ONOMÁZEIN
Revista de lingüística, filología y traducción

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE LETRAS

Variable subject pronoun expression in Andean Spanish: a drift from the acrolect

Álvaro Cerrón-Palomino
Arizona State University
United States of America

ONOMÁZEIN 42 (diciembre de 2018): 53-73
DOI: 10.7764/onomazein.42.02
ISSN: 0718-5758

Álvaro Cerrón-Palomino: School of International Letters and Cultures, Arizona State University, USA.
| Correo electrónico: acerropa@asu.edu

Fecha de recepción: marzo de 2017
Fecha de aceptación: agosto de 2017

Abstract

This study is one of the first variationist accounts of subject pronoun expression (SPE) in Andean Spanish, in particular, in the variety spoken in Huancayo, Peru. The data analyzed consist of sociolinguistic interviews of sixteen participants, equally divided by sex.

The results show that the overt subject personal pronoun (SPP) rate in Huancayo Spanish is the lowest one attested: 16.2%, even lower than that of Lima, the capital (16.8%), which situates Andean Spanish among conservative varieties regarding SPE.

In order to determine if *Huancaínos* are following *Limeños* as a model in their SPE use and therefore sharing the same constraints, I conducted two separate regression analyses: one for Huancayo and one for the Limeño acrolect using Goldvarb X. The results show that Lima's dialect is sensitive to two factor groups that are neutralized in Huancayo, which suggests that *Huancaínos* do not take *Limeños* as a role model for their overt SPP production.

Keywords: Andean Spanish; comparative sociolinguistics; subject expression; language prestige.

1. Introduction

In Spanish, variable subject pronoun expression (SPE), i.e. the alternation between overt subject personal pronouns (SPPs) and null subjects, has been studied by a number of researchers in different varieties, to the extent that it has been categorized as the showcase variable of Spanish sociolinguistics (Bayley et al., 2012: 49-50). In these studies, SPE has been found to be robustly constrained by morphosyntactic, pragmatic and processing factors such as grammatical person and number, subject referential switch and priming, whereas verb class, tense/aspect/mood (TAM) endings and reflexivity of the verb have been less constant predictors of SPP production.

On the one hand, studies of diverse monolingual varieties show two well-established patterns for monolingual varieties: a) Latin American varieties display higher overt SPP rates than Peninsular varieties, and b) Caribbean varieties exhibit higher percentages of overt SPPs than mainland Latin American varieties (Orozco and Guy, 2008: 70). On the other hand, in analyses of bilingual Spanish communities, concretely in varieties in contact with English, some studies suggest that SPE in Spanish is affected by the mandatory subject expression of English (Lapidus Shin and Otheguy, 2005; Otheguy and Zentella, 2012), while others suggest that there is no such influence (Silva-Corvalán, 1994; Torres Cacoullos and Travis, 2010; Á. Cerrón-Palomino, 2016).

In variationist studies of variable SPE in Spanish varieties in contact with languages other than English, some of the analyses offer evidence of language contact effects on SPE (Barnes, 2010; Michnowicz, 2015; Á. Cerrón-Palomino, 2018), whereas others do not (Carvalho and Child, 2011; Carvalho and Bessett, 2015; De Prada Pérez, 2015).

With the purpose of shedding more light on the research on Spanish SPE, this study is motivated by the following questions: a) how does the overt SPP rate of Huancayo Andean Spanish compare to the rest of the Hispanic world? b) what are the linguistic factors constraining SPE in Huancayo Spanish?, and c) is Huancayo Spanish SPE modelled after the prestigious *Limeño* variety?

This paper is organized as follows. In section 2, I characterize the variety of Spanish spoken in Huancayo. I explain the methodology followed in section 3, and the factors and predictions analyzed in section 4. Section 5 is devoted to the results of the variationist analysis, and section 6 compares the linguistic constraints operating on SPE in Huancayo with those conditioning SPE in the Peruvian acrolect spoken in Lima. Finally, I present the concluding remarks to this study.

2. Huancayo Spanish

Huancayo is a city in the Central Andes of Peru, capital of the Junin region. The city is located in the Mantaro Valley, and it is the commercial heart of the entire Central Andes Region

as well as the main food source for Lima, the country's capital. Its population consists of 340,000 to 380,000 inhabitants. According to the 2007 National Census (the last one), 87.7% of the Huancayo population spoke Spanish, whereas 12% spoke Quechua, and 0.3% spoke other native tongues such as Aymara or the Amazonian language Ashaninka.

The Spanish spoken in the city is characterized as a variety of Andean Spanish, a macrodialect used across the Andes Mountains region, from the Colombian southwest to the northeast of Chile and Argentina, but the majority of its speakers concentrates in Bolivia, Ecuador and Peru. An important difference between the status of Andean Spanish in these three countries is that it is spoken in the capitals of Bolivia and Ecuador, where it constitutes the acrolect, whereas it is not spoken in the capital of Peru, where a variety of *ribereño* Spanish is spoken, which represents the Peruvian acrolect (Escobar, 2016: 354). I will return to this subject in section 6.

Andean Spanish possesses a number of features which, taken all together, set this variety aside from most Spanish dialects. Some of these characteristics are not due to contact with Quechua. At the phonetic/phonological level, some contact-free characteristics of Andean Spanish are the maintenance of the palatal lateral approximant phoneme /ʎ/, the assibilation of /r/, the strengthening of intervocalic /x/, the maintenance of coda /s/, and the reduction of unstressed vowels (Escobar, 2016). At the morphosyntactic level, there are Spanish-internal features like double possession (Merma, 2004) and *leísmo* (Godenzzi, 1986).

