Andrea Pešková

Experimenting with *Pro-drop* in Spanish[1]

Abstract

This paper investigates the omission and expression of pronominal subjects (PS) in Buenos Aires Spanish based on data from a production experiment. The use of PS is a linguistic phenomenon demonstrating that the grammar of a language needs to be considered independently of its usage. Despite the fact that the observed Spanish variety is a consistent pro-drop language with rich verbal agreement, the data from the present study provide evidence for a quite frequent use of overt PS, even in non-focal, non-contrastive and non-ambiguous contexts. This result thus supports previous corpus-based empirical research and contradicts the traditional explanation given by grammarians that overtly realized PS in Spanish are used to avoid possible ambiguities or to mark contrast and emphasis. Moreover, the elicited semi-spontaneous data indicate that the expression of PS is optional; however, this optionality is associated with different linguistic factors. The statistical analysis of the data shows the following ranking of the effects of these factors: grammatical persons > verb semantics > (syntactic) clause type > (semantic) sentence type.

1. Introduction

It is well known that Spanish is a pro-drop ("pronoun-dropping") or null-subject language whose grammar permits the omission of pronominal

subjects (PS). The widespread term "pro-drop" emerged from the Principles and Parameters model of language within the Generative framework (Chomsky 1981). The newest Generative typology of the pro-drop parameter (see Biberauer, Holmberg, Roberts & Sheenan 2010: 6–13) suggests four identifiable types of null-subject language that facilitate the omission of subjects: Expletive null-subject languages (e.g. German), partial null-subject languages (e.g. Finnish), discourse (radical) pro-drop languages (e.g. Chinese), and consistent null-subject languages (e.g. Spanish). The latter group characteristically shows "rich" verbal agreement inflection (see Section 2).

The use of PS in Spanish constitutes not only a source of contexts with empty subject pronouns, but it also offers many examples of contexts in which the subject is but does not have to be expressed. The subject position must always be empty (Ø) in impersonal and generic structures such as (1a–b), as well as in sentences with inanimate reference (1c):

(1)   a. *Ø Está   lloviendo.*
         Ø be.3SG.PRES.IND   rain.GERUND
         'It is raining.'

      b. *Ø Llaman   a la puerta.*
         Ø call.3PL.PRES.IND   at the door
         'There is someone at the door.'

      c. *Ø Está   sobre la mesa (el libro).*
         Ø be.3SG.PRES.IND on the table   (the book)
         'It is on the table (the book).'

Personal sentences in Spanish, however, show null-overt subject pronoun variation, i.e. a subject pronoun can be overtly realized or not (see Example 2):

(2)   a. *Ø Estoy   cantando.*
         Ø be.1SG.PRES.IND sing.GERUND
         'I am singing.'

      b. *Yo   hablo   español.*
         I.NOM   speak.1SG.PRES.IND Spanish
         'I speak Spanish.'

Moreover, there are instances in Spanish in which null-overt subject pronoun variation does not occur, i.e. in which an overt pronoun or a null subject is always required. The first case arises when the pronominal subject is interpreted as a focus (3a) or as a contrastive topic (3b); the second case occurs, for instance, in coordination clauses (2ⁿᵈ position) (3c) or in imperatives (3d):

(3)   a. *Yo   estoy   cantando   y   no Juan.*
          I.NOM be.1SG.PRES.IND sing.GERUND   and not John
          'I am singing and not John.'

       b. *Juan habla   checo, pero yo   hablo   eslovaco.*
          John speak.3SG.PRES.IND Czech but   I.NOM speak.1SG.PRES.IND Slovak
          'John speaks Czech, but I speak Slovak.'

       c. *Pedro canta   y   Ø toca   la guitarra.*
          Peter   sing.3SG.PRES.IND and Ø play.3SG.PRES.IND the guitar
          'Peter sings and plays the guitar.'

       d. *¡ØAbre   la puerta!*
          Ø open.2SG.IMP the door
          'Open the door!'

Extensive research on the use of PS in Spanish combines different perspectives that are usually treated separately in the literature. Traditional Hispanic grammar throughout the past century (e.g. RAE 1973; Badia Margarit 1988; Alarcos Llorach 1994) asserted that the lack of PS is admissible in Spanish due to verbal affixes, which function as proper subjects in clauses (compare with the typological account of *World Atlas of Language Structures* in Dryer 2005). In contrast to this perspective, the Generative framework introduced the empty category *pro*, which occupies the subject position in finite sentences (e.g. *pro hablo* 'I speak') (see e.g. Chomsky 1981, 1995; Rizzi 1982, 1986 for general pro-drop properties; for Spanish see e.g. Bosque 1989; Fernández Soriano 1999; Luján 1999). In addition, this approach investigates the different morphosyntactic and licensing conditions for *pro* and its overt counterpart (see e.g. Montalbetti 1984), it describes cross-linguistic variation and explains language change (see e.g. Fischer 2010) as well as language acquisition (see e.g. Isabelli 2004). The traditional Generative (e.g. Rizzi 1986) as well as Hispanic (e.g. Alarcos Llorach 1994) view on the omission of pronominal subjects in

languages like Spanish is that it is conditioned by "rich" verbal agreement (e.g. *canto* 'I sing'; *cantamos* 'we sing', etc.). In contrast to the omission of PS, the expression of PS has been explained using reasons such as ambiguity resolution, contrast, and emphasis. While the first reason addresses the syncretism in the inflectional marking of subject person in certain tenses or moods (e.g. *hablaría* 'I/(s)he would speak'; *hablaba*, 'I/(s)he talked'), the latter deal with their pragmatic functions in discourse. It is important to note that the PS must always be phonetically realized when interpreted as a focus or a contrastive topic (cf. 3a and 3b), whereas their expression is not obligatory in cases of ambiguity (grammatical features), as the inflectional syncretism can be disambiguated by the context (see e.g. Silva-Corvalán 2001).

In comparison to the grammarians' – mostly descriptive non-empirical – point of view, numerous corpus-based studies have examined the variable use of subject pronouns in Spanish (e.g. Bentivoglio 1987; Bayley & Pease-Álvarez 1997; Silva-Corvalán 2001; Lipski 2002; Amaral & Schwenter 2005; Orozco & Guy 2008; Posio 2008, 2011; Aijón Oliva & Serrano 2010, among many others). These empirical analyses have indicated that the variation between null and overt subject pronouns is primarily motivated by internal factors – structural features of a language or dialect – such as grammatical person, morphological and contextual ambiguity, verb semantics, clause type or switch-reference. External – social – factors, which are usually addressed in the variationist studies of the Labovian (sociolinguistic) tradition, do not seem to play a decisive role in the observed phenomenon (see e.g. Bentivoglio 1987; Silva-Corvalán 2001). Regarding the use of PS, the finding of a cross-dialectal variation can be considered to be one of empirical investigations' most interesting contributions (see Table 1):

Table 1. Overt subjects as percentages of the total in different locations (from Otheguy, Zentella & Livert 2007)

Rate of overtly realized PS	Location
19%	Mexico
24%	Colombia
27%	Ecuador
33%	Cuba
35%	Puerto Rico
41%	Dominican Republic

As can be observed, the Caribbean dialects (e.g. Puerto Rico, Dominican Republic) exhibit the highest rates of overt pronoun usage among the Spanish dialects. This suggests that these varieties realize overt PS very frequently, even in pragmatically neutral contexts in which the close-to-standard varieties (e.g. Spain, Mexico) prefer empty pronouns (cf. RAE 2010: §33.4c). Some Caribbean varieties also allow for the realization of pronominal subjects in impersonal structures or sentences with inanimate reference (cf. 1a–c). For instance, it is possible to say *Ello está lloviendo* ('It is raining') or *Él* (*el libro*) *está sobre la mesa* ('He (the book) is on the table') in Dominican Spanish (see Henríquez Ureña 1939; Toribio 2000; Hinzelin & Kaiser 2006 for these dialects).

