

# Variable First Person Singular Subject Expression in Spoken Valencian Catalan\*

Peter Herbeck

Bergische Universität Wuppertal & Universität Wien  
peter.herbeck@univie.ac.at



Received: February 6, 2022  
Accepted: September 9, 2022

## Abstract

In this paper, I analyze first person singular subject pronoun expression in spoken Valencian Catalan varieties. I present a quantitative and qualitative analysis of 7 interviews from the corpus *Parlars*, examining 1sg subject rates regarding tense markings (syncretic vs. non-syncretic forms), verb type, and discourse type (monologues vs. conversations). It will be shown that the factor of morphological syncretism as well as verb type influence 1sg subject expression rates. Differently from what has been observed in some previous studies on Spanish, cognitive verbs are not associated with particularly high 1sg subject expression rates, but verbs of saying are. Looking at frequent verb forms in the spoken interviews, it will be shown that *(jo) dic* ‘(I) say’ plays an important role, direct speech being a commonly used mechanism in the data examined here. Furthermore, discourse type (monologue vs. conversation) affects subject expression rates. These results suggest that 1sg subject expression is influenced by an interaction of verb morphology, verb type and specific verb forms, but that subject expression rates and the factors influencing them vary depending on particular *pro*-drop varieties, discourse types and the type of data. Looking at the contexts in which the strong pronoun *jo* ‘I’ is used with verbs of saying in spoken Valencian Catalan, I argue that subject expression triggers a perspectival, rather than referential shift in several cases.

**Keywords:** strong pronouns; Valencian Catalan; subject expression; *pro*-drop; perspective; quantitative study; corpus study

\* Acknowledgements: Parts of this investigation have been conducted during a research stay at the Departament de Filologia Catalana of the Universitat de València in 2021. I would like to thank the members of the department and, especially, Andreu Sentí for providing me access the data of the corpus *Parlars* and for his continuous support during my work with the data of spoken Valencian Catalan. This research has been supported by a postdoc grant from the DAAD (*Deutscher Akademischer Austauschdienst*) which has made my stay at the Universitat de València possible. Parts of this research have been conducted in the research group “Linguistic variation in Catalan (VaLingCat)”, University of València (Ref: GIUV2017-397). Furthermore, I would like to thank two anonymous reviewers for helpful comments.

**Resum.** *L'expressió variable de la forma de subjecte de la primera persona singular en varietats orals del valencià*

Aquest article estudia la realització del pronom de subjecte de primera persona singular en varietats orals del valencià. Es presenta una anàlisi quantitativa i qualitativa de 7 entrevistes del corpus oral *Parlars* i s'examinen les freqüències dels pronoms explícits i implícits tenint en compte les variables de temps verbal (formes amb sincretisme i sense), tipus de verb i tipus de discurs (monòleg o conversa). Es demostrarà que tant el factor del sincretisme morfològic com el factor del tipus de verb influeixen en la realització del pronom de primera persona singular. A diferència dels resultats d'alguns estudis previs de l'espanyol peninsular, els verbs cognitius en les dades del valencià oral no es correlacionen amb freqüències altes de subjectes explícits de primera persona singular, però sí que s'hi associen els verbs de comunicació. En l'anàlisi de les formes verbals concretes, es mostrarà que (*jo*) *dic*, un mecanisme de narració en estil directe, és freqüent i té un paper important en aquest tema. A més, el tipus de discurs (monòleg o conversa) també afecta la realització del subjecte. Aquests resultats suggereixen que tant la morfologia verbal, com el tipus de verb i les formes verbals concretes influeixen en la realització del subjecte de primera persona singular, però que els factors poden variar segons el tipus de *pro*-drop, de discurs i de dades. L'anàlisi amb detall dels contextos en què el pronom fort *jo* es combina amb verbs de comunicació ens fa pensar que l'expressió del subjecte de primera persona singular provoca un canvi de perspectiva (i no pas un canvi referencial) en molts casos.

**Paraules clau:** pronoms forts; valencià; expressió del subjecte; subjecte nul; perspectiva; estudi quantitatiu; estudi de corpus

### Table of Contents

1. Introduction	6. Frequent first-person singular verb forms and subject expression
2. Background: morphological ambiguity between 1sg and 3sg	7. Discussion
3. The factor of verb type and verb forms	8. Conclusion
4. The study: methodology, data and coding	References
5. Results	Corpus
	Software/Tools

## 1. Introduction

Catalan, as is well-known (e.g., Rigau 1988; Solà 1992; Rosselló 2000), is a language that allows subjects to be dropped, similarly to other Romance languages like European Portuguese, Italian and Spanish (Rizzi 1982; Barbosa 1995). The variable expression of subjects has been linked to several factors, such as emphasis and contrast (Luján 1999), topic continuity (Bentivoglio 1983), and switch reference (Cameron 1994, 1995; Silva-Corvalán 2001; Travis & Torres Cacoullos 2012; etc.).

It has been argued that pronominal or 'strong' agreement morphology is one factor for the formal possibility of null subjects (see e.g. Rizzi 1982) in 'consistent'

*pro*-drop languages (in the sense of Holmberg 2005).<sup>1</sup> Relatedly, one factor that has been assumed to influence the use of an explicit or implicit subject pronoun is (morphological) ambiguity (see Hochberg 1986). That is, overt pronouns are used if the reference of the subject cannot be recovered from verb morphology or from context. Such a reasoning predicts that overt subject pronouns are more frequently used in contexts of morphological syncretism, in which verb inflection cannot discriminate between two (or more) potential referents. The hypothesis that syncretism in verbal morphology correlates with higher subject expression rates has come to be known as the *functional hypothesis* (Hochberg 1986).