With respect to features ascribed to contact with Quechua, they are more numerous at the morphosyntactic level, such as the use of direct object clitics with intransitive verbs, and the use of the verb *dice* as a referential validator (Merma, 2004). Some other Andean Spanish features attributed to contact with Quechua are the use of the direct object clitic *lo* regardless of the gender and number of its co-referential noun phrase (Godenzzi, 1991). There are also some syntactic calques such as noun phrase gender and number disagreement, preverbal objects and adverbials, and subordinate clauses preceding main clauses (R. Cerrón-Palomino, 2003). Finally, there are also some pragmatic markers of Andean Spanish, like the use of *pues* for clarification or confirmation, and the use of diminutives to signal modesty or deference (Escobar, 2016).

With respect to a possible influence of Quechua in Andean Huancayo Spanish SPE, Á. Cerrón-Palomino (2018) compared the production of overt SPPs in Spanish monolinguals and Quechua-Spanish bilinguals, and found that the rates for both groups were comparable, with no statistically significant difference. By the same token, Á. Cerrón-Palomino (2018) reports that the factor groups favoring overt SPPs in both groups are exactly the same, which suggests no influence from Quechua. However, at the factor group ranking level, Á. Cerrón-Palomino suggests that the subordinating switch reference marker *-pti* that Quechua possesses boosts the use of overt pronominal subjects in the bilinguals in contrast to the monolinguals, which could be interpreted as a case of indirect transfer.

Since Andean Spanish varieties are similar at their core (Escobar, 2016: 354), the study of SPE in a Central Andes variety may shed light on how this variable is constrained at the Andean Spanish macro-dialectal level. Moreover, the study of SPE in Huancayo may offer valuable insight on the regional distribution of patterns affecting the production of overt SPPs proposed by Orozco and Guy (2008: 70), presented in the introduction.

3. Methodology

In this section, I address the nature of the data, the data elicitation technique, the characteristics of the participants of the study, and the envelope of variation of the variable analyzed.

3.1. Data and participants

This is a quantitative sociolinguistic study which follows the variationist framework as outlined by Tagliamonte (2006); therefore, the main eliciting technique consisted of interviews ranging from thirty minutes to one hour.

Sixteen speakers of the Spanish variety of Huancayo, recruited through the friend-of-a-friend/snowball sampling, were interviewed in May-June 2012, following the sociolinguistic interview model (Tagliamonte, 2006), which consists of question modules about the participants' childhood, school, hobbies, etc. The sample consists of eight men and eight women who had lived all or most of their lives in Huancayo; their ages range from 22 years of age to 78.

3.2. Envelope of variation

Following the variationist methodology (Tagliamonte, 2006), only contexts in which the two SPE variants could occur were included in the study. The variable under scrutiny is further circumscribed to include only instances of tensed verbal forms with a human subject referent, expressed or null, as in (1).

(1) Pero es que cuando **tú** los compras pequeñitos, **Ø** tienes que hacerlos crecer.
'But (it) happens that when **you** buy them at a young age, **(you)** have to raise them.'[1]
(HC07F25S)

1 A pronominal subject in parentheses in the translation indicates the occurrence of a null subject in the original Spanish utterance.

Contexts in which no overt SPPs appeared in the data were excluded from the study, such as subject relative clauses and non-specific 3rd person plural subjects.

The selection of utterances thus described resulted in a total of 2258 tokens of Huancayo Spanish analyzed in this study. The data were coded and then entered into the Goldvarb X multivariate software (Sankoff et al., 2005), which allows to address relationships between the dependent linguistic variable and multiple independent variables (factor groups) through logistic regressions, reporting the outcomes as proportion-like probabilities which range from zero to one.

4. Factors analyzed

In this section, I will introduce the factor groups included in the present analysis, all of them of linguistic/internal nature, and the effects that the factors/constraints within them are expected to exert on the variable under study. Overall, I expect the SPP rate of Huancayo to be close to the 16.8% found in Lima by Á. Cerrón-Palomino (2014), since the capital's variety is the most prestigious one and serves as a model for the entire country (Caravedo, 2014: 292).

Following the ample variationist literature on variable SPE in Spanish (Silva-Corvalán, 1994; Travis, 2005, 2007; Otheguy and Zentella, 2012; Lastra and Martín Butragueño, 2015), a total number of six internal (linguistic) factor groups were incorporated in this analysis: grammatical person and number, verb class, discourse connection, ambiguity of the TAM ending, previous realization of a SPP (priming), and reflexive use of the verb. In the following lines, they are described in more detail.

4.1. Grammatical person/number

As Carvalho, Orozco and Shin (2015: xiv) have indicated, there is variation in the verbal suffixes of person and number favoring overt SPPs across dialects, which results in a single main generalization: singular verbal forms favor the overt variant more than their plural counterparts. Therefore, I expect for this pattern to be followed in Andean Huancayo Spanish.

4.2. Verb class

In this study, I followed Travis' (2005, 2007) verb semantic classification, which groups verbs in five categories: psychological, like *saber* (know), *creer* (believe), *pensar* (think); speech acts, like *decir* (say), *llamar* (call); copulas, like *ser, estar* (be); motion, like *ir* (go), *venir* (come); and other, which encompasses verbs not suitable in the previous groups.