Most empirical studies are based on spoken (e.g. Barrenechea & Alonso 1977) or written (e.g. Lu 1997) language corpora. This paper offers a novel viewpoint pertaining to empirical research on the use of PS in Spanish. It investigates the realization of subject pronouns in semi-spontaneous speech obtained through an elicited production task conducted with 13 native speakers of *Porteño*, the Spanish variety spoken in Buenos Aires. Its aims are twofold: (1) to determine whether a correlation exists between the use of pronominal subjects and selected (intra-)linguistic factors, and (2) to demonstrate that the use of a production task may lead to a better understanding of the usage of grammar in a natural, subconscious way. The study has two hypotheses: First, it is predicted that the overt realization of PS is possible, even in spite of non-focal, non-ambiguous and non-contrastive contexts (**hypothesis 1**). This assumption contradicts the standard explanation given by grammarians and confirms the findings of previous corpus-based empirical research. From this point of view, it thus supports Newmeyer's statement (2003: 25) that "Grammar is Grammar and Usage is Usage", i.e. that the grammar of a language must be characterized independently of its usage. The idea underlying this assumption is that "grammar contributes to an explanation of language use, but usage, frequency, and so on are not represented in the grammar itself" (Newmeyer 2003: 6). But what does this mean? In terms of the phenomenon under investigation, the grammar of Spanish (in a narrow sense syntax and/or morphosyntax) assumes that the pronoun subject must be omitted. Native Spanish speakers are usually conscious of this property when comparing it to a second language such as English (non-pro-drop language), in that the PS must be almost always realized (e.g. '**He** told me that **he** was in Argentina'). However, the use of PS in Spanish, especially in spoken language, seems to be subconscious, and is varied with regards to different,

mainly semantic, discourse and pragmatic, factors. As mentioned above, most previous corpus-based empirical studies have shown that it is necessary to consider various intervening factors concerning the use of PS in Spanish. My investigation concentrates on four linguistic variables: (1) grammatical person (*yo* 'I', *vos* 'you-SG (familiar)', *él/ella* 'he/she', *usted* 'you-SG (formal)', *nosotros* 'we', *ustedes* 'you-PL', *él/ella* 'they'), (2) verb semantics (epistemic verbs vs. perceptive verbs), (3) type of sentence (declarative, absolute interrogative, *wh*-interrogative), and (4) type of clause according to its structural complexity (matrix clause with or without subordinate clause, subordinate clause) (see e.g. Barrenechea & Alonso 1977; Lu 1997; Otheguy; Zentella & Livert 2007 for an investigation of similar factors). One of the main questions of my study is whether a statistical correlation exists between the overt pronoun rate and the selected factor groups. Special focus will be placed on determining whether grammatical persons (factor 1) demonstrate the same overt pronoun rate under the same conditions (factors 2–4). The grammatical person as a relevant factor in the use of PS has been supported by many empirical studies. For example, a study by Barrenechea and Alonso (1977) on the usage of PS in the spoken language of *Porteño* found a higher overt pronoun rate with singular persons than in the persons of the plural as well as a greater probability of overt pronouns with 1ˢᵗ and 2ⁿᵈ persons than with 3ʳᵈ persons. In agreement with Barrenechea and Alonso (1977) and other earlier corpus-based empirical research (e.g. Lu 1997), it is expected that the overt pronoun rate is different for every grammatical person, often despite the same or very similar contexts (**hypothesis 2**). The factor "grammatical person" and the three remaining factors will be presented in more detail in Section 2.

The present paper is organized as follows: Section 2 describes the pronominal and the verbal properties of *Porteño* Spanish, outlines existing findings on the use of PS in this dialect and presents the point of departure for this study. Section 3 then describes the methodology and data, while Section 4 offers the results, followed by a discussion in Section 5. Finally, the paper ends with concluding remarks in Section 6.

2.   Expression of pronominal subjects in *Porteño* Spanish

In this section I will describe the results of several existing (corpus-based) empirical studies on the use of PS in *Porteño* Spanish. The overt-null subject pronoun variation in this dialect is only possible in personal

constructions, i.e. in sentences with referential (animate) subjects. Example (4) presents the pronominal and the morphologically "rich" verbal system of *Porteño* Spanish in the present tense. The possible omission of the subject pronouns is indicated by parentheses:

(4) a. (*yo*) *canto*
      I.NOM sing.1SG.PRES.IND
      'I sing.'

    b. (*vos*) *cantás*
      you.NOM sing.2SG.PRES.IND
      'You (informal) sing.'

    c. (*él* / *ella* / *usted*) *canta*
      he.NOM / she.NOM / you.NOM sing.3SG.PRES.IND
      'He/she sings. / You (formal) sing.'

    d. (*nosotros / nosotras*) *cantamos*
      we.NOM.M/F sing.1PL.PRES.IND
      'We sing.'

    e. (*ellos* / *ellas* / *ustedes*) *cantan*
      they.NOM.M / they.NOM.F / you.NOM sing.3PL.PRES.IND
      'They sing. / You (informal, PL) sing.'

Notice that for etymological reasons, the second persons *usted* and *ustedes* are conjugated in the third person in Spanish. These pronouns are derived from the honorific address *vuestra merced*, 'your mercy', which was used until the middle of the 19th century. The high overt pronoun rate of *usted* found in different empirical analyses (see e.g. Enríquez 1984, 76%) is usually attributed to the historical origin of this pronoun (RAE 2010: §16.1.b, §33.5h, §16.14g) and/or its formality (DPD 2005: 531). Whereas the singular form *usted* is exclusively formal, the plural form *ustedes* is used for both familiar and formal speech in *Porteño*. In Peninsular Spanish, *ustedes* is the plural formal address, a counterpart to the informal 2nd person plural *vosotros*, which is absent in almost all of Latin America (including Argentina). Another feature of Argentinean Spanish is the so-called *voseo*: The usage of the 2nd person singular pronoun *vos* instead of *tú* (Standard Spanish) and the corresponding *voseo*-verb conjugation (e.g. *cantás* instead of *cantas* '(you) sing'; *sos* instead of *eres* '(you) are', etc.).

Regarding the pro-drop characteristics of *Porteño* Spanish, this variety reflects the grammar of standard Spanish and thus exhibits properties typical of null-subject languages such as postverbal subjects, *that*-trace effect, null-expletives, etc. (see e.g. Rizzi 1986; Biberauer et al. 2010). Nevertheless, there are some differences between *Porteño* and other Spanish varieties with respect to the usage of pronominal subjects in spoken language. Comparing the overt pronoun rates from previous empirical findings (see Barrenechea & Alonso 1977 for *Porteño* Spanish; Hochberg 1986 for Puerto Rican Spanish; Soares da Silva 2006 for *Porteño* and Peninsular Spanish), the overt pronoun rate of *Porteño* seems to lie between those of the Caribbean and close-to-standard dialects. For instance, a study by Pešková (2011) on the use of the 2nd person singular showed that *Porteño* tends to realize subject pronouns in certain sentences in which Peninsular Spanish (specifically the Madrid dialect) prefers their omission. Examples are given in (5) and (6) (taken from Pešková 2011: 57):

(5)   a. *Porteño* Spanish:
       Vos   sabés   que a mí   me   gustan...
       you.NOM.SG know.2SG.PRES.IND that to me.DAT me.DAT like.3PL.PRES.IND
       'You know that I like…'

       b. Peninsular Spanish:
       Sabes   que a mí   me   gustan...
       know.2SG.PRES.IND that to me.DAT me.DAT like.3PL.PRES.IND
       'You know that I like…'

(6)   a. *Porteño* Spanish:
       ¿Vos   me   querés   decir   flaca?
       you.NOM.SG   me.DAT want.2SG.PRES.IND tell.INF thin
       'Are you wanting to tell me (that I am) thin?'

       b. Peninsular Spanish:
       ¿Me   quieres   decir   flaca?
       me.DAT want.2SG.PRES.IND tell.INF thin
       'Are you wanting to tell me (that I am) thin?'

These differences between the two varieties were detected on the basis of a comparative analysis of the Argentinean comic book *Maitena* and its Peninsular Spanish counterpart. Although the findings do not tell us much about the systematic use of overt subject pronouns, the examples in (5)–(6)

provide clear evidence of a possible realization of the subject pronouns in *Porteño*, even in pragmatically non-marked, i.e. non-contrastive and non-focused, contexts. The question is how this use of the overt PS should be interpreted. The presence of the pronoun *vos* in *Porteño* probably has some special pragmatic function; it expresses the speaker's attitude and indicates a connection between him and the hearer or other participants in a conversation. The pronoun may be a part of a kind of emphatic or emotional expression. Speakers of the peninsular dialect seem to prefer the omission of PS, probably only using intonation or other element such as *sólo* ('only'). The latter case is demonstrated in (7) (from Pešková 2011: 57):

(7)   a. *Porteño* Spanish:
       ¿Todavía ahí?   ¡Vos   vivís   para trabajar!
       still   there you.NOM.SG live.2SG.PRES.IND for   work.INF
       'Still there? You live for work!'

      b. Peninsular Spanish:
       ¿Todavía ahí?   ¡Sólo vives   para trabajar!
       still   there only live.2SG.PRES.IND for   work.INF
       'Still there? You live (only) for work!'

In addition, an empirical comparative study by Soares da Silva (2006) found dissimilarities between the Spanish spoken in Buenos Aires and in Madrid with regards to the overt pronoun rate. His findings are summarized in Table 2:

Table 2. Overt pronoun rates in Peninsular and *Porteño* Spanish in a study by Soares da Silva (2006)

Grammatical Persons		Peninsular	*Porteño*
First persons	*yo*	35%	37%
	nosotros	11%	38%
Second persons	*tú*/*vos*	22%	22%
	usted	31%	40%
	ustedes	33%	37%
Third persons	*él/ella*	12%	19%
	ellos/ellas	9%	23%

Comparing these dialects, Soares da Silva's results show a slightly higher rate of overt pronouns in *Porteño* Spanish for all grammatical persons.