In Catalan (as in European Portuguese and Spanish), a prediction would be that first person singular (1sg) and third person singular (3sg) subject expression rates are higher in the indicative imperfective and conditional as well as the subjunctive present and imperfective paradigms, where a syncretism between these person/number specifications can be found, than in paradigms without syncretism. Table 1 presents this difference for the verb *cantar* ‘sing’.<sup>2</sup>

**Table 1.** Verb inflection in (Valencian) Catalan with and without syncretism between 1sg and 3sg

	1sg = 3sg	1sg ≠ 3sg
	<b>present (subjunctive)</b>	<b>present (indicative)</b>
1sg	<i>cant-e</i>	<i>cant-e</i>
2sg	<i>cant-es</i>	<i>cant-es</i>
3sg	<i>cant-e</i>	<i>cant-a</i>
	<b>conditional</b>	<b>future</b>
1sg	<i>cantar-ia</i>	<i>cantar-é</i>
2sg	<i>cantar-ies</i>	<i>cantar-às</i>
3sg	<i>cantar-ia</i>	<i>cantar-à</i>
	<b>imperfect (indicative)</b>	<b>periphrastic past</b>
1sg	<i>cantav-a</i>	<i>vaig cantar</i>
2sg	<i>cantav-es</i>	<i>vas cantar</i>
3sg	<i>cantav-a</i>	<i>va cantar</i>
	<b>imperfect (subjunctive)</b>	
1sg	<i>cant-ara</i>	
2sg	<i>cant-ares</i>	
3sg	<i>cant-ara</i>	

1. The relation between agreement and null subjects is more complex, as an anonymous reviewer points out, given that there are languages that do not have strong agreement morphology (e.g., Chinese; see Huang 1984; Jaeggli & Safir 1989 for discussion) but they allow (discourse-identified) null subjects. However, it might be the case that different factors are involved in sanctioning null subjects in different *pro*-drop types, agreement morphology being a relevant factor in Catalan-type *pro*-drop systems. In fact, as I will argue throughout my study of 1sg subject expression in Valencian Catalan, even within the group of ‘consistent’ *pro*-drop system, considerable variation can be found in subject expression among varieties, genres, and verb types on a quantitative level.
2. Verb inflection is subject to dialectal variation in Catalan (see Sanchis Guarner 1950; Veny 1984; Beltran & Segura 2019, and references, for an in-depth discussion).

In this paper, I present a study of 1sg verb forms and their implicit and explicit subjects in a corpus of spoken Valencian Catalan varieties. A focus on 1sg verb forms is motivated by the observation that 1sg/2sg speaker/addressee pronouns and third person discourse anaphoric ones correspond to very different functions. In fact, it has been argued that a fruitful area of research is considering subject expression separately for 1sg, 2sg and/or 3<sup>rd</sup> person verb forms (cf. Travis & Torres Cacoullos 2012; Geeslin & Gudmestad 2016; Posio 2018). Furthermore, 1sg verb forms are particularly frequent in spoken discourse so that they provide an interesting field for the investigation of the relation between pronoun expression rates, verb semantics, and routinized subject+verb sequences.

1sg subject pronoun expression with frequent verb forms has received considerable attention in the study of spoken Spanish, above all in the context of verbs of cognition (e.g. *(yo) creo* ‘(I) think/believe’, *(yo) sé* ‘(I) know’) and verbs of saying (e.g. *(yo) digo* ‘(I) say’); Davidson (1996); Aijón Oliva & Serrano (2010); Posio (2011); Herbeck (2021). In spoken discourse, these verb forms are highly frequent, and it has been argued that high or low rates of a verb form have consequences for subject expression frequencies (see Erker & Guy 2012).

However, for Catalan, only few quantitative studies of subject expression exist (Casanova 1998; de Prada Pérez 2010, 2015) and a study of Valencian Catalan from such a perspective is, to the best of my knowledge, absent. A detailed examination of Valencian Catalan varieties is interesting because there is an additional paradigm with syncretism between 1sg and 3sg in the indicative perfect in some varieties (*jo/ell h[a] cantat* ‘I/he have/has sung’; Veny 1984: 164; Beltran & Segura 2019).

Furthermore, the question of whether the same verb types or verb forms trigger high and low 1sg strong pronoun rates across different Romance languages and varieties is only seldomly addressed. Some evidence for the existence of variation among closely related languages with similar subject-verb agreement paradigms can be found in Posio’s (2013, 2014) study of Peninsular Spanish and European Portuguese. The author observes that 1sg *creer* ‘think’ is frequently associated with an overt pronoun in Peninsular Spanish if compared with *decir* ‘say’, but *dizer* ‘say’ has high 1sg subject pronoun frequencies in European Portuguese while *achar* ‘think’ has average 1sg subject pronoun rates in this language.

On an empirical level, the present paper intends to provide new data about (i) 1sg subject expression in a sample of spoken Catalan varieties and (ii) the relevance of the factors of morphological syncretism, verb type, discourse type (monologues vs. conversations) and highly frequent verb forms in spoken language for yielding subject expression rates. To reach these goals, the 1sg verb forms and their subjects of 7 interviews of the corpus of spoken Valencian Catalan *Parlars* (Montserrat & Segura 2020; Esplà-Gomis & Sentí *in prep*) have been analyzed taking into account the above mentioned factors. With respect to verb forms, a special focus will be on (i) *jo dic* ‘(and) I say’, which frequently appears with an overt pronoun. Here, I examine the discourse contexts in which this form is used.

The results of the present study seem to show that morphological syncretism plays a role for 1sg subject expression frequencies, but that verb type and, furthermore, specific verb forms have to be considered. In particular, if we compare

the results of the spoken Valencian Catalan varieties investigated here with those obtained in previous studies on Peninsular Spanish, different verb types and verb forms trigger high and low 1sg subject expression rates. Thus, while verbs of cognition such as *(jo) creo* ‘I think’ have been observed to have high subject expression frequencies in Peninsular Spanish (Posio 2011; Herbeck 2021), it is verbs of communication and especially *(jo) dic* ‘I say’ that has high 1sg subject expression rates in the spoken Valencian Catalan data examined here, similarly to what Posio (2013, 2014) observes for European Portuguese. This might indicate that verb types and specific verb forms differ among Romance varieties and languages with respect to their 1sg subject expression rates. However, a further result of the study is that subject expression is influenced by the data and discourse type: subject pronoun frequencies are significantly lower in monologues than in conversations, indicating that a high degree of interactivity favors 1sg subject expression rates.