I expect, following most SPE studies (Carvalho, Orozco and Shin, 2015: xv), psychological verbs to favor the presence of SPPs significantly, in contrast to the rest.

4.3. Referential connection (same/switch reference)

Overt SPPs have shown to be favored when there is a referential switch of the second of two consecutive subjects (Travis, 2005: 329), as in the differently indexed subjects of *son* and *soy* in (2).

(2) **Ellos**$_i$ ya son finados. **Yo**$_j$ soy huérfano de padre y madre, como se dice.
'**They**$_i$, are already gone. **I**$_j$ am an orphan to mother and father, as (they) say.'
(HC01M69S)

In accordance with previous studies on Spanish SPE (Silva-Corvalán, 1994; Abreu, 2012; Otheguy and Zentella, 2012; Alfaraz, 2015), I expect that a change in the reference of the subject will also favor the presence of overt SPPs, whereas continuity of reference will promote the use of null subjects.

4.4. Ambiguity of the tense/aspect/mood (TAM) form

In Spanish, some TAM endings stand in a homophony relation with others, e.g., the 1st person singular, the deferential 2nd person singular (*usted*) and the 3rd person singular of the imperfect indicative. In some dialects, ambiguous TAM endings have proven to be statistically significant in promoting the use of overt SPPs (Cameron, 1993; Á. Cerrón-Palomino, 2014; Lastra and Martín Butragueño, 2015; De Prada Pérez, 2015). It is thus hypothesized that ambiguous TAMs will favor overt SPPs to clarify the referent of their subjects, as opposed to non-ambiguous ones.

4.5. Previous realization of an overt SPP (priming)

Cameron and Flores-Ferrán (2004) characterize priming as the lingering over time of a previously uttered form. In this study, I consider as part of the process not only co-referential subjects of immediately consecutive verbs, but also co-referential subjects of verbs across intervening clauses (Travis, 2007). Thus in (3), the overt SPP of the verb *estaba* is the trigger of the overt SPP of the verb *estudiaba*, which is its target. By the same token, the overt SPP of the verb *salía* is the target of the overt SPP of the verb *estudiaba*, and at the same time it is the trigger of the null subject of the verb *tomaba*, where the overt SPP priming effect stops.

(3) Cuando **yo**$_i$ estaba en la escuela, y eso que **yo**$_i$ estudiaba aquí en Huancayo, **yo**$_i$ salía cinco y cuarenta y cinco, a mi casa, que **Ø**$_i$ tomaba carro a las seis.
'When **I**$_i$ was in school, and **I**$_i$ even studied here in Huancayo, **I**$_i$ used to leave at five forty-five headed home, such that (**I**$_i$) would take the bus at six.'

(HC05F28S)

Studies including syntactic priming in their analyses (Abreu, 2012; Cameron and Flores-Ferrán, 2004; Carvalho and Child, 2011; Orozco, 2015; Travis, 2005 and 2007; Torres Cacoullos and Travis, 2010; Á. Cerrón-Palomino, 2016) have shown that previously mentioned overt SPPs promote the presence of more co-referential overt SPPs. Because of the consistency of the priming effect on variable SPE attested in the literature, the prediction for this factor group is that the priming effect will also take place in the Huancayo variety.

4.6. Reflexive use of the verb

In some studies (Carvalho and Child, 2011; Michnowicz, 2015; Otheguy et al., 2007), it was found that verbs used reflexively as *paso* in (4), favored null subjects; in contrast, verbs lacking the reflexive pronoun promoted the use of more overt SPPs, like *entro*.

(4) **Yo** entro a las ocho y **Ø** me paso hasta las dos de la mañana haciendo un trabajo.
'**I** get in at eight and **(I)** stay until two in the morning working on homework.'

(HC06M22S)

Reflexive verbs favor null subjects because the co-referential object clitic accompanying the verb eases the speaker's task of identifying the subject's referent; in contrast, non-reflexive verbs prompt speakers to use SPPs to make the subject's referent more easily retrievable. I hypothesize that the Huancayo variety will follow the aforementioned pattern.

5. Results and discussion

Table 1 shows the distribution of the variants studied in this paper: 16.2% of the verbs analyzed have overt SPPs in Andean Huancayo Spanish, in contrast with the 83.8% of null subjects produced. In order to answer my first research question about where Andean Spanish's SPP rate stands in comparison to other communities studied, I present table 2.

As table 2 shows, the rate of Huancayo is the lowest among all of the communities studied, patterning roughly with Mexican, Peninsular and Peruvian varieties. Moreover, the Peruvian communities exhibit the lowest overt SPP rates recorded, which demands a reformulation

TABLE 1
Distribution of pronominal and null subjects in the Spanish of Huancayo, Peru

Variant	N	%
+ Pro	365	16.2
– Pro	1893	83.8
Total	2258	

of the generalization made by Orozco and Guy (2008: 70) that Peninsular varieties show lower rates than continental American varieties. In this respect, assuming that Cameron (1993) and Holmquist (2012) are right and a lower rate of SPPs evidences the conservative nature of a variety, Andean Spanish seems to be amongst the most conservative dialects. This clustering of Mexican, Peninsular and Peruvian varieties regarding their overt SPP rate was pointed out by Á. Cerrón-Palomino (2014), who suggests it is presumably due to the tight connections between Spain and its two main viceroyalties in America, Mexico (New Spain) and Peru, in contrast with the rest of its colonies.