Notice that the formal second person *usted* does not exhibit higher overt pronoun rates as is usually indicated in other studies (see e.g. Rosengren 1974: 56%; Enríquez 1984: 76%; Lu 1997: 90,8%). But what is even more noticeable is the difference between the overt pronoun rates of the 1st person plural (38% in *Porteño* vs. 11% in Peninsular). Unfortunately, no interpretation of this seemingly significant difference is to be found. Even though the rather high rate of the use of *nosotros* in *Porteño* or its very low rate in Peninsular Spanish remain quite suspicious, the results imply cross-dialectal variation within one pro-drop language.[2] It is problematic to say, however, to what extent we can rely on the differences in overt pronoun rate between dialects, as there might be also a disparity within each individual dialect. For example, comparing two empirical studies on the use of PS in *Porteño* (Barrenechea & Alonso 1977; Soares da Silva 2006), we can observe inconsistencies in the rates of overt PS:

Table 3. Overt pronoun rates according to grammatical persons in *Porteño* Spanish in the studies by Soares da Silva (2006) and Barrenechea & Alonso (1977)

Grammatical Persons		Soares da Silva (2006)	Barrenechea & Alonso (1977)
First persons	*yo*	37%	24%
	nosotros	38%	20%
Second persons	*vos*	22%	36%
	usted	40%	56%
	ustedes	37%	59%
Third persons	*él/ella*	19%	11%
	ellos/ellas	23%	17%

In Barrenechea and Alonso (1977), the overt pronoun rates are lower with regards to the first and third persons, whereas the second persons exhibit higher percentages for the realization of subject pronouns. Nevertheless, both studies show the following tendency: Formal second persons and first persons (at least in Soares da Silva 2006) show higher overt pronoun rates, while the third persons have lower rates of overt pronoun use. This might

---

[2] Interestingly enough, cross-linguistic variation has also been attested with respect to the use of PS in pro-drop languages. For example, a comparative empirical study by Posio (2012) detected several systematic differences between Peninsular Spanish and the typologically related European Portuguese (both Romance consistent null-subject languages). See also Biberauer et al. (2010) for differences between Spanish and Italian in the Minimalist framework.

have something to do with the function of the subject pronouns in the speech act; second persons are "hearers", first persons are "speakers", and third persons are commonly "non-speakers, non-hearers". The low frequency of the pronouns of 3rd persons is usually attributed to the fact that these are anaphoric, i.e. interpreted as given information or familiar topics in some explanations (see e.g. Frascarelli 2007), and are frequently used in narrative. Additionally, 1st and 2nd persons are deictic (see Cabredo-Hofherr 2006 for a typology of *pro*) and generally appear in interactive speech that involves a greater probability of overt subject pronouns (see Balasch 2008). Furthermore, it should be pointed out that both studies (Barrenechea & Alonso 1977 and Soares da Silva 2006) base their results on the same corpus of spontaneous speech produced by educated speakers from Buenos Aires. The corpus contains free interviews collected in the 1970's and edited by Barrenechea (1987) under the title *El habla culta de la ciudad de Buenos Aires*. The question arises as to why there is such a discrepancy in the realization of overt pronouns between the two studies. One potential explanation might be a methodological problem: The investigators used different corpus sizes for their analyses, chose different speakers and/or used different methods to analyze the data. For instance, maybe they did not exclude all contexts in which the pronominal subject must be obligatorily empty or overt; this means that those contexts showing no null-overt subject pronoun variation. While spontaneous data certainly offer a very important resource for the investigation of linguistic variation, problems may emerge when we attempt to extract information from such data. The corpus-data may be limited to the occurrences of certain linguistic phenomena or they do not establish a comparable and controllable set of data. As for the use of PS, the spoken data do not generate the same conditions for the factors which might influence the use of PS. For instance, the 1st person singular is found in this type of data very often with epistemic verbs: This increases the probability of an overt subject (e.g. *yo creo*, 'I think'). However, one is less likely to find the use of epistemic verbs in the third person (e.g. *ella cree*, 'she thinks'). It will be shown that the production experiment used in the present investigation creates similar contexts for all grammatical persons. Of course, the "data control" is also obtainable through spontaneous speech, in which the independent contribution of the linguistic factor groups is obtained using multivariate methods within different software packages (Varbrul, SPSS, etc.) in many corpus-based empirical studies (e.g. Otheguy, Zentella & Livert 2007). Nevertheless, I agree with Goodall (2010), who points out

that different experimental techniques "can give us more certainty about the status of data where there have been disputes or doubts, as well as more precision in dealing with subtle contrasts among sentences" (Goodall 2010: 233–234).

A manifest discrepancy in the empirical results discussed above served as the point of departure for the present experimental study. As already observed in Tables 2 and 3, empirical data show variation in overt pronoun rates between different grammatical persons (**factor 1**). Regarding the verb semantics (**factor 2**), a small production experiment in the context of Pešková (2011) indicates that the 2nd person singular *vos* is overtly very often realized with epistemic verbs, but very seldom with perceptive verbs. The considerable influence of epistemic verbs on subject pronoun realization has been previously attested in various empirical studies (cf. Enriquéz 1984; Bentivoglio 1987; Lu 1997; Hurtado 2001; Otheguy, Zentella & Livert 2007; Posio 2012). With respect to the type of sentence (**factor 3**), Pešková's study (2011) on the use of PS of the 2nd person singular indicates higher rates in overt pronoun use with declaratives and absolute interrogatives than with *wh*-questions. Finally, the variable "type of clause" (**factor 4**) exhibits a higher overt pronoun rate in simple (matrix) clauses than in complex clauses (Pešková 2011). However, there are some empirical studies stating that these two factors (type of sentence, type of clause) do not in fact play any role in the usage of PS in Spanish (e.g. Lu 1997: 126–127). In order to test statistical significance of the findings, I will use simple chi-square statistics and extend the analysis through the application of multiple regression models with random effects. In addition to presenting the overt pronoun rates, the syntactic and discourse properties of the overt PS will be briefly discussed on the basis of the collected experimental data.

In the following section I turn to the presentation of the methodology and data used in this study in order to contribute my own empirical view on this matter.

3.   Methodology and data

This section describes the spoken language corpus which was used as a source of data for the present study. The data stem from a **production experiment** conducted with 13 monolingual *Porteño* speakers in 2009 in

Buenos Aires.[3] The participants were six men and seven women, aged 20–45, all of whom were born and brought up in Buenos Aires. They had all a university degree or were university students at the time of interview, and were totally naïve as to the purpose of the experiment. In order to determine the relevance of the four selected linguistic factors (grammatical person, verb semantic, type of sentence and type of clause), a questionnaire with everyday non-contrastive and non-ambiguous situations was created to acquire target-sentences in which the speaker had the option to either say a subject pronoun or omit it. As an example, a task like *Preguntale a tu padre qué Ø opina de Buenos Aires* ('Ask your father what (he) thinks about Buenos Aires') was intended to lead to a target-question from the participant such as *¿Qué opinás (vos) de Buenos Aires?* ('What do [you] think of Buenos Aires?'). All contexts were presented by the researcher with empty subjects (Ø) and were repeated no more than three times. Despite the given context, the speakers were asked to express the target-sentences as naturally as possible. It should also be emphasized that the target-utterances considered for the final analysis had no generic reading or inanimate references. The experiment took no more than 45–60 minutes and the data were recorded for later transcription.

The questionnaire was designed as follows: 18 very similar situations were used for each of the seven subject pronouns (*yo, vos, él/ella, usted, nosotros, ellos/ellas, ustedes*) (**factor 1**). The pronoun *ustedes*, however, was only used as an informal (familiar) address in the experiment. Thus, the questionnaire consisted of a total of 126 situations. An example of the same situation for production with second persons is given in (8) (the target-verb which should be produced with or without the PS is underlined):

(8) Questionnaire Example:
    a. Situation Nr. 2 (*vos*): 'you', SG informal
       Estás hablando con tu padre. Preguntale qué <u>opina</u> de Buenos Aires.
       '[You] are talking with your father. Ask him what [he] thinks of Buenos Aires.'

---

[3] The methodology was inspired by the production task used in Pešková (2011) and in the intonation survey proposed by Prieto (2001). As for other experiments with pronouns, different grammaticality judgments tests (see e.g. Pešková 2011; Alonso-Ovalle, Fernández-Solera, Frazier et al. 2002 for Spanish; Carminati 2002 for Italian) or visual-world eye-tracking experiments were also carried out (see e.g. Kaiser & Trueswell 2008; Kaiser 2011 for English and Finnish).

b. Situation Nr. 20 (*usted*): 'you', PL formal
   Preguntale al Sr. Brandoni, profesor de la UBA, qué opina de la Argentina.
   'Ask Mr. Brandoni, Professor at UBA, what [he] thinks of Argentina.'

c. Situation Nr. 38 (*ustedes*) 'you', PL informal
   Estás hablando con tus padres. Preguntales qué opinan de Puerto Madero.
   '[You] are talking with your parents. Ask them what [they] think of Puerto Madero.'