On a theoretical level, I argue that 1sg strong pronouns indicate a perspectival, rather than referential or topic shift in several cases. I implement this observation by assuming that overt 1sg subject pronouns can be located in the specifier of a SA (Speech Act) phrase (Speas & Tenny 2003), where they are sanctioned by a [+shift] feature. That is, differently from 3rd person subject pronouns, which have been argued to indicate topic shift (Frascarelli 2007), 1sg subject pronouns indicate perspectival shift with certain verb types (cf. Herbeck 2021 for discussion of Spanish). Empirical evidence is drawn from an in-depth analysis of *(jo) dic* ‘(I) say’, where an overt subject pronoun correlates with a shift from external narrator to internal speaker perspective in several cases.

This paper is structured as follows: in section 2, I discuss verb inflection and syncretism between 1sg and 3sg in Valencian Catalan varieties. In section 3, the factor of verb type for subject expression will be discussed. Section 4 presents the data that have been used for the study and the methodology that has been applied. In section 5, I present the main results with respect to subject expression rates, tense marking, verb type, and discourse type. Furthermore, I discuss how the most frequent verb forms behave with respect to subject expression in section 6. Section 7 is devoted to a deeper discussion of the results, focusing on frequent 1sg subject+verb combinations in the light of contextual factors determining their use, especially with verbs of communication and cognition. In this section, I also propose a syntactic analysis of 1sg subject pronouns with verbs of communication, arguing that these pronouns trigger a perspectival shift and are licensed by a [+shift] feature in the specifier of the functional category SAP.

## 2. Background: morphological ambiguity between 1sg and 3sg

Verb inflection has been singled out as one factor for the possibility of implicit subjects in finite clauses. Furthermore, the *functional hypothesis* (Hochberg 1986) predicts that syncretism of verbal morphology leads to a higher frequency of explicit subjects.

In Valencian Catalan varieties, as in Standard Catalan and Spanish, there are cases of syncretism between 1sg and 3sg in the (indicative and subjunctive) imper-

fect, subjunctive present and the conditional paradigms, but not in (most)<sup>3</sup> indicative present, future, and indicative past ones (see Table 1 above). One prediction on a quantitative basis would be that those paradigms with syncretism sanction higher 1sg subject pronoun frequencies because, even though the discourse context disambiguates reference to the speaker and (co-)reference to a discourse antecedent in some cases, ambiguity arises in others:

- (1) *quan mon pare tenia la botiga / jo fea aiualimon en una  
 when my father had the shop I made water-lemon in an  
 gelaora  
 icemachine  
 ‘when my father had the store, I made ‘aiualimon’ (= iced lemon drink) with  
 an ice machine’* (Parlars, Algemesí)

In (1), if 1sg *jo* wasn’t overtly expressed with the verb form *fea* (‘I/he/she made’) the most natural interpretation would be co-reference with *mon pare* ‘my father’, but reference to the speaker would be possible given the appropriate discourse conditions. Insertion of *jo* ‘I’ resolves this potential ambiguity. Note that ambiguity is, of course, not the only factor triggering 1sg subject pronoun use with these verb forms:

- (2) *i... mos casàrem allí perquè jo no volia casar-me aci  
 and REFL married.1PL there because I not wanted marry.INF-REFL.1SG here  
 ‘and... we got married there because I didn’t want to get married here’*  
 (Parlars, Alginet)

In this sentence, even though the verb is morphologically ambiguous between 1sg/3sg, the ambiguity is resolved clause-internally by means of the 1sg reflexive *me* so that insertion of the strong pronoun would not be necessary for reasons of disambiguation. In (2), insertion of the strong pronoun *jo* ‘I’ is related to the notion of contrast because a set of (topic) alternatives (Büring 2003; Mayol 2010) containing the speaker and her husband is created. The insertion of the strong pronoun selects the speaker out of this set.

In spoken Valencian Catalan varieties, there is an interesting case of syncretism between 1sg and 3sg that does not arise in Standard Central Catalan or Standard Peninsular Spanish in the indicative present perfect, which is formed by means of the auxiliary *haver* ‘have’ and a participle. In Standard Catalan, 1sg and 3sg are distinguished by means of the forms *he* ‘(I) have’ and *ha* ‘(he/she) has’, respectively. In some spoken Valencian Catalan varieties, however, this paradigm shows a syncretism between 1sg and 3sg (cf. Veny 1984: 164; Beltran & Segura 2019) because *he* (1sg) and *ha* (3sg) are both realized as [a] (e.g., *(jo/ell/ella) [a] cantat*

3. In the regular inflection of the second conjugation paradigm, there is a syncretism between 1sg/3sg in some varieties (*jo/ell cap* ‘I/he fit(s)’), while the forms are distinguished in others (*jo cabo / ell cap*). See also fn. 6.

‘(I/he/she) have/has sung’). The following represent two examples from the interviews of Sollana and Agost, respectively:

- (3) *ara jo tamé m- m’ha portat conforme*  
 now I also me.TRUNC me-have.1SG/3SG acted according  
*ha cregut jo que me tenia que portar*  
 have.1SG/3SG thought I that REFL.1SG had to act (*Parlars*, Sollana)<sup>4</sup>  
 ‘Now, I have also acted according to how I believed that I had to act’
- (4) *Jo no m’ ha menejat*  
 I not REFL.1SG have.1SG/3SG moved  
 ‘I haven’t moved’ (*Parlars*, Agost)

This situation constitutes an interesting case because it has been argued that it is the discourse function of TAM markings that influence subject expression rates (see Silva Corvalán 2001) rather than morphological ambiguity. Silva Corvalán (2001) argues that past preterit forms put the focus on the event – they are “foregrounded”, while events in imperfective, conditional and subjunctive forms are “backgrounded” (cf. Silva Corvalán 2001: 161; Table 4.16). Given that overt realization of the subject puts the focus on the subject referent (and not the event), it is expected that pronoun frequencies are lower in preterit, where the event is ‘foregrounded’ than in imperfective and conditional forms, where the event is ‘backgrounded’ (cf. *ibid.* 162).