Moving on to my second research question, I ran a regression analysis using Goldvarb X to determine which constraints favor overt SPE production in Huancayo. Table 3 displays the results of the analysis comprising the factors that contribute to the occurrence of the overt variant. As it was mentioned in section 3.2, the probabilities that Goldvarb reports range from zero to one, but the general convention in interpreting such results is that a factor weight above .5 favors the variant described—in this case, overt SPP occurrence—whereas weights lower than .5 disfavor its occurrence, and a weight of .5 has no effect on the variant. The ranges state the relative strength of each factor group.

Table 3 shows that the 2nd person singular and 1st person singular are the strongest predictors of overt SPPs in this variety, while a subject referential switch is the second strongest predictor, and a previous realization of a co-referential pronoun, i.e. priming, is the third and last one.

In what follows, I will address individually the internal factors that contribute significantly to overt SPP use in the Spanish of Huancayo. Regarding the most reliable predictors of the overt variant, table 4 shows that verbs inflected for 2nd person singular and 1st person singular were the ones favoring overt SPPs with probabilities of .61 and .56, respectively. This study confirms the claim by Carvalho, Orozco and Shin (2015: xiv) that no specific grammatical person favors overt SPPs universally, but that verbs inflected in singular favor more overt SPPs than plural verbs.

The second strongest linguistic constraint of overt SPPs is a change of the referent of the subject of two successive verbs; conversely, when the subject referent in two consecutive

TABLE 2
Overt SPP Rates by City/Country

City/origin of speakers	Rate
San Juan, Puerto Rico (Cameron, 1993)	44.7%
Santo Domingo (Alfaraz, 2015)	42.3%
Dominicans in New York (Otheguy and Zentella, 2012)	41%
Puerto Ricans in New York (Otheguy and Zentella, 2012)	39%
Isabela, Puerto Rico (Abreu, 2012)	39%
Cubans in New York (Otheguy and Zentella, 2012)	38%
Santiago, Chile (Cifuentes, 1980, quoted by Silva-Corvalán, 2001)	38%
Baranquilla, Colombia (Orozco and Guy, 2008)	35.7%
Colombians in New York (Otheguy and Zentella, 2012)	32%
Buenos Aires (Soares de Silva, 2006, quoted by Carvalho and Bessett, 2015)	29%
Ecuadorians in New York (Otheguy and Zentella, 2012)	28%
Castañer, Puerto Rico (Holmquist, 2012)	28%
Mexican-Americans in San Antonio (Bayley et al., 2012)	27%
Chipilo, Mexico (Barnes, 2010)	26%
Mexican-Americans in Los Angeles (Silva Corvalán, 2001)	25.1%
Rivera, Uruguay (Carvalho and Bessett, 2015)	25%
Jalapa, Mexico (Orozco, 2016)	24.8%
Puente Genil, Spain (Ranson, 1991)	24%
Mexicans in New York (Otheguy and Zentella, 2012)	22%
Mexico City (Lastra and Martín Butragueño, 2015)	21.7%
Castilla, Spain (Rosengren, 1974, quoted by Holmquist, 2012)	21%
Madrid, Spain (Cameron, 1993)	20.9%
Yucatan, Mexico (Michnowicz, 2016)	19.7%
Mexican-Americans in Phoenix (Author, 2016)	17.8%
Granada, Spain (Manjón-Cabeza et al., 2016)	17.5%
Lima, Peru (Á. Cerrón-Palomino, 2014)	16.8%
Huancayo, Peru (this study)	16.2%

TABLE 3
VARBRUL analysis of the linguistic factors contributing to the occurrence of overt pronominal subjects in the Spanish of Huancayo, Peru

Corrected mean	0.137
Log likelihood	-916.552
Significance	0.000
Total N	2258

	Factor Weight	%	N
Person and number			
tú + usted	.61	23.8	164
yo	.56	18.9	1223
él + ella	.49	13.1	289
ellos + ellas	.48	13.4	186
nosotros + nosotras	.29	7.3	381
Range	***32***		
Referential connection			
switch reference	.63	22.2	1226
same reference	.35	9	1032
Range	***28***		
Previous realization (priming)			
pronominal	.68	29.2	359
not pronominal	.47	13.7	1899
Range	***21***		

verbs remains the same, a null subject tends to occur. In this respect, the Huancayo variety patterns with all of the dialects accounted for in the literature.

The last predictor of the overt variant is priming, the prior occurrence of a co-referential overt SPP. These results are consistent with most studies that have included priming as a factor, regardless of the person/number analyzed.

In fact, the factors mentioned above are the three most robust cross-dialectal predictors of overt SPPs, as mentioned in the introduction, and they have shown a similar effect on SPE in virtually every variety in which they have been examined.

In contrast, the semantic class of the verb, the reflexive use of the verb, and the TAM ambiguity, absent from the table, turned out to have no effect on the variable studied. Not surprisingly, these factor groups have not been consistent across varieties, and can be therefore considered moderate cross-dialectal SPP predictors of the overt variant.