Furthermore, each of the 18-situations sets was controlled according to the verb semantics (**factor 2**): Consequently, there were nine situations with epistemic verbs (such as *creer* 'think', *saber* 'know') and nine similar situations with perceptive verbs (such as *escuchar* 'listen to', *mirar* 'to watch'). Additionally, these two verb groups were also controlled for the type of sentence (declaratives, absolute questions and *wh*-questions) (**factor 3**) as well as for the type of clause according to its structural complexity (simple matrix clause, matrix clause with subordinate clause, subordinate clause) (**factor 4**).[4] As for factor 3, the 1st persons (*yo* and *nosotros*) were controlled only by declarative sentences and not by interrogatives. For a better understanding of the method used in the present study, see Appendix I with an example of the 18-situations set used for elicitation of the 2nd person singular *vos*. Observe that every situation led to a target-sentence which was controlled by all four factors (indicated in brackets; the verb-targets are underlined).

Finally, a **statistical analysis** was performed using STATA, version 11 (STATA Corp., Texas, USA). First, I examined whether the relationship between the use of pronominal subjects (dependent variable) and the selected factors (independent variable) was statistically significant. Then I tested the interaction that describes the simultaneous influence of two independent variables (grammatical person plus another of the three factors) on the dependent variable (use of PS). Moreover, the intraclass correlation was incorporated into all regression models with speakers (participants in the experiment) as well as into all models with interaction. The intraclass correlation should explain the extent to which units in the same group resemble one another. Lastly, the strength of the effects of the

---

[4] Clarifying examples are provided here for a better understanding of the factor "type of clause", where the target-verb appears (underlined): a) simple (matrix) clause (e.g. *Escucho música argentina* 'I listen to Argentinean music'), b) matrix clause with subordinate clause (e.g. *Creo que va a llover* 'I think that it is going to rain'), c) subordinate clause (e.g. *Me gusta lo que escuchás* 'I like what you listen to').

examined factors was observed in order to determine a possible improvement to the intercept model, i.e. a simple model without intervening factors which calculates or predicts the probability of the (non-) realization of PS.

4. Results

This section presents and summarizes the results obtained from the analysis of the empirical data. A total of 1638 tokens were recorded (126 situations per 13 speakers), 12 of which had to be discarded due to speaker failure ($N = 1626$). The results show that the speakers omitted subject pronouns in 52% of all contexts (845 cases). This means that the subject pronouns were overtly realized in 48% of all contexts (781 cases).

For **factor 1**, singular persons were overtly realized more often than plural persons (51% vs. 44%), a finding which supports those common in empirical studies (see e.g. Rosengren 1974; Hochberg 1986). As expected, the grammatical persons differed in overt pronoun rates, in spite of having been controlled for by the same or very similar contextual conditions (factors 2–4), see Table 4:

Table 4. Overt pronoun rate according to grammatical person

Person	%	N
yo (1SG)	47%	109/234
nosotros (1PL)	36%	85/233
vos (2SG, fam.)	33%	76/232
usted (2SG, form.)	70%	164/234
ustedes (2PL, fam.)	47%	110/232
él/ella (3SG)	55%	128/232
ellos/ellas (3PL)	48%	109/229
Total	**48%**	**781/1626**

While the overt pronoun rate was quite similar for first persons (*yo* 47% vs. *nosotros* 36%) and third persons (*él/ella* 55% vs. *ellos/ellas* 48%), there were remarkable differences with regards to second persons (*vos* 33%; *usted* 70%; *ustedes* 47%). The relationship between the grammatical persons and the realization of PS showed statistical significance (Model $\chi^2(6) = 84.7$, $p < 0.001$) and remained significant even after controlling for the heterogeneity in speakers in the random effects logistic regression

model. Thus, we can consider the factor "grammatical person" to be important in the use of PS.

In terms of verb semantics (**factor 2**), the results showed a higher overt pronoun rate with epistemic verbs (57%) than with perceptive verbs (39%) (Model $\chi^2(1) = 48.4$, p < 0.01), see Table 5:

Table 5. Overt pronoun rate according to verb semantics

Verb	%	N
Epistemic	57%	463/816
Perceptive	39%	318/810
Total	**48%**	**781/1626**

Interestingly, the preference for the use of PS with epistemic verbs can be observed for all grammatical persons, see Table 6:

Table 6. Overt pronoun rate according to grammatical person and verb semantics

Verb/Person	yo	nosotros	vos	usted	ustedes	él/ella	ellos/ellas
epistemic (%)	58%	44%	36%	77%	52%	65%	65%
perceptive (%)	35%	29%	30%	63%	43%	45%	30%
epistemic (N)	68/117	51/117	41/115	90/117	61/117	76/117	75/115
perceptive (N)	41/117	34/116	35/117	74/117	49/115	52/115	34/114

The difference between the use of PS with epistemic verbs and the use of PS with perceptive verbs was statistically significant for 1st persons, 3rd persons and the pronoun *usted* (*yo* and *ellos/ellas*: p < 0.001, *él/ella*: p < 0.01, *nosotros* and *usted*: p < 0.05). However, the interaction between verb semantics and grammatical persons was not statistically significant, suggesting that the magnitude of the differences between epistemic and perceptive verbs does not vary significantly between grammatical persons. The factor "verb semantics" can also be considered important in the use of PS.

Concerning **factor 3** (type of sentence), the global overt pronoun rate of subject pronouns was higher in interrogatives (*wh*-interrogatives 53%; absolute interrogatives 52%) than in declaratives (48%). Recall that the 1st persons were not examined for the factor "type of sentence", as the 1st persons mostly appear in declarative sentences, see Table 7:

Table 7. Overt pronoun rate according to type of sentence

Type of sentence	%	N
declarative	48%	186/387
absolute interrogative	52%	277/536
wh-interrogative	53%	124/236
Total (without 1st persons)	**51%**	**587/1159**

The relationship between the type of sentence and the use of PS was not statistically significant (Model $\chi^2(2) = 1.60$, p = 0.449), a finding further supported by the random effects logistic regression (Model $\chi^2(2) = 1.79$, p = 0.409). When comparing the overt pronoun rate of each individual grammatical person according to the type of sentence, the results show the following picture (Table 8):

Table 8. Overt pronoun rate according to grammatical person and type of sentence

Sentence	*vos*	*usted*	*ustedes*	*él/ella*	*ellos/ellas*
declarative (%)	38%	71%	46%	50%	36%
absolute interrogative (%)	34%	68%	44%	61%	52%
wh-interrogative (%)	22%	75%	57%	51%	56%
declarative (N)	29/77	56/79	35/77	39/78	27/76
absolute interrogative (N)	36/106	73/108	47/108	66/109	55/105
wh-interrogative (N)	11/49	35/47	27/47	23/45	27/48

As can be observed, the ranking of these factors according to overt pronoun rate as *wh*-interrogatives > absolute interrogatives/declaratives is supported by the 2nd persons *usted* and *ustedes*, as well as by the 3rd person plural *ellos/ellas*. Here, the results exhibited a significant interaction (Model $\chi^2(8) = 16.34$, p = 0.038). This indicates that the interaction between sentence type and grammatical person is more important in explaining the use of PS than the effect of the sentence type.

As for **factor 4**, there was a tendency to realize subject pronouns more frequently in simple matrix clauses (50%) or matrix clauses with subordinate clauses (54%) than in subordinate clauses (35%) (Model $\chi^2(2) = 32.1$, p < 0.001), see Table 9:

Table 9. Overt pronoun rate according to type of clause

Type of clause	%	N
simple matrix	50%	359/720
matrix with subordinate clause	54%	299/553
subordinate clause	35%	123/353
Total	**48%**	**781/1626**

The overt pronoun rate according to grammatical person and type of clause is summarized in Table 10:

Table 10. Overt pronoun rate according to grammatical person and type of clause

Type of clause	yo	nosotros	vos	usted	ustedes	él/ella	ellos/ellas
simple matrix cl. (%)	48%	38%	32%	77%	54%	47%	49%
matrix with subord. cl. (%)	51%	40%	38%	67%	51%	79%	61%
subordinate cl. (%)	35%	27%	27%	59%	29%	42%	26%
simple matrix cl. (N)	39/81	29/77	37/114	89/116	60/112	53/113	52/107
matrix with subord. cl. (N)	52/102	42/105	26/69	45/67	36/71	53/67	44/72
subordinate cl. (N)	18/51	14/51	13/49	30/51	14/49	22/52	13/50

The relationship between the overt pronoun rate according to grammatical person and type of clause was highly statistically significant for the 3rd person plural (*ellos/ellas*; p<0.001) and less so for the pronouns *usted* and *ustedes* (p < 0.05). The random effects logistic regression confirmed the significant main effect of grammatical person (Model $\chi^2(6) = 18.67$, p = 0.0048), but not of the clause (Model $\chi^2(2) = 3.45$, p = 0.18). Neither any significant interaction was found here. Therefore, I consider the factor 'clause' to be quite important in explaining the dependent variable.