While the imperfective, conditional and subjunctive forms have been investigated in the light of this reasoning in Spanish, the present perfect in Valencian Catalan poses a new situation. In 6 out of 7 interviews of the Valencian Catalan varieties examined here, the 1sg and 3sg form of the auxiliary *haver* is not clearly distinguishable or the auxiliary is even sometimes fully dropped (*jo anat* ‘I [have] gone’).<sup>5</sup> These data might thus be interesting in the light of the question whether morphological syncretism plays a role for subject expression rates in 1sg.<sup>6</sup>

4. I add the gloss TRUNC if a word was interrupted/not completed by a speaker, e.g. in cases of hesitation as in (3).
5. In one of the interviews – the town of Tírig – the 1sg and 3sg forms of the auxiliary *haver* are systematically distinguished.
6. There are some further paradigms in Valencian Catalan varieties that show syncretism between 1sg and 3sg. For example, in the simple past, there is syncretism between 1sg and 3sg in the third conjugation paradigm (*jo/ell/ella dormí* ‘I/he/she slept’). However, even though in 4 interviews, the simple past is used, it is often substituted by the periphrastic past in 1sg (see also Veny 1984; Beltran & Segura 2019 for discussion), so that only few data points are available, which do not allow an exhaustive investigation of these verb forms.

Another case is the present tense in the regular inflection of the second conjugation paradigm, where a syncretism between 1sg and 3sg can be observed in some varieties (*jo/ell cap* ‘I/he fit(s)’) while in others these forms are distinguished (*jo cabo / ell cap*). However, there were no 1sg cases of these verb forms in the sample.

Lastly, in Standard Central Catalan, verbs in the third conjugation like *llegir* ‘read’ and *fregir* ‘fry’ contain an inchoative extension and 1sg and 3sg are distinguished morphologically – *jo llegeixo* ‘I read’ / *ell llegeix* ‘he reads’ or *jo renyeixo* ‘I quarrel’ / *ell renyeix* ‘he quarrels’. In some (but not all) Valencian Catalan varieties, on the other hand, these verbs have the same form in 1sg and 3sg: *jo/ell llig* ‘I/he read(s)’, *jo/ell riny* ‘I/he quarrel(s)’ (cf. GEIEC 2018; *Gramàtica Essencial*

In the next section, I discuss a further factor that has been examined in the present study: verb type and specific verb forms.

### 3. The factor of verb type and verb forms

It is well known that verb semantics or verb type have an important impact on subject expression frequencies (Enríquez 1984; Morales 1997; Posio 2011; among many others). In most studies on variable subject expression in Spanish, cognitive verbs and, especially, verbs that express the personal opinion of the speaker, have high subject pronoun frequencies in first person singular (Enríquez 1984; Travis & Torres Cacoullos 2012). However, the task of defining coherent (semantic) verb classes to investigate overt subject expression has notoriously been a difficult task (see Posio 2018 for discussion).

Thus, semantic verb classes do not behave uniformly with respect to subject expression rates if individual verb forms are considered. For example, even though *creer* ‘think’ and *saber* ‘know’ are both cognitive verbs, their subject expression rates in 1sg differ considerably in spoken Peninsular Spanish (see Posio 2011; Herbeck 2021 and references therein). In the studies of Posio (2011) and Herbeck (2021), 1sg subject pronoun expression rates are significantly higher with *(yo) creo* ‘(I) think’ than with *(yo) sé* ‘(I) know’. Enríquez (1984) observes that 1sg subject expression rates are higher with cognitive verbs that express the personal opinion of the speaker than in those that do not. Furthermore, variation in subject expression rates can be observed even with the same verb form, depending on its concrete discourse functions: Aijón Oliva & Serrano (2010) report that Spanish *(yo) creo* ‘(I) think’ has lower overt 1sg subject rates when it has the epistemic meaning of ‘think’, while it has higher expression rates if it has an argumentative use.

In addition to the observed variation within semantic verb classes, divergent results have been observed in different studies and dialects: As Posio (2018: 294) discusses, in Orozco’s (2015) study on Colombian Spanish, 1sg *saber* ‘know’ favors overt subject expression, contrary to Peninsular Spanish. Furthermore, Erker & Guy (2012) observe especially high subject pronoun frequencies with *(tú) sabes* ‘you know’ with Spanish speakers in New York, but 2sg *saber* has low overt subject rates in studies on Peninsular Spanish (Posio 2011; Herbeck 2021).

Some evidence that frequent verb forms might show varying patterns of subject expression across Romance languages comes from Posio’s (2013, 2014) comparison between European Portuguese (EP) and Peninsular Spanish: In Posio (2013), 1sg subjects of *dizer* ‘say’ appear in 29/37 (= 78%) of the cases in his EP data, but only in 33/110 cases (= 33%) with 1sg *decir* in Peninsular Spanish. Furthermore, Posio (2014: 16) argues that subject expression with *creo* ‘(I) think’ is twice as high as the average 1sg subject expression rate in Spanish, while *acho* ‘(I) think’ in EP has average 1sg subject expression rates.

---

*de la Llengua Catalana* 5.2; 5.5.2). These forms were not frequent in 1sg in the sample and the few ones that were produced recourse to alternative forms, maintaining a difference between 1sg and 3sg: *jo lliçc* ‘I read’ and *jo rinyç* ‘I quarrel’ (both in *Parlars*, Sollana).