6. SPE in Andean Spanish vis-à-vis *Limeño* Spanish

The third research question of this study examines the possibility that Andean Spanish speakers follow the patterns of SPE used in Lima. It is well-known that the most prestigious variety of Spanish in Peru is the one spoken in Lima, the country's capital (Caravedo, 2014). In contrast, Andean Spanish is held in low esteem by *Limeño* speakers. On the one hand, Andean Spanish speakers are fully aware of this stigmatization of their variety; on the other, they perceive the Lima acrolect as more prestigious than their own (Caravedo, 2014: 289-296).

In this respect, R. Cerrón-Palomino (2003) hypothesizes that the linguistic features that are most stereotypically attributed to Andean Spanish are the ones that will be removed more promptly by its users, whereas the ones that escape the perception of the average *Limeño* should be more resilient to change. In this regard, phonetic-phonological variables should be more prone to change than morphosyntactic ones, since they are cognitively more prominent than the latter.

The negative view that *Limeños* have of Andean Spanish (Caravedo and Rivarola, 2011) has in fact caused bilingual and monolingual speakers of Andean Spanish, who have migrated to Lima, to suppress some of the distinctive features of the variety, in particular, the palatal lateral phoneme /ʎ/ and the assibilated pronunciation of /r/, and to adopt coda /s/ elision, originally absent from Andean Spanish (Klee and Caravedo, 2006).

Furthermore, Caravedo (1999) had already noted that some of these stigmatized features, in particular, the palatal lateral /ʎ/ was being lost among speakers of Andean Spanish who had not migrated to Lima. This accommodation of speakers of Andean Spanish to *Limeño* Spanish without moving to Lima is promoted by the media, in particular television, which takes place exclusively in Lima headquarters. Even local radio stations undergo a process of de-Andeanization of their Spanish, as witnessed by myself when I interviewed two local radio commentators who had copied the Lima phonetic and phonological features to the extent that their interviews had to be excluded from the study, due to lack of representativeness. It is not far-fetched, then, to suggest that *Huancaínos* could adapt their SPE patterns to the one of *Limeños* after the daily input they are exposed to in the media.

This hypothesis has to be examined by a comparison of overt SPP rates in both varieties, and by comparing also the factor groups that favor the overt variant and the constraints within them.

First, as we have observed in section 7, the overt SPP rates of Huancayo and Lima are very similar (16.2% and 16.8%, respectively). This correspondence could suggest *prima facie* that *Huancaínos* are imitating *Limeños* even in a variable that is not as cognitively salient as the phonetic and phonological ones. Moreover, if we observe the rates in table 1, the two Peruvian communities are clearly the ones showing the smallest percentage of overt variants, forming a homogeneous group and challenging Orozco and Guy's (2008) claim that Peninsular varieties display the lowest rates of pronominal subjects.

However, a rate comparison is not enough to establish the conditioning of a variable (Otheguy and Zentella, 2012). For such reason, I performed a regression analysis of variable SPE in *Limeño* Spanish using the same data as in Á. Cerrón-Palomino (2014), but including two factors that were not analyzed in the aforementioned study: reflexivity of the verb and priming[2], for the sake of accuracy in comparing SPE in the two dialects.

In table 4, we can see the results of the analysis. The factors that favor overt SPPs in Lima are verbs inflected for 2nd and 1st person singular, switch reference, psychological and speech act verbs, ambiguous TAMs, and priming of a previous co-referential overt SPP. The constraints are thus not the same as in Huancayo: there are more overt SPP predictors in Lima. The verb class and ambiguous TAM factor groups were discarded in the regression analysis in Huancayo, whereas they both exert an effect on SPE in Lima.

With regard to the similarities between the two varieties, the most striking one is that in both varieties the highest ranked factor group is that of person and number, and within said factor group, 2nd person singular and 1st person singular are the strongest predictors of overt SPPs in both dialects. Some authors attribute the primacy of said persons/numbers in favoring overt SPPs to the dialogic nature of sociolinguistic interviews, which engages the interviewer and interviewee in a greater use of overt pronominal forms *yo* and *tú* (Enríquez, 1984: 191; Manjón-Cabeza et al., 2016: 191).

Another plausible explanation to the prevalence of 2nd and 1st person singular overt SPPs is Kuno and Kaburaki's (1977) empathy hierarchy (EH). The authors consider empathy to be "the speaker's identification with varying degrees (ranging from degree 0 to 1), with a person who participates in the event that he describes in a sentence". The rationale for the EH is the assumption that speakers are mostly interested in themselves, followed by their interlocutors, followed then by any other person or object in the context.

2 By the same token, for the sake of comparison of identical constraints, the factor of verb hierarchy, which turned out to be non-significant in Á. Cerrón-Palomino (2014) and was not examined in the Huancayo data, was excluded from the new Lima regression analysis.

TABLE 4

VARBRUL analysis of the linguistic factors contributing to the occurrence of overt pronominal subjects in the Spanish of Lima, Peru

Corrected mean			0.139
Log likelihood			-646.515
Significance			0.031
Total N			1584
	Factor Weight	**%**	**N**
Person and number			
tú + usted	.73	39.6	53
yo	.59	19.9	853
él + ella	.43	17.7	277
nosotros + nosotras	.43	10.1	267
ellos + ellas	.19	5.4	130
Range	***54***		
Referential connection			
switch reference	.64	25.1	802
same reference	.36	9.6	781
Range	***28***		
Verb class			
psychological	.61	27	178
speech act	.52	20.7	92
other	.50	16.7	989
copula	.48	17.1	211
movement	.33	7.9	114
Range	***28***		
TAM			
ambiguous	.66	23.2	551
not ambiguous	.42	14.4	1033
Range	***24***		
Previous realization (priming)			
pronominal	.64	31.5	216
not pronominal	.48	15.3	1368
Range	***16***		