As a next step, I further observed the use of overt pronoun subjects by the interviewed speakers. Of course, the effect of the speakers was already controlled for all regression models. Nevertheless, it is interesting to show that the overt pronoun rate of the 13 interviewed speakers laid between 35% and 79% (Model $\chi^2(12) = 86.9$, p < 0.001). The Table 11 summarized this "between-speaker" variability.

It is, however, difficult to state whether a systematic speaker effect exists or whether each speaker uses the null-overt subject pronoun

randomly. For example, comparing Speaker_08 and Speaker_09, which differ greatly in overt pronoun rates (35% vs. 79%), I observed that the factor "grammatical person" still has a strong effect on the use of PS. Both speakers exhibit different tendencies, however: While Speaker_08 prefers to express the PS with 1st persons, Speaker_09 prefers to express PS with 2nd persons (especially *usted* and *ustedes*). Since I performed the random effects logistic regressions and determined that 7% of the variance in the propensity to use a pronoun can be attributed to the individuals, I consider the effect of the speakers on the use of PS to be lower than the combination of four observed factors.

Table 11. Overt pronoun rate according to speaker

Speaker	%	N
Pers_01	44%	55/126
Pers_02	60%	74/123
Pers_03	44%	54/124
Pers_04	56%	71/126
Pers_05	56%	70/126
Pers_06	37%	46/124
Pers_07	42%	51/122
Pers_08	35%	44/126
Pers_09	79%	99/125
Pers_10	40%	50/126
Pers_11	40%	50/126
Pers_12	49%	62/126
Pers_13	44%	55/126
Total	**48%**	**781/1626**

In a final step, I measured the strength of effects of the examined factors (Pseudo R = McFadden's R2): The greatest effect was shown by the grammatical person (Pseudo R = 0.04), followed by the type of verb (Pseudo R = 0.025) and the type of clause (Pseudo R = 0.018). The effect of the type of sentence was smaller than that of the remaining factors (Pseudo R = 0.002). The ranking of the strength of the effects of the factors under observation is exposed in (9):

(9) Scale of the strength of effects of examined factors
GRAMMATICAL PERSON > TYPE OF VERB > TYPE OF CLAUSE > TYPE OF SENTENCE

Technically, this analysis was conducted using the intercept model, i.e. a model without any independent variables. This model predicted that the speakers do not realize the PS in 52% of the cases (recall that the speakers omitted the PS in 52% of the cases and realized the PS in 48% of the cases (N = 1626)). After including the intervening factors, I observed a 10% improvement in predictive power compared to the model with intercept only (Pseudo R = 0.10). Comparing the influence of each individual factor, I added the independent variables incrementally to the model only with intercept term. First, I introduced the grammatical person, which improved the model by 4%. Second, I included the sentence type and improved the model by 0.2%. Third, I added the type of verb, which further improved the model by 2.5%. Finally, I included the type of clause, which improved the model by 1.8%.

Appendix II is also of interest in this context, which illustrates the random effects model with only the main effects in order to keep the model as simple as possible. Here, all the four factors are summarized with their odds ratios (OR) and probability (p-) values. The OR-values are a measure of effect size, describing the strength of the association between the dependent and independent variables. The odds represent the *probability* of an event occurring *divided* by the *probability* of an event not occurring. Using the model we can observe, for instance, that the odds (chance) for the realization of *usted* (OR = 4.283) are 4.28 times higher than the odds for *yo* ($p < 0.001$).

5. Discussion

In this section I will discuss the primary findings (5.1), i.e. the results presented in Section 4, as well as the secondary findings presented in 5.2 with regards to topics such as word order and the pragmatic interpretation of the PS found in the data.

5.1 Primary findings

In the production experiment utilized in this study, the interviewed speakers only omitted PS in 52% of the cases. This means that the PS was realized in 48% of all cases, a number which seems to be quite high for a pro-drop language. For such reasons, some scholars have proposed a reformulation of the term "null-subject language". For example, Posio (2012) suggests the terminology "**languages with variable subject**

expression, as subject pronoun expression need not be in any way more marked or less frequent than their omission in a language pertaining to this category, and it is indeed subject to considerable variation" (Posio 2012: 6). Nevertheless, the term "variable subject expression" would only apply to personal sentences in Spanish, but not to impersonal sentences or sentences with inanimate reference (cf. Example 1).

Notice that the overt pronoun rate obtained by the production experiment (48%) in this study is much higher than that in the data analyzed from the empirical studies on *Porteño* by Barrenechea and Alonso (1977, 21%) and Soares da Silva (2006, 32%). Such a difference could be due to the sort of data utilized in my study, which did not stem from free interviews but rather were based on semi-spontaneous speech. In the experiment carried out here, no narrative style was given and a considerable use of epistemic verbs was observed (recall that a narration supports a decrease in the overt pronoun rate, while epistemic verbs supports an increase in the overt pronoun rate).

Two hypotheses were tested and confirmed by the experimental data. The **first hypothesis** suggested that the overt realization of PS depends not only on focus, contrast and morphological ambiguity, but also on other linguistic factors. As for the focus and contrast, the results clearly show that subject pronouns are not always contrastive or focal expressions of their null counterparts. I will discuss this issue further in Section 5.2. In terms of verbal syncretism, it was demonstrated that the use of null/overt PS is not subordinated to the verbal morphology: The PS can be realized in spite of a "rich" verbal agreement inflection without ambiguous interpretation. This also implies that the rich subject-verb agreement is not necessarily a direct cause of the pro-drop, as stated by some grammarians (see e.g. Alarcos Llorach 1994). I would thus support Ackema and Neeleman's statement that "there is an indirect relation between rich agreement and pro drop: rich agreement facilitates pro drop in more contexts" (Ackema & Neeleman 2007: 81).

I now turn to the use of PS according to the four observed linguistic factors. As for **factor 1**, the grammatical person showed the strongest effect on the use of PS. The results also confirmed the **second hypothesis** stating that the overt pronoun rate varies among the grammatical persons, despite these having been controlled for the same or very similar conditions (factors 2–4). Regarding the correlation between the overt pronoun rate and grammatical persons, 3rd and 1st persons showed rather balanced frequencies (3rd singular 55% vs. 3rd plural 48%; 1st singular 47% vs. 1st

plural 36%), whereas 2nd persons showed noticeable differences between *vos* (33%), *usted* (70%) and *ustedes* (47%). Interestingly enough, the 1st person plural *nosotros* and the 2nd person singular (familiar) *vos* exhibited the lowest overt pronoun rates of all grammatical persons (*nosotros* 36%; *vos* 33%). As for 3rd persons, these showed much higher overt pronoun rates (48% singular; 55% plural) compared to those taken from the free interview data (see Tables 2 and 3). This implies that the 3rd persons are expressed more frequently in interactive speech, despite being interpreted as given or familiar information. Furthermore, the expression of 3rd person pronouns seems to exhibit a sort of repetition or echo effect between speaker and interlocutor in statement/question sequences (a similar finding was reported in a study on the use of full NPs as subjects by Dumont 2006). As for 2nd persons, the high rate of overt pronoun of *usted*, *ustedes* can probably be explained by the nominal origin *vuestra merced* ('your grace'), as indicated in the literature (e.g. DPD 2005; RAE 2010). The fact that the pronoun *usted* constitutes a formal address and its plural form *ustedes* is informal (in the data of the present study) might explain the differences between their overt pronoun rates (70% vs. 47%). It is worth mentioning that in some studies, the pronouns of 2nd persons are also interpreted as vocatives for attracting the attention of the addressee (see e.g. Platzack & Rosengren 1994; Alonso-Cortés 1999). Interestingly, the speakers combined the subject pronouns of the 2nd person together with the first name (or surname, in the case of *usted*) and/or the Argentinean colloquial vocative expression *che* ('hey' or 'man'). However, I did not observe any correlation between the null-overt subject variation and the usage of other nominal vocative expressions. Concerning verb semantics (**factor 2**), which exhibited the second strongest effect, the following tendency could be observed across all grammatical persons: Epistemic verbs clearly exhibited higher overt pronoun rates than perceptive verbs. This finding coincides with those from several previous studies (e.g. Pešková 2011). As for **factors 3** and **4**, the type of sentence and clauses did not show as a clear tendency as did factor 2. For instance, comparing the overt pronoun rates of *ellos/ellas* (3PL) vs. *él/ella* (3SG), we can observe that *él/ella* was overtly realized in 50% of declaratives, whereas *ellos/ellas* is found in only 36% of declaratives, despite both grammatical persons being controlled by exactly the same conditions. Due to this lack of consistency, I will not attempt to provide an explanation for this interaction. Interestingly enough, this factor had the smallest effect on the

usage of PS of all those examined. This result contests the findings of some previous studies (e.g. Lu 1997).