Considering the necessity of more data from other Ibero-Romance languages and varieties, the present study attempts to provide new data in this context by means of an investigation of spoken Valencian Catalan.

#### 4. The study: methodology, data and coding

In this section, I present the methodology of the study of variable 1sg subject expression in Valencian Catalan varieties.

##### 4.1. Data: *The corpus Parlars*

I have examined 7 interviews of different localities of the corpus *Parlars* (Montserrat & Segura 2020; Esplà-Gomis & Sentí *in prep*). This corpus represents spoken Catalan of speakers older than 60 years (cf. Montserrat & Segura 2020 for discussion). The 7 examined interviews contain data from different localities: Alcoi, Agost, Algemesí, Alginet, Catarroja, Sollana and Tírig. Out of these interviews, 2 were (quasi-)monologues and 5 were semi-spontaneous conversations. In the former, an interviewer asks questions to guide the informant, but interferes as little as possible, while the informant provides longer answers in a continuous way. In the 5 conversations, a (linguist) interviewer proposes some topics of conversation (mainly past experiences, past life in the village, customs, activities, etc.), but the main conversation is maintained between the informants. The difference between these discourse types is crucial because the lower interactivity of monologues might have the consequence of yielding lower frequencies of expressed 1sg subjects (see section 5.3).

Table 2 summarizes the interviews that have been used for the study and their classification with respect to variety (following Beltran & Segura 2019) and discourse type.

Table 2. Annotated interviews of the corpus *Parlars*

interview	variety	length (approx.)	discourse type
Tírig	<i>valencià tortosí</i>	68:41 min	conversation
Catarroja	<i>valencià central</i>	37:59 min	conversation
Sollana	<i>valencià central</i>	62:45 min	(quasi-)monologue <sup>7</sup>
Algemesí	<i>valencià central</i>	39:23 min	conversation
Alginet	<i>valencià central</i>	74:09 min	conversation
Alcoi	<i>valencià meridional</i>	41:21 min	monologue
Agost	<i>valencià alacantí</i>	58:57 min	conversation

In the next section, the annotation categories of the data will be outlined.

7. In this interview, there were two informants but one of them almost didn't produce turns, so that the overall discourse type is more similar to a monologue than to a conversation. This is reflected by the fact that out of 315 1sg verb forms 312 were produced by the main informant and only 3 were produced by the other one.

#### 4.2. Data annotation

The texts already contained an orthographic transcription (see Beltran et al. 2019), and an automatic tokenization, lemmatization and morpho-syntactic annotation (see *ibid.*). All morpho-syntactic information referring to the verb (person/number, TAM) and the subject has been manually revised by myself. Furthermore, I have added the information with respect to subject expression and verb type fully manually.

The data have been annotated in a stand-off format in ELAN (Brugman & Russel 2004) with respect to the following categories for the study of subject expression: (i) person/number, (ii) TAM, (iii) subject expression, and (iv) verb type. The first three categories are straightforward: in the case of (i), all 1sg verb forms have been extracted for the current study. In category (ii), the following values have been differentiated: conditional, future, imperfect (indicative), imperfect (subjunctive), periphrastic past, simple past, present (indicative), present (subjunctive), and present perfect (indicative). Regarding (iii), the verb has been annotated for whether it appeared with an overt subject pronoun or not.

Category (iv) – verb type – was the most complex one, given the ambiguity of several verbs. Thus, this annotation has been carried out manually in a final step after morpho-syntax annotation. Table 3 shows the categories of verb type that have been analyzed.

This classification of verbs has benefited from the following sources: first, it is based on Enríquez's (1984) seminal classification of verbs into external states, external actions, internal cognition and verbs of personal opinion. It has further integrated findings of studies that posit more fine-grained classifications, such as the role of a subject as a 'communicator' (verbs of saying), a 'cognizer' (verbs of cognition), or a 'possessor' (verbs of possession), etc. (see Posio 2011). Lastly, in several cases I have consulted preexisting classifications of corpus projects on the annotation of verb semantics, such as Adesse for Spanish (cf. García-Miguel et al. 2010) and AnCora for Catalan (Aparicio et al. 2008).

#### 4.3. Excluded sentences

For the study of 1sg verb forms, the fixed sequences *no sé què* (... *no sé quants*) 'lit. (I) don't know what', and *vull dir* 'that is; lit. (I) wanna say' have been excluded from the data set, because they do not freely allow overt subjects and *jo què sé* ('What do I know?') has been excluded because it obligatorily requires an overt subject pronoun.

Furthermore, the following data was excluded: contexts when the speakers spoke Spanish, those parts where the speakers sang a song, truncated sentences where subject expression is unclear, direct repetitions of verbs, and all cases where the annotation category could not unambiguously be determined. Lastly, the data from the interviewees were excluded because they did not always represent the same age range or the same spoken variety of Valencian Catalan as the informants.