In consequence, the prediction of the hierarchy, when applied to SPE, would be that 1st person singular would be the strongest overt SPP predictor, followed by 2nd person singular would favor overt SPPs, and if 3rd person singular were also a favoring constraint, it would be ranked third. The results presented in tables 3 and 4 show that 3rd person singular is actually the cut-off point for both varieties, since it does not favor pronominal subjects, along the lines of the EH. However, in the Spanish of Huancayo and Lima, it is the 2nd person singular that contributes more to overt pronominal subjects than the 1st person singular. In this respect DeLancey (1981) shows evidence of languages where phenomena like split intransitivity are controlled by a hierarchy where 2nd person outranks 1st person. In other words, a weaker version of the EH simply states that 1st and 2nd persons singular, which DeLancey (1981: 627) calls Speech-Act Participants (SAPs), outrank 3rd person singular, without there being predictions of which of the SAPs is ranked higher. This version of the EH accounts better for the factor group of person/number in SPE in the two varieties studied in this paper.

The second similarity between *Huancaínos* and *Limeños* is that the factor group ranked second for both varieties is referential connection, with the switch reference constraint being the overt SPP predictor, in congruence with the majority of studies on Spanish SPE.

The last similarity found between the two varieties is the significant effect that the factor group of previous realization of a co-referential subject has on SPE, despite the fact that it was ranked third in Huancayo and fourth in Lima. Priming of a previously uttered overt SPP is another stable predictor of an overt pronominal subject.

In the following lines, I will present the differences found in the two regression analyses. A possible reason why ambiguous TAM is significant in Lima but not in Huancayo is that *Limeños* do elide coda /s/ (Caravedo, 1990), which can cause ambiguity, whereas coda /s/ deletion is almost inexistent in Andean Spanish (Klee and Caravedo, 2006). A finding supporting this explanation is that the second person singular, which is the strongest SPP predictor for both groups and—coincidentally—the only verbal ending sensitive to morphemic *-s*, has a much higher frequency of overt SPPs in Lima (39.6%) than in Huancayo (23.8%). It is possible, then, that a potential four-way ambiguity system including verbs agreeing with SPP *tú* in Lima prompts the use of more pronouns in the speech of the community, whereas the three-way ambiguity existing in Huancayo is not extreme enough to promote the overt variant.

With regard to the second difference, the effect of the verb semantic class, the Lima variety patterns with most dialects in being favored by psychological verbs, although it is also favored by speech act verbs. This arrangement contrasts with Huancayo Spanish, where there is no verb class effect operating on the choice of grammatical subjects. In this respect, it is worth to underscore the fact that the semantic class of the verb is not a uniform SPP predictor across varieties. For instance, Michnowicz (2015) found that mental verbs exerted no influence on SPE in the Spanish monolinguals and Maya-Spanish bilinguals in Yucatan, Mex-

ico. By the same token, the Mexican children's production studied by Shin and Erker (2015) revealed no SPE conditioning by the semantic class of the verb. In addition, Orozco (2016) found copulative, perception and motion verbs—in that order—to favor overt SPPs in Jalapa, Mexico. Moreover, Orozco reports that within each verb semantic category there are verbs that behave differently, which leads him to propose, alongside other researchers (Erker and Guy, 2012; Posio, 2015; Silva-Corvalán, 2015), that there may be a lexical effect at work rather than a genuine semantic class effect.

In sum, Huancayo Spanish does not follow the same set of SPE constraints that the Spanish of Lima exhibits, supporting R. Cerrón-Palomino's (2003) hypothesis that only the most cognitively prominent Andean Spanish features are replaced. It can be concluded, therefore, that Huancayo Spanish is not adjusting to the acrolect regarding its SPE, contrary to what happens with phonetic and phonological variables.

7. Conclusions

This study was inspired by three research questions: a) how does the overt SPP rate of Andean Spanish compare to the rest of the Hispanic world?, b) what are the linguistic factors constraining SPE in Huancayo Andean Spanish?, and c) is Andean Spanish SPE modelled after the prestigious *Limeño* variety?

Regarding the first objective, Huancayo Spanish shows a rate of 16.2% of overt SPPs, the lowest one yet attested, patterning with the 16.8% reported for Lima (Á. Cerrón-Palomino, 2014). This extremely low rate of overt SPPs places Andean Huancayo Spanish amongst varieties that are conservative with regard to SPE (Cameron, 1993; Holmquist, 2012), which are for the most part Peninsular, Mexican and Peruvian. Moreover, it calls for a reformulation of the characterization of Peninsular varieties as being the ones with the lowest overt SPP rates (Orozco and Guy, 2008).

In order to answer my second research question, I conducted a separate regression analysis with Goldvarb X to examine the linguistic constraints conditioning SPE in Huancayo. The analysis revealed that the linguistic factors favoring overt SPP are verbs inflected for 2nd and 1st person singular, switch reference and priming, which have proven to be the most robust SPP predictors across varieties.