5.2   Secondary findings

In this section I will very briefly discuss the properties of the overt pronominal subjects at the syntax-pragmatics (and partly prosody) interface. The data indicate interesting variation in word order with respect to the position of the overt PS in a sentence. Syntactically, the subjects were realized as left- or right-dislocated elements (10a–b) or as clause-internal arguments (10c) in interrogatives:

(10)   a.   *¿Ustedes   qué   opinan,   chicos?*
             you.NOM.PL what think.3PL.PRES.IND   boys
             'What do you think, boys?'

        b.   *¿Vieron   la última película de Almodóvar   ustedes   en cine?*
             see.3PL.PAST.IND the last movie by Almodóvar   you.NOM.PL in cinema?
             'Did you see the last movie by Almodóvar at the cinema?'

        c.   *¿Mira   usted   programas deportivos en la tele?*
             watch.3SG.PRES.IND   you.NOM.SG sports programs   on the TV?
             'Do you watch sports programs on TV?'

As for the right- and left-dislocated subjects, we can suggest that the specifier position within the inflectional phrase (IP) is occupied by an empty subject category *pro* (in Generative terminology), which is co-referent with the dislocated subject (e.g. [CP *Ustedes*i *qué* [IP proi *opinan*]], 'What do you think?').

In terms of preverbal subjects in transitive sentences with unmarked word order SVO (for example *Vos mirás programas raros en la tele*, lit. 'You watch strange programs on TV'), it is more difficult to determine whether the preverbal subjects are left-dislocated elements or internal arguments (for a discussion on this issue see e.g. Vallduví 1993; Alexiadou & Anagnostopoulou 1998; Gutiérrez Bravo 2007; López 2009). Furthermore, the SVO order was also observed frequently in yes-no questions (e.g. *¿Usted mira programas deportivos en la tele?*, lit. 'You watch sports programs on TV?'). Perhaps even more interestingly, the preverbal subject was stated four times, even in *wh*-questions (e.g. *¿Qué tipo de música usted escucha?*; lit. 'What kind of music you listen to?').

Nevertheless, the subject-verb inversion in *wh*-questions is supposedly obligatory in Spanish (see e.g. Torrego 1984). According to the present data, the intervening subject only seems to be tolerated in *Porteño* with a complex *wh*-word such as *qué tipo de música* ('what kind of music'). Interestingly, Goodall (2010: 237) established the following hierarchy for the ability of a *wh*-phrase to allow an intervening subject in Spanish: *why > complex wh-phrase > how > where/when > what/who*. It is also noteworthy that in nine (out of 186) declarative sentences, the subject occurred post-verbally in the sentence-final position (i.e. VOS or VS):[5]

(11) a. *No, escuchamos   músicaclásica nosotros.*
          no listen.1PL.PERS.IND music classicalwe.NOM/M
          'No, we listen to classical music.'

      b. *Che, qué música linda que   escuchás   vos.*
          hey   what music   nice   that listen.2SG.PERS.IND you.NOM.SG
          'Hey, what a nice music you are listening to.'

In Spanish, the subject pronouns shifted to the rightmost position of the sentence (with transitive verbs) can be interpreted either as (1) focused constituents which obligatorily bear nuclear stress and are thus prosodically prominent (see e.g. Zubizarreta 1998; Zagona 2002; Gabriel 2007; Gabriel, Feldhausen, Peskova, Colantoni, Lee, Arana & Labastia 2010), or as (2) right-dislocated elements (topics) lacking prosodic prominence (see e.g. Bosque & Gutiérrez-Rexach 2009; Gabriel 2010). I assume that all subject pronouns on the right edge of the sentence such as those in (11) should be interpreted as "afterthought" topics and not as the focus of the present data. There are two reasons for this: First, the situations presented to the speaker included a non-focused subject (pragmatic argumentation), and second, these pronominal subjects were deaccented (prosodic argumentation).

    We can thus say that all of the overt subjects in the data as well as their empty counterparts are either familiar (given) or aboutness-shift (new) topics (cf. Frascarelli & Hinterhölzl 2007). They neither present disambiguation nor fulfill the domain of focus or contrastive topics (cf. 3a

---

[5] The situations presented to the speaker were: (11a) *Ustedes están hablando con un amigo sobre música. Le dicen que escuchan música clásica* ('You [informal, plural] are talking with a friend about music. You tell him that you listen to classical music'); (11b) *Entrás en casa de tu amiga. Ella pone música. Decile que es muy lindo lo que escucha* ('You [informal, singular] enter your friend's house. She turns on some music. Tell her that what she is listening to is very nice').

and 3b). Nevertheless, the explicit PS might present a hidden contrast. The RAE (2010: §33.5e) argues that verbs describing an opinion may establish a hidden contrast between the speaker and other persons in the sequence as in *Yo creo que va a llover* ('I think it is going to rain'). Such an explanation would, however, not apply for other verbs (e.g. perceptive) with non-obligatorily expressed pronominal subjects.

In sum, the overt expression of PS is not required, but rather is optional in many contexts. In addition, the present production experiment has shown that the degree of "optionality" in the use of subject pronouns or empty categories depends on the following intervening factors: Grammatical person (strongest effect) > Verb semantics > Type of clause > Type of sentence (smallest effect).

6.   Concluding remarks

This paper investigated the use of PS in semi-spontaneous data as obtained by a production experiment with 13 monolingual speakers of *Porteño* Spanish. There were two main goals to the study: First, to determine whether a correlation exists between an overt pronoun rate and four selected intra-linguistic factors (grammatical person, verb semantic, type of sentence, type of clause), and second, to show that investigations using experimental data may lead to a better understanding of the examined phenomenon and of the usage of grammar in a natural way. As for the first goal, the experiment provided evidence for the expression of PS, which not only had reasons such as ambiguity, contrast or emphasis, but was also dependent on other intervening factors such as grammatical person or verb semantics. As for the second goal, it was shown that experimental techniques play a very important role in allowing us to uncover this type of evidence. As the inductive method used in the present paper also generates a comparable and controllable set of data, it can also be applied to a fine-grained analysis of further intra-linguistic and extra-linguistic factors, as well as in cross-dialectal or cross-linguistic research and in doing so significantly facilitates the empirical corpus-based investigation of the use of PS in Spanish and other null-subject languages.

References

Ackema, Peter & Neeleman, Ad (2007) Restricted pro drop in early modern Dutch. *The Journal of Comparative Germanic Linguistics* 10 (2): 81–107.

Aijón Oliva, Miguel Ángel & Serrano, María José (2010) El hablante en su discurso: Expresión y omisión del sujeto de *creo* [Speaker in his discourse: Expression and omission of the subject of *I think*]. *Oralia* 13: 7–38.

Alarcos Llorach, Emilio (1994) *Gramática de la lengua española* [Grammar of the Spanish Language]. Madrid: Espasa Calde.

Alexiadou, Artemis & Anagnostopoulou, Elena (1998) Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16: 491–539.

Alonso-Cortés, Ángel (1999) *La exclamación en español. Estudio sintáctico y pragmático* [Exclamation in Spanish. Syntactic and Pragmatic Study]. Madrid: Minerva Ediciones.

Alonso-Ovalle, Luis; Fernández-Solera, Susana; Frazier, Lyn & Clifton, Charles Jr. (2002) Null vs. overt pronouns and the topic-focus articulation in Spanish. *Rivista di Linguistica* 14 (2): 1–19.

Amaral, Patrícia Matos & Schwenter, Scott A. (2005) Contrast and the (non-) occurrence of subject pronouns. In David Eddington (ed.), *Selected Proceedings of the 7th Hispanic Linguistics Symposium*, pp. 116–127. Somerville, MA: Cascadilla Proceedings Project.

Badia Margarit, Antoni M. (1988) La omisión del sujeto en español [Omission of the subject in Spanish]. In Pedro Peira (ed.), *Homenaje a Alonso Zamora Vicente*, pp. 361–367. Madrid: Castalia.

Balasch, Sonia (2008) La conectividad discursiva en el discurso interactivo [Discourse connectivity in the interactive discourse]. In Joyce Bruhn de Garavito & Elena Valenzuela (eds.), *Selected Proceedings of the 10th Hispanic Linguistics Symposium*, pp. 300–311. Somerville, MA: Cascadilla Proceedings Project.

Barrenechea, Ana María (ed.) (1987) *El habla culta de la ciudad de Buenos Aires* [Educated Speech of the City of Buenos Aires]. Buenos Aires: Universidad Nacional de Buenos Aires.

Barrenechea, Ana María & Alonso, Alicia (1977) Los pronombres personales sujetos en el español hablado en Buenos Aires [Pronominal subjects in the Spanish spoken in Buenos Aires]. In Juan M. Lope Blanch (ed.), *Estudios sobre el español hablado en las ciudades principales de América*, pp. 333–349. México: Universidad Autónoma de México.