Table 3. Annotated (sub-)categories of verb type

main verb type	subtype	example
agentive	---	<i>fer</i> 'make', <i>comprar</i> 'buy', <i>llegir</i> 'read', <i>jugar</i> 'play', <i>riure</i> 'laugh', <i>treballar</i> 'work', etc.
unaccusative verbs and passives	unaccusative	<i>anar</i> 'go', <i>eixir</i> 'go out', <i>venir</i> 'come', <i>arribar</i> 'arrive', <i>entrar</i> 'enter', <i>caure</i> 'fall', <i>néixer</i> 'be born', etc.
	passive	all passives with <i>se/es</i>
verbs of communication	---	<i>dir</i> 'say', <i>fer</i> 'make' (introducing direct speech), <i>raonar</i> 'talk', <i>explicar</i> 'explain', <i>xerrar</i> 'talk', ...
statives without an accusative complement	attribution	<i>estar</i> 'be', <i>resultar</i> 'result', <i>parèixer</i> 'seem', ...
	existential	<i>viure</i> 'live', <i>existir</i> 'exist', <i>estar</i> 'be', ...
	<i>ser</i>	<i>ser</i> 'be'
statives + accusative complement	possession	<i>tenir</i> 'have', <i>necessitar</i> 'need', ...
cognitive verbs	thinking	<i>creure</i> 'think/believe', <i>pensar</i> 'think', ...
	knowledge	<i>saber</i> 'know', <i>reconèixer</i> 'recognize', <i>conèixer</i> 'know / be acquainted with', <i>recordar(-se'n)</i> 'remember', ...
	perception	<i>veure</i> 'see', <i>mirar</i> 'look', <i>escoltar</i> 'listen', <i>sentir</i> 'hear/feel', ...
	others	<i>voler</i> 'want/love', <i>tenir il·lusió</i> 'to be excited', <i>passar-s'ho bé</i> 'to have a good time', ...
dative psych-verbs	---	<i>agradar</i> 'like', <i>encisar</i> 'excite', <i>fer il·lusió</i> 'to be excited', ...
modal/aspectual + infinitive	modal	<i>voler</i> (+inf) 'want', <i>deure</i> 'shall/ought to', <i>poder</i> 'can', <i>tenir que/de</i> 'have to', <i>haver de</i> 'have to'
	aspectual	<i>acabar de</i> 'to finish; just', <i>començar a</i> 'to start', ...

#### 4.4. Statistical analysis

I analyzed the binary dependent variable *subject expression* with the values [overt] or [null] in relation to the independent variables *morphological syncretism*, *verb type* and *discourse type*. In the category *morphological syncretism*, all TAM categories that have morphologically equal forms between 1sg and 3sg (= [syncretism]: imperfective past (indicative + subjunctive), present subjunctive, conditional, and present perfect where applicable) were distinguished from those that differentiate 1sg and 3sg inflection (= [-syncretism]: present indicative, simple + periphrastic past, future, and present perfect where applicable). In the category *discourse type*, the interviews Agost, Algemesí, Catarroja, Sollana, and Tirig were grouped together as [conversation] and the interviews Alcoi and Sollana as [monologue]. The variable *verb type* includes the semantic verb classes described in Table 3 under the category "main verb type": agentive verbs, unaccusative/passivized verbs, verbs of communication, cognitive verbs, stative verbs without an accusative comple-

ment, statives with an accusative complement, and periphrasis containing modal/aspectual verbs.

I applied a binomial logistic regression model in R (R Core Team 2018). The reference category with respect to which the results are interpreted are the following: [null] for *subject expression*, [-syncretism] for the category *morphological syncretism*, [agentive verb] for *verb type*, and [conversation] for *discourse type*.

I first present the results descriptively for each variable and indicate the relevant coefficients and *p*-values. The results of the statistical model will be summarized for all variables in section 5.4.

## 5. Results

In this section, I present the results of the analysis of TAM and morphological syncretism (§ 5.1), verb type (§5.2), and discourse type (§5.3) with respect to 1sg subject expression.

### 5.1. TAM and morphological syncretism

Table 4 shows 1sg subject expression rates according to TAM for the 1607 analyzed sentences from the 7 interviews. The highest overt subject rates with 1sg verb forms arise with conditional (63%), imperfect (51% indicative, 45% subjunctive), present perfect (47%), and subjunctive present (46%). The average total overt pronoun rate is as high as 43%. It is interesting that those overt 1sg subject rates above the average are forms that are syncretic between 1sg and 3sg. This is especially interesting with the present perfect: as I have mentioned, in most of the Valencian Catalan varieties examined, syncretism between 1sg/3sg was found (not in Tírig, where the difference between 1sg and 3sg is systematically maintained in the perfect auxiliary). However, the TAM marking with the highest frequency –

**Table 4.** Overt and null subjects with 1sg verb forms according to TAM markings

TAM	null	overt	total	%-overt
conditional	3	5	8	63%
imperfective (indicative)	162	167	329	51%
present perfect	95	85	180	47%
present (subjunctive)	7	6	13	46%
imperfective (subjunctive)	17	14	31	45%
present (indicative)	474	327	801	41%
simple past	24	14	38	37%
periphrastic past	113	61	174	35%
future	24	9	33	27%
total	919	688	1607	43%

conditional – has only very few data points. Also subjunctive present and simple past forms were not frequent in 1sg in the sample.

Looking at the variable of *morphological syncretism* for 1sg subject expression, syncretic verb forms have higher 1sg subject expression frequencies (272/550 = 49.5%) than non-syncretic ones (416/1057 = 39%). The difference is significant, the probability of an overt 1sg subject pronoun increasing with syncretic verb forms ( $p < 0.001$ ; coefficient in logits: 0.49828; reference category: [-syncretism]).

## 5.2. Verb type

Figure 1 presents the descriptive results of 1sg subject expression and main verb type.<sup>8</sup> As can be seen, 1sg subject expression with cognitive verbs (42%) is close to the average overt 1sg pronoun rate in the sample (= 43%). Furthermore, agentive (39%) and unaccusative/passivized (41%) verbs also have similar rates of 1sg overt subjects. Interestingly, verbs of communication have higher than average 1sg subject pronoun frequencies in the examined data (53%), followed by (non-possessive) stative verbs (46%). According to the statistical analysis, the only significant difference with respect to the reference category [agentive verb] arises with verbs of communication (coefficient in logits: 0.70994;  $p < 0.001$ ). The other verb classes do not significantly differ in subject expression rates from the reference category ( $p > 0.05$ ).