The third research question, asking whether *Huancaínos* follow SPE patterns of *Limeños* or not, demanded a logistic regression of the Spanish spoken in Lima, examining the same linguistic factor groups studied for Andean Spanish. Therefore, I re-analyzed the data scrutinized in Author (2014) in order to have comparable results. The overt SPP predictors for Lima were second and first person singular, switch reference, psychological and speech act verbs,

ambiguous TAMs, and priming of a previous co-referential overt SPP. As is evident, the verb class and ambiguous TAM factor groups do play a significant role in *Limeño* Spanish, while they do not affect SPE in Andean Spanish. The aforementioned results suggest that Huancayo Spanish does not follow the Peruvian acrolect in this syntactic variable, in contrast to cognitively salient phonetic and phonological features, in accordance with R. Cerrón-Palomino's prediction (2003), which states that stereotypical features of Andean Spanish, in particular, phonetic and phonological ones, are abandoned by speakers faster than those less salient cognitively, such as morphosyntactic structures.

To conclude, this study has contributed to a better understanding of subject pronoun expression in Spanish, by examining the linguistic factors constraining this variable in an understudied variety which, nonetheless, represents a macrodialect spoken across the vast Andes Mountains. This study also shows that Huancayo Spanish has the lowest rate of overt SPPs attested across dialects, displacing Peninsular varieties, and that this Andean Spanish variety drifts from the patterns of the *Limeño* acrolect, since it does not follow the same set of SPE constraints.

8. Works cited

Abreu, Laurel, 2012: "Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish" in Kimberly Geeslin and Manuel Díaz-Campos (eds.): *Selected proceedings of the 14th Hispanic Linguistics Symposium*, Somerville, MA: Cascadilla Proceedings Project, 1-8.

Alfaraz, Gabriela, 2015: "Variation of overt and null subject pronouns in the Spanish of Santo Domingo" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 3-16.

Barnes, Hilary, 2010: "Subject pronoun expression in bilinguals of two null subject languages" in Karlos Arregi, Zsuzsanna Fagyal, Silvina A. Montrul, and Annie Tremblay (eds.): *Romance Linguistics 2008: Interactions in Romance. Selected papers from the 38th Linguistic Symposium on Romance Languages (LSRL), Urbana-Champaign, April 2008*, Amsterdam: John Benjamins, 9-22.

Bayley, Robert, Norma Cárdenas, Belinda Treviño, and Carlos Vélez, 2012: "Spanish Dialect Contact in San Antonio, Texas: An Exploratory Study" in Kimberly Geeslin and Manuel Díaz-Campos (eds.): *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, Somerville, Massachusetts: Cascadilla Proceedings Project, 48-60.

Cameron, Richard, 1993: "Ambiguous agreement, functional compensation, and nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain", *Language Variation and Change* 5, 305-334.

Cameron, Richard, and Nidia Flores-Ferrán, 2004: "Perseveration of subject expression across regional dialects of Spanish", *Spanish in Context* 1, 41-65.

Caravedo, Rocío, 1990: *Sociolingüística del español de Lima*, Lima: Pontificia Universidad Católica del Perú.

Caravedo, Rocío, 1999: *Lingüística del corpus: cuestiones teórico-metodológicas aplicadas al español*, Salamanca: Ediciones Universidad de Salamanca.

Caravedo, Rocío, 2014: *Percepción y variación lingüística. Enfoque sociocognitivo*, Madrid/Frankfurt: Iberoamericana/Vervuert.

Caravedo, Rocío, and José Luis Rivarola, 2011: "El español andino, ¿variedad real o mental?" in Willem Adelaar, Pilar Valenzuela y Roberto Zariquiey (eds.): *Estudios sobre lenguas andinas y amazónicas. Homenaje a Rodolfo Cerrón-Palomino*, Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú, 369-389.

Carvalho, Ana M., and Michael Child, 2011: "Subject pronoun expression in a variety of Spanish in contact with Portuguese" in Jim Michnowicz and Robin Dodsworth (eds.): *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, Somersville, MA: Cascadilla Proceedings Project, 14-25.

Carvalho, Ana M., and Ryan M. Bessett, 2015: "Subject pronoun expression in Spanish in contact with Portuguese" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 143-168.

Carvalho, Ana M., Rafael Orozco, and Naomi Shin (eds.), 2015: *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*, Washington: Georgetown University Press.

Cerrón-Palomino, Álvaro, 2014: "Ser o no ser pro: variación en la expresión de sujeto pronominal en el castellano limeño", *Lingüística* 30 (1), 61-83.

Cerrón-Palomino, Álvaro, 2016: "(No) English interference on U.S. Southwest Spanish? A look at variable subject expression in Phoenix Spanish-English bilinguals", *Sociolinguistic Studies* 10 (3), 383-408.

Cerrón-Palomino, Álvaro, 2018: "Null Subject Encounter: Variable Subject Pronoun Expression in the Spanish of Quechua-Spanish Bilinguals in the Central Peruvian Andes", *International Journal of Bilingualism* [online first: https://journals.sagepub.com/doi/10.1177/1367006918763175].

Cerrón-Palomino, Rodolfo, 2003: *Castellano andino. Aspectos sociolingüísticos, pedagógicos y gramaticales*, Lima: Fondo Editorial Pontificia Universidad Católica del Perú, Cooperación Técnica Alemana GTZ.

De Prada Pérez, Ana, 2015: "First person singular subject pronoun expression in Spanish in contact with Catalan" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 121-142.

DeLancey, Scott, 1981: "An interpretation of Split Ergativity and Related Patterns", *Language* 57 (3), 626-657.