Bayley, Robert & Pease-Álvarez, Lucinda (1997) Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9: 349–371.

Bentivoglio, Paola (1987) *Los sujetos pronominales de primera persona en el habla de Caracas* [Pronominal Subjects of the First Person in the Speech of Caracas]. Caracas: Universidad Central de Venezuela.

Biberauer, Theresa; Holmberg, Anders; Roberts, Ian & Sheenan, Michelle (2010) *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: Cambridge University Press.

Bosque, Ignacio (1989) Clase de sujetos tácitos [Type of Unexpressed Subjects]. *Philologica* II: 91–111.

Bosque, Ignacio & Delmonte, Violeta (eds.) (1999) *Gramática descriptiva de la lengua española* [Descriptive Grammar of Spanish]. Madrid: Espasa.

Bosque, Ignacio & Gutiérrez Rexach, Javier (2009) *Fundamentos de sintaxis formal* [Foundations of Formal Syntax]. Madrid: Akal.

Cabredo-Hofherr, Patricia (2006) Arbitrary pro and the theory of pro-drop. In Peter Ackema, Patrcik Brandt, Maaike Schoorlemmer & Fred Weerman (eds.), *Arguments and Agreement*, pp. 230–260. Oxford: Oxford University Press.

Carminati, Maria Nella (2002) *The Processing of Italian Subject Pronouns*. University of Massachusetts: Amherst.

Chomsky, Noam (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

—— (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.

DPD = Real Academia Española (RAE) / Asociación de Academias de la lengua española (ed.) (2005) *Diccionario panhispánico de dudas* [Pan-Hispanic Dictionary of Doubts]. Madrid: Santillana.

Dryer, Matthew S. (2005) Expression of pronominal subjects. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Dumont, Jenny (2006) Full NPs as subjects. In Almeida Jacqueline Toribio & Nuria Sagarra (eds.), *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, pp. 286–296. Somerville, MA: Cascadilla Proceedings Project.

Enríquez, Emilia V. (1984) *El pronombre personal sujeto en la lengua española hablada en Madrid* [Personal Pronoun Subject in the Spanish Language Spoken in Madrid]. Madrid: Consejo Superior de Investigaciones Científicas.

Fernández Soriano, Olga F. (1999) El pronombre personal. Formas y distribuciones. Pronombres átonos y tónicos [Personal pronoun. Forms and distributions. Atonic and tonic pronouns]. In Ignacio Bosque & Violeta Delmonte (eds.), *Gramática descriptiva de la lengua española*, pp. 1209–1273. Madrid: Espasa.

Fischer, Susann (2010) *Word-Order Change as a Source of Grammaticalisation*. Amsterdam: John Benjamins.

Frascarelli, Mara (2007) Subjects, topics and the interpretation of referential pro: An interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory* 25 (4): 691–734.

Frascarelli, Mara & Hinterhölzl, Roland (2007) Types of topics in German and Italian. In Susanne Winkler & Kerstin Schwabe (eds.), *On Information Structure, Meaning and Form*, pp. 87–116. Amsterdam: John Benjamins.

Gabriel, Christoph (2007) *Fokus im Spannungsfeld von Phonologie und Syntax. Eine Studie zum Spanischen*. Frankfurt am Main / Madrid: Vervuert.

—— (2010) On focus, prosody, and word order in Argentinean Spanish. A minimalist OT account. *Revista Virtual de Estudos da Linguagem* 4: 183–222.

Gabriel, Christoph; Feldhausen, Ingo; Pešková, Andrea; Colantoni, Laura; Lee, Su Ar; Arana, Valeria & Labastía, Leopoldo (2010) Argentinian Spanish intonation. In Pilar Prieto & Paolo Paolo Roseano (eds.), *Transcription of Intonation of the Spanish Language*, pp. 285–317. München: Lincom.

Goodall, Grant (2010) Experimenting with wh-movement in Spanish. In Karlos Arregi, Zsuzsanna Fagyal, Silvina A. Montrul & Annie Tremblay (eds.), *Romance*

Linguistics 2008: Interactions in Romance, pp. 233–248. Amsterdam: John Benjamins.

Gutiérrez Bravo, Rodrigo (2007) Prominence scales and unmarked word order in Spanish. *Natural Language and Linguistic Theory* 25: 235–270.

Henríquez Ureña, Pedro (1939) ELLO [It]. *Revista de filología hispánica* 1 (3): 209–229.

Hinzelin, Marc-Olivier & Kaiser, Georg A. (2006) El pronombre *ello* en el léxico del español dominicano [The pronoun *it* in the vocabulary of the Dominican Spanish]. In Wiltrud Mihatsch & Monika Sokol (eds.), *Language Contact and Language Change in the Caribbean and Beyond / Lenguas en contacto y cambio lingüístico en el Caribe y más allá*, Sprachen, Gesellschaften und Kulturen in Lateinamerika 9, pp. 171–188. Frankfurt am Main: Peter Lang.

Hochberg, Judith G. (1986) Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62 (3): 609–621.

Hurtado, Luz Marcela (2001) *La variable expresión del sujeto en el español de los colombianos y colombo-americanos residentes en el condado de Miami-Dade* [The Variable Expression of the Subject in the Spanish of Colombian and Colombo-American Residents in the County of Miami-Dade]. Unpublished doctoral dissertation. University of Florida. <http://etd.fcla.edu/UF/anp4021/final.PDF> (24 June 2013).

Isabelli, Casilde A. (2004) The acquisition of NSP properties in SLA: Some effects of positive evidence in a natural learning setting. *Hispania* 87: 150–162.

Kaiser, Elsi (2011) Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus. *Language and Cognitive Processes* 26: 1625–1666.

Kaiser, Elsi & Trueswell, John C. (2008) Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes* 23 (5): 709–748.

Lipski, John (2002) Subject pronoun usage among Spanish dialects. Unpublished manuscript. <http://www.personal.psu.edu/jml34/pronouns.pdf> (24 June 2013).

López, Luis (2009) *A Derivational Syntax for Information Structure*, Oxford Studies in Theoretical Linguistics 23. Oxford: Oxford University Press.

Lu, Hui-Chuan (1997) La presencia y ausencia del pronombre personal [Presence and absence of the personal pronoun]. *Epos. Revista de filología* 13: 117–133.

Luján, Marta (1999) Expresión y omisión del pronombre personal [Expression and omission of personal pronoun]. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, pp. 1275–1315. Madrid: Espasa Calpe.

Montalbetti, Mario M. (1984) *After binding. On the interpretation of pronouns*. Unpublished doctoral dissertation. Massachusetts Institute of Technology.

Newmeyer, Frederick J. (2003) Grammar is grammar and usage is usage. *Language* 79 (4): 682–707.

Orozco, Rafael & Guy, Gregory R. (2008) El uso variable de los pronombres sujetos: ¿qué pasa en la costa Caribe colombiana? [Variable use of pronominal subjects: What happens in the Caribe Colombian Coast?]. In Maurice Westmoreland &

Juan Antonio Thomas (eds.), *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*, pp. 70–80. Somerville, MA: Cascadilla Proceedings Project.

Otheguy, Ricardo; Zentella, Ana Celia & Livert, David (2007) Language contact in Spanish in New York toward the formation of a speech community. *Language* 83 (4): 770–802.

Pešková, Andrea (2011) La omisión y la expresión del pronombre sujeto vos en el español porteño [Omission and expression of subject pronoun vos in Porteño Spanish]. In Ángela Di Tullio & Rolf Kailuweit (eds.), *El español rioplatense*, pp. 49–76. Frankfurt am Main/Madrid: Vervuert.

Platzack, Christer & Rosengren, Inger (1994) On the subject of imperatives: A minimalist account of the imperative pronoun and negated imperatives. *Sprache und Pragmatik* 34: 26–67.

Posio, Pekka (2008) *Uso del pronombre personal sujeto de la primera persona del singular en español y portugués hablados: factores semánticos y pragmáticos* [Use of Personal Pronoun of the First Person Singular in Spoken Spanish and Portuguese: Semantic and Pragmatic Factors]. Unpublished master's thesis. University of Helsinki.

—— (2011) Spanish subject pronoun usage and verb semantics revisited: First and second person singular subject pronouns and focusing of attention in spoken Peninsular Spanish. *Journal of Pragmatics* 43: 777–798.

—— (2012) *Pronominal subjects in Peninsular Spanish and European Portuguese. Semantics, pragmatics, and formulaic sequences*. Doctoral dissertation. Helsinki, University of Helsinki Press, <http://urn.fi/URN:ISBN:978-952-10-83203-7> (24 June 2013).

Prieto, Pilar (2001) L'entonació dialectal del català: el cas de les frases interrogatives absolutes [Dialectal intonation of Catalan: The case of yes-no questions]. In August Bover, Maria-Rosa Lloret & Mercè Vidal-Tibbits (eds.), *Actes del Novè Col·loqui d'Estudis Catalans a Nord-Amèrica*, pp. 347–377. Barcelona: Publicacions de l'Abadia de Montserrat.