Let us have a closer look at the subclasses of verbs of cognition. Table 5 presents the descriptive results of verbs of thinking, knowledge, perception and ‘oth-

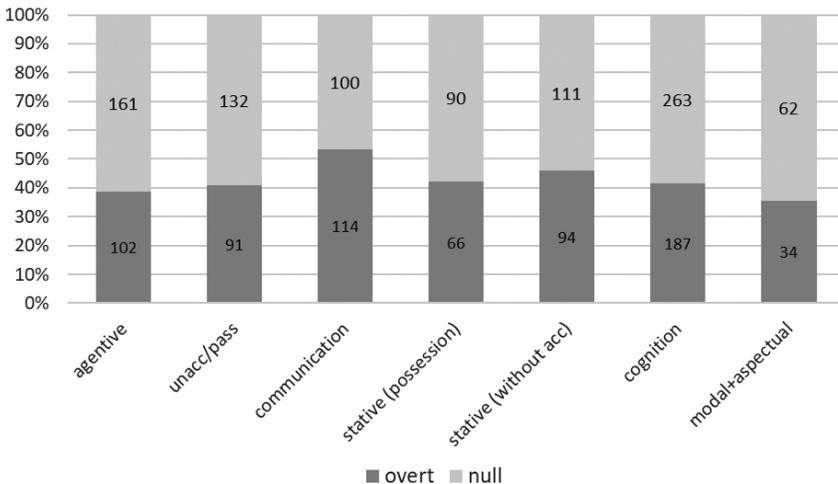


Figure 1. Null and overt 1sg subject pronouns according to verb type

8. There were no cases of dative psych verbs with 1sg inflection.

**Table 5.** Subject expression with 1sg cognitive verbs

verb of cognition	null	overt	total	%-overt
perception	19	25	44	57%
others	35	27	62	44%
thinking	33	22	55	40%
knowledge	176	113	289	39%
total	263	187	450	42%

ers'. Verbs of thinking like *creure* are generally not frequent in the corpus sample and the subject expression rate (40%) is similar to the average rate in the whole sample. Verbs of knowledge are the most frequent verb forms and their 1sg subject pronoun rate (39%) is also close to average. Lastly, verbs of perception have the highest subject pronoun rate ( $25/45 = 57\%$ ).

In section 6, I will have a look at specific verb forms in 1sg but before, I examine whether the type of interview influences subject pronoun expression in 1sg.

### 5.3. Discourse type and individual interviews

As I have mentioned, different interviews might have varying subject expression rates because they represent different discourse types. Recall that two interviews are either a monologue (Alcoi) or a quasi-monologue (Sollana), while the other five interviews were conversations between 2-4 speakers so that the latter involve a higher degree of interactivity. In contrast to the (quasi-)monologues, where the interviewer asks few questions and the informant offers longer responses, the conversations contained several turn takings and debates and, thus, overt subject frequencies might be higher in the latter interview type. Table 6 shows the descriptive results of 1sg subject expression in the different interviews.

It can be observed that the interviews of Algemesí (57%), Agost (50%), Catarroja (49%), and Tirig (46%) have the highest 1sg subject expression rates. Alcoi (33%), Alginet and Sollana (both 39%) have the lowest subject expression frequencies in 1sg. Recall that two of the interviews with lowest 1sg subject expression rates – Alcoi and Sollana – are (quasi-)monologues. In fact, the monologue Alcoi, which has the highest overall frequency of 1sg verb forms, has the lowest 1sg subject pronoun rate. In this interview, the informant talks about past (personal) experiences from his perspective in a continuous way, so that neither perspectival shifts nor turn taking are expected to have an impact on 1sg subject expression.

According to the statistical model outlined in section 4.4, the group of (quasi-) monologues has significantly lower 1sg subject expression frequencies than the interviews representing conversations ( $254/713 = 36\%$  vs.  $434/894 = 49\%$ , respectively;  $p < 0.001$ ; coefficients in logits:  $-0.59286$ , reference category: [conversation]).

**Table 6.** 1sg subject expression in different interviews

interview	null	overt	total	%-overt
Algemesí	53	69	122	57%
Agost	179	178	357	50%
Catarroja	72	70	142	49%
Tírig	80	68	148	46%
Alginet	76	49	125	39%
Sollana	191	124	315	39%
Alcoi	268	130	398	33%
total	919	688	1607	43%

#### 5.4. Interim summary

The results indicate that morphological syncretism (1sg/3sg) of the verb form as well as the type of interview (monologue vs. conversation) influences 1sg subject expression rates in the data examined here. This is plausible if conversations have a higher degree of interactivity, in which turn taking and perspectival changes are especially relevant. In monologues, on the other hand, the speaker's perspective is maintained continuous so that neither referential nor perspectival shift is expected to be as relevant as in conversations. With respect to verb type, verbs of communication have highest 1sg subject pronoun expression rates (and not verbs of cognition), differently from what has been found in several studies on Spanish.

Figure 2 summarizes the results of the quantitative analysis of 1sg subject expression rates in the sample of spoken Valencian Catalan according to the statistical model outlined in section 4.4.<sup>9</sup>

In Figure 2, those values with blue lines have a higher probability and those with red lines have a lower probability of triggering an overt 1sg subject pronoun if compared with the reference category. Those values marked with an asterisk show a significant effect (\*\*\*) standing for  $p < 0.001$ , while those without an asterisk are not significant ( $p > 0.05$ ). Thus, verbs of communication significantly favor 1sg subject expression (compared to the reference category [agentive verbs];  $p < 0.001$ ) while the probability of an overt 1sg subject does not significantly increase with the other verb types. Furthermore, syncretic verb forms significantly favor 1sg subject expression ( $p < 0.001$ ) and monologues disfavor overt 1sg subjects if compared to conversations ( $p < 0.001$ ).

