de Puerto Rico*. San Juan, Puerto Rico: Editorial Playor. 89–100.
- Navarro Tomás, Tomás. (1948). *El español en Puerto Rico*. Río Piedras: Editorial Universitaria.
- Otheguy, Ricardo, Zentella, Ana Celia, & Livert, David. (2007). Language and dialect contact in New York: Toward the formation of a speech community. *Language* 83(4):770–802.
- Picó, Fernando. (1985). *Amargo café*. Río Piedras: Ediciones Huracán.
- Pintzuk, Susan. (1988). *Varbrul programs for MS Dos*. Philadelphia: University of Pennsylvania.
- Poplack, Shana. (1980). Deletion and disambiguation in Puerto Rican Spanish. *Language* 56:371–385.
- Rosengren, Per. (1974). *Presencia y ausencia de los pronombres personales sujetos en el español moderno*. Stockholm: Almqvist & Wiksell.
- Silva-Corvalán, Carmen. (1977). A discourse study of word order in the Spanish spoken by Mexican-Americans in West Los Angeles. M.S. thesis, University of California, Los Angeles.
- \_\_\_\_\_. (2001). *Sociolingüística y pragmática del español*. Washington, D.C.: Georgetown University Press.
- Torres, Lourdes (2002). Bilingual discourse markers in Puerto Rican Spanish. *Language in Society* 31:65–83.