RAE (1973) = Real Academia Española (1973) *Esbozo de una nueva gramática de la lengua española* [Draft of the New Grammar of Spanish Language]. Madrid: Espasa Calpe.

RAE (2010) = Real Academia Española (2010) *Nueva gramática de la lengua española* [New Grammar of the Spanish Language]. Madrid: Espasa Calpe.

Rizzi, Luigi (1982) *Issues in Italian Syntax*. Dordrecht: Foris.

—— (1986) Null object in Italian and the theory of pro. *Linguistic Inquiry* 17: 501–557.

Rosengren, Per (1974) *Presencia y ausencia de los pronombres personales sujeto en español moderno* [Presence and Absence of the Personal Pronoun Subjects in the Modern Spanish]. Stockholm: Almqvist and Wiksell International.

Silva-Corvalán, Carmen (2001) *Sociolingüística y pragmática del español* [Sociolinguistics and Pragmatics of Spanish]. Washington: Georgetown University Press.

Soares da Silva, Humberto (2006) *O parâmetro do sujeito nulo: Confronto entre o português o e espanhol* [Null-Subject Parameter: Confrontation Between

Portuguese and Spanish]. Unpublished master's thesis. Rio de Janeiro: Faculdade de Letras/UFRJ.

StataCorp. (2009) *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Toribio, Almeida Jacqueline (2000) Setting parametric limits on dialectal variation in Spanish. *Lingua* 10: 315–341.

Torrego, Esther (1984) On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15: 103–129.

Vallduví, Enric (1993) *The Informational Component*. New York: Garland.

Zagona, Karen (2002) *The Syntax of Spanish*. Cambridge: Cambridge University Press.

Zubizarreta, María Luisa (1998) *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.

Appendix I. Questionnaire for the elicitation of the 2nd person singular vos

1. *Estás hablando con tu amigo Manuel. Preguntale qué <u>opina</u>: ¿va a terminar pronto la crisis mundial?*
   'You are talking to your friend Manuel. Ask him what he thinks: will the world crisis be over soon?'
   [2nd pers. SG; epistemic verb; *wh*-question; simple matrix clause]

2. *Estás hablando con tu padre. Preguntale qué <u>opina</u> de Buenos Aires.*
   'You are talking to your father. Ask him what he thinks about Buenos Aires.'
   [2nd pers. SG; epistemic verb; *wh*-question; simple matrix clause]

3. *Estás jugando con tu sobrinito Felipe. Preguntale si <u>cree</u> en los fantasmas.*
   'You are playing with your nephew Felipe. Ask him if he believes in ghosts.'
   [2nd pers. SG; epistemic verb; absolute question; simple matrix clause]

4. *Estás hablando con tu amiga Mariana sobre diferentes países. Preguntale si <u>se imagina</u> estar viviendo en otro país.*
   'You are talking to your friend Mariana about different countries. Ask her if she can imagine living in another country.'
   [2nd pers. SG; epistemic verb; absolute question; simple matrix clause]

5. *Estás por salir. Ves que hay nubes y le preguntás a tu madre si <u>cree</u> que va a llover.*
   'You are going out. You see it is cloudy and ask your mother if she thinks that it is going to rain.'

[2nd pers. SG; epistemic verb; absolute question; matrix with subordinate]

6. *Estás hablando con tu amigo Manuel sobre Juan y Mariana. Preguntale si <u>sabe</u> que se van a casar.*
'You are talking with your friend Manuel about Juan and Mariana. Ask him if he knows that they are going to get married.'
[2nd pers. SG; epistemic verb; absolute question; matrix with subordinate]

7. *Es lunes. Por la tarde tenés el curso de inglés. Tu hermano te llama para salir. Decile que <u>sabe</u> que los lunes nunca podés.*
'It is Monday. In the afternoon you have English lessons. Your brother calls you to go out. Tell him that he knows that you cannot on Mondays.'
[2nd pers. SG; epistemic verb; declarative; matrix with subordinate]

8. *Empezaste a estudiar rumano. Tu hermano cree que se parece bastante al castellano. Decile que es mucho más difícil de lo que <u>piensa</u>.*
'You have started to learn Romanian. Your brother thinks it is very similar to Spanish. Tell him that it is more complicated than he thinks.'
[2nd pers. SG; epistemic verb; declarative; subordinate clause]

9. *Tu hermana te cuenta sobre su intención de pedir un préstamo. Pero no conoce los detalles. Decile que si <u>piensa</u> que es tan fácil, se equivoca.*
'Your sister tells you about her intention to ask for a loan. But she does not know all the details. Tell her if she thinks that it is so simple, she is wrong.'
[2nd pers. SG; epistemic verb; declarative; subordinate clause]

10. *Querés poner música. Preguntale a tu amiga Mariana qué <u>escucha</u>: ¿tango o bossa nova?*
'You want to turn on some music. Ask your friend Mariana what she listens to: tango or bossa nova?'
[2nd pers. SG; perceptive verb; *wh*-question; simple matrix clause]

11. *Preguntale a tu amigo Manuel qué estación de radio <u>escucha</u>.*
'Ask your friend Manuel what radio station he listens to.'
[2nd pers. SG; perceptive verb; *wh*-question; simple matrix clause]

12. *Estás hablando con tu hermano sobre el cine. Preguntale si <u>vio</u> la última película con Brad Pitt.*
'You are talking with your brother about the cinema. Ask him if he

saw the last movie with Brad Pitt.'
[2nd pers. SG; perceptive verb; absolute question; simple matrix clause]

13. *Preguntale a tu madre si <u>mira</u> programas políticos en la tele.*
'Ask your mother if she watches political programs on TV.'
[2nd pers. SG; perceptive verb; absolute question; simple matrix clause]

14. *Estás hablando con Manuel sobre los amigos de ustedes. Preguntale si <u>se enteró</u> de que Fernando se va a casar con Elena.*
'You are talking with Manuel about your friends. Ask him if he heard that Fernando was going to get married to Elena.'
[2nd pers. SG; perceptive verb; absolute question; matrix with subordinate]

15. *Hablás con tu compañero sobre la complicada situación en el trabajo. Preguntale si <u>se da cuenta</u> de cuál es el problema.*
'You are talking with your colleague about the complicated situation at work. Ask him if he realizes what the problem is.'
[2nd pers. SG; perceptive verb; absolute question; matrix with subordinate]

16. *Vivís con una amiga. Ella siempre se la pasa mirando programas raros en la tele. Un día le decís que <u>mira</u> programas rarísimos.*
'You live with one friend. She always spends her time watching weird programs on TV. One day you tell her that she watches very weird programs.'
[2nd pers. SG; perceptive verb; declarative; matrix with subordinate]

17. *Entrás en casa de tu amiga. Ella pone música. Decile que es muy lindo lo que <u>escucha</u>.*
'You enter the house of your friend. She turns on some music. Tell her that what she is listening to is very nice.'
[2nd pers. SG; perceptive verb; declarative; subordinate]

18. *Hace tiempo que no te ves con tus amigos. Te encontrás con Manuel. Decile que si <u>ve</u> a los chicos un día, que les mande muchos saludos.*
'You have not seen your friends for a while. You meet Manuel. Tell him that if he sees the boys one day, he should send them your regards.'
[2nd pers. SG; perceptive verb; declarative; subordinate]

Appendix II. Random Effects Logistic Model of the collected data

N = 1626	**Odds Ratio**	Std. Err.	z	**P-Value**	95% CI for Odds Ratio	
					Lower Limit	Upper Limit
Grammatical Person						
1SG‡	–	–	–	–	–	–
1PL	0.623	0.125	-2.36	0.018	0.420	0.923
2SG	0.740	0.175	-1.27	0.205	0.465	1.178
vd	4.283	1.029	6.06	<0.001	2.675	6.859
vds	1.486	0.346	1.7	0.088	0.942	2.345
3SG	2.109	0.492	3.2	0.001	1.335	3.332
3PL	1.492	0.348	1.72	0.086	0.945	2.357
Sentence						
declarative sentence*	–	–	–	–	–	–
absolute question	0.585	0.108	-2.9	0.004	0.408	0.841
wh-question	0.633	0.141	-2.05	0.04	0.409	0.979
Verb						
epistemic*	–	–	–	–	–	–
perceptive	0.451	0.050	-7.16	<0.001	0.363	0.561
Clause						
subordinate clause*	–	–	–	–	–	–
matrix with sub.	3.061	0.573	5.98	<0.001	2.121	4.418
simple matrix	2.886	0.547	5.59	<0.001	1.990	4.184

The estimated intraclass correlation is 0.073

McFadden's R^2=0.09

* Reference category

Contact Information:

Andrea Pešková
Universität Hamburg
Institut für Romanistik
Von-Melle-Park 6
20146 Hamburg
e-mail: andrea(dot)peskova(at)uni-hamburg(dot)de