Enríquez, Emilia, 1984: *El pronombre personal sujeto en la lengua española hablada en Madrid*, Madrid: CSIC.

Erker, Daniel, and Gregory Guy, 2012: "The role of lexical frequency in syntactic variability: variable subject personal pronoun expression in Spanish", *Language* 88 (3), 526-557.

Escobar, Anna María, 2016: "Dialectos del español de América: español andino" in Javier Gutiérrez-Rexach (coord.): *Enciclopedia de Lingüística Hispánica* 2, New York: Routledge, 353-362.

Flores-Ferrán, Nidia, 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 (1), 49-73.

Godenzzi, Juan Carlos: 1986. "Pronombres de objeto directo e indirecto del castellano en Puno", *Lexis* 10 (2), 187-201.

Godenzzi, Juan Carlos: 1991. "Discordancias gramaticales del castellano andino en Puno (Perú)", *Lexis* 15 (1), 107-118.

Holmquist, Jonathan, 2012: "Frequency rates and constraints on subject personal pronoun expression: Findings from the Puerto Rican highlands", *Language Variation and Change* 24 (1), 203-220.

Klee, Carol, and Rocío Caravedo, 2006: "Andean Spanish and the Spanish of Lima: Linguistic variation and change in a contact situation" in Clare Mar-Molinero and Miranda Stewart (eds.): *Globalization and Language in the Spanish-speaking World. Macro and Micro Perspectives*, Hampshire and New York: Palgrave MacMillan, 94-113.

Kuno, Susumu, and Etsuko Kaburaki, 1977: "Empathy and Syntax", *Linguistic Inquiry* 8 (4), 627-672.

Lapidus Shin, Naomi, and Ricardo Otheguy, 2005: "Overt nonspecific *ellos* in Spanish in New York", *Spanish in Context* 2 (2), 157-174.

Lastra, Yolanda, and Pedro Martín Butragueño, 2015: "Subject pronoun expression in oral Mexican Spanish" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 39-58.

Manjón-Cabeza, Antonio, Francisca Pose, and Francisco Sánchez, 2016: "Factores determinantes en la expresión de sujeto pronominal en el corpus PRESEEA de Granada", *Boletín de Filología* 51 (2), 181-207.

Merma, Gladys, 2004: "Lenguas en contacto: Peculiaridades del español andino peruano. Tres casos de interferencia morfosintáctica", *Estudios de Lingüística* 18, 191-211.

Michnowicz, Jim, 2015: "Subject pronoun expression in Yucatan Spanish" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 101-120.

Orozco, Rafael, and Gregory Guy, 2008: "El uso variable de los pronombres de sujetos: ¿qué pasa en la costa Caribe colombiana?" in Maurice Westmoreland and Juan A. Thomas (eds.): *Selected proceedings of the 4th workshop on Spanish sociolinguistics*, Somerville: Cascadilla Proceedings Project, 70-80.

Orozco, Rafael, 2015: "Pronominal Variation in Colombian Costeño Spanish" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject pronoun expression in Spanish: A cross-dialectal perspective*, Washington: Georgetown University Press, 17-37.

Orozco, Rafael, 2016: "Subject pronoun expression in Mexican Spanish: ¿Qué pasa en Xalapa?" in *Proceedings of the Linguistic Society of America* 1, 7, 1-15.

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

Otheguy, Ricardo, and Ana Celia Zentella, 2012: *Spanish in New York. Language contact, dialectal leveling, and structural continuity*, New York: Oxford University Press.

Posio, Pekka, 2015: "Subject Pronoun Usage in Formulaic Sequences: Evidence from Peninsular Spanish" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 59-79.

Ranson, Diana, 1991: "Person marking in the wake of /s/ deletion in Andalusian Spanish", *Language Variation and Change* 3, 133-52.

Sankoff, David, Sali Tagliamonte, and Eric Smith, 2005: *Goldvarb X*. Department of Linguistics, University of Toronto, Toronto, Canada.

Shin, Naomi, and Daniel Erker, 2015: "The Emergence of Structured Variability in Morphosyntax: Childhood Acquisition of Spanish Subject Pronouns" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 169-189.

Silva-Corvalán, Carmen, 1994: *Language Contact and Change. Spanish in Los Angeles*, New York: Oxford University Press.

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

Silva-Corvalán, Carmen, 2015: "The Acquisition of Grammatical subjects by Spanish-English bilinguals" in Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin (eds.): *Subject Pronoun Expression in Spanish. A Cross-Dialectal Perspective*, Washington: Georgetown University Press, 211-229.

Tagliamonte, Sali, 2006: *Analysing Sociolinguistic Variation*, New York: Cambridge University Press.

Torres Cacoullos, Rena, and Catherine Travis, 2010: "Variable *yo* expression in New Mexico: English influence?" in Susana Rivera-Mills and Daniel Villa (eds.): *Spanish of the U.S. Southwest: A Language in Transition*, Madrid: Vervuert/Iberoamericana, 185-206.

Travis, Catherine, 2005: "The yo-yo effect" in Randall Gess and Edward J. Rubin (eds.): *Theoretical and Experimental Approaches to Romance Linguistics: Selected papers from the 34th Linguistic Symposium on Romance Languages (LSRL)*, Amsterdam: John Benjamins, 329-349.

Travis, Catherine, 2007: "Genre effects on subject expression in Spanish: priming in narrative and conversation", *Language Variation and Change* 19 (2), 101-136.