## 6. Frequent first-person singular verb forms and subject expression

Apart from subject expression with different verb types, the question arises how the most frequent 1sg verb forms behave in the examined sample. I first extracted

9. The graph has been created with the sjPlot package (Lüdecke 2021).

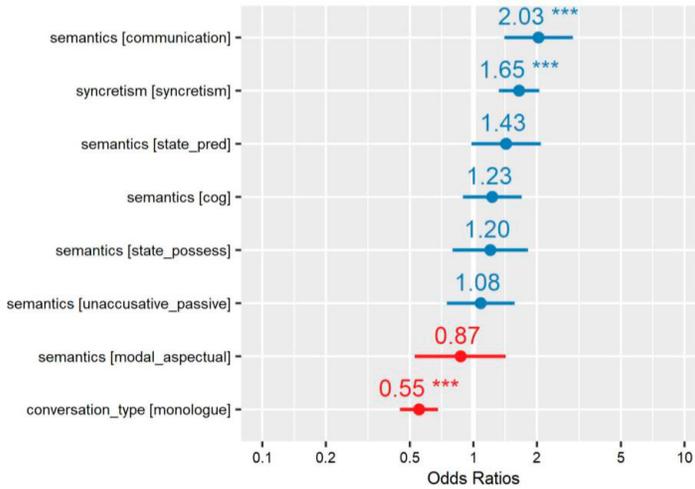


Figure 2. Binomial logistic regression model of 1SG subject pronoun expression frequencies in a sample of spoken Valencian Catalan

a verb list together with their annotations and determined the most frequent forms with the program AntConc (Anthony 2018). Thereafter, overt subject frequencies have been calculated and will be presented descriptively in this section.

With respect to 1sg verb forms, there were in total 1607 tokens and 329 types (including periphrastic tenses). Table 7 presents the most frequent verb forms in the investigated sample with a minimum of 20 tokens, together with their subject expression frequencies.

Recall that the average rate of overt 1sg pronouns is 43%. It is interesting to observe that *(jo) crec* ‘(I) think/believe’ (17/42 = 40%) does not have especially high subject expression rates and this verb form is generally not frequent in the sample. The reason might be that the expression of personal opinions is not frequent in the examined interviews, which are often about past experiences of the speakers, events and states. Volitional *(jo) vullc* ‘(I) want’ (12/23 = 52%) is among the cognitive verbs with the highest 1sg subject expression rate, even though the total number of this verb form is generally low in the sample. The most frequent verb forms of cognitive verbs in 1sg are *(jo) sé* ‘(I) know’ and *(jo) recorde* ‘(I) remember/recall’. However, subject pronoun frequencies differ considerably between these two (41/140 = 29% vs. 40/89 = 45%, respectively). The second most frequent verb form in 1sg in the data examined here is the verb of saying *(jo) dic* ‘(I) say’, which has high subject pronoun frequencies (65/114 = 57%). This verb form is often used by the speakers in past narrations introducing direct speech (see section 7).

Furthermore, there is an interesting situation with *anar* ‘go’: Subject expression is higher in imperfective *(jo) anava* (55%) than in the periphrastic past *(jo) vaig anar* (39%) and it is lowest with present tense *(jo) vaig* ‘(I) go’ (29%). Interestingly,

Table 7. Most frequent verb forms in 1sg and subject expression

rank	verb form	null	overt	total	%-overt
1	(jo) sé '(I) know'	99	41	140	29%
2	(jo) dic '(I) say'	49	65	114	57%
3	(jo) recorde '(I) remember'	49	40	89	45%
4	(jo) tinc '(I) have'	42	38	80	48%
5	(jo) tenia '(I) had' (imperfective)	38	26	64	41%
6	(jo) anava '(I) went' (imperfective)	24	29	53	55%
7	(jo) crec '(I) think/believe'	25	17	42	40%
8	(jo) vaig '(I) go'	25	10	35	29%
9	(jo) vaig anar '(I) went'	20	13	33	39%
10	(jo) era '(I) was' (imperfective)	15	17	32	53%
11	(jo) estic '(I) am'	20	10	30	33%
12	(jo) sóc '(I) am'	12	17	29	59%
13	(jo) he/ha anat '(I) have gone'	6	17	23	74%
13	(jo) vullc '(I) want'	11	12	23	52%
	(jo) estava '(I) was'	8	14	22	64%
14	(imperfective)				
14	(jo) puc '(I) can'	16	6	22	27%
	(jo) podia '(I) could'	14	7	21	33%
15	(imperfective)				
	(jo) fe(ia) '(I) made'	6	14	20	70%
16	(imperfective)				

subject pronouns are most frequent in present perfect (*jo*) *he/ha anat* 'I have gone' (17/23 = 74%), where a syncretism can be observed in 6 of the 7 examined interviews. A similar situation obtains with *estar* 'be' (64% overt 1sg pronouns in imperfect *estava* and 33% in present tense *estic*). However, syncretism in verb morphology cannot be the only factor for the observed frequencies: for example, indicative present (*jo*) *tinc* and imperfective (*jo*) *tenia* both have overt subject rates over 40%, with the present tense form having a slightly higher percentage (48% vs. 41%, respectively).

Another interesting case is the imperfective form of *fer* 'make' which has the second highest subject expression frequency with 14/20 = 70% overt subjects. As will be discussed in the next section, one factor is that in some of these cases, *fer* 'make' is in fact used as a verb of direct speech in (spoken) Catalan (see Schmid 2006), a use that also appears in the current Valencian Catalan sample:

- (5) *però perquè jo volia*                      *eh jo fea*                      “*Jo vuic...*”  
 but because I wanted.1SG/3SG eh I made.1SG/3SG I want.1SG  
 ‘but since I wanted, I said [lit. I made]: “I want...”                      (*Parlars*, Sollana)

This indicates that it is neither verb morphology, verb semantics, nor verb form *alone* that determines subject expression, but these factors jointly have an impact on 1sg pronoun rates.

## 7. Discussion

In this section, I discuss the results in the context of previous studies (section 7.1) and examine the discourse contexts in which subject pronoun + 1sg verb sequences appear, with a special focus on the sequence (*jo*) *dic* ‘(I) say’ and on verbs of cognition (see section 7.2). In section 7.3, I provide a syntactic analysis of overt 1sg subjects with verbs of communication.

### 7.1. Previous studies on subject expression in Catalan (and Spanish)

There is only a small number of previous quantitative studies on 1sg subject expression in Catalan: In Casanova (1998), the general rate of expressed subjects is 28%, but this number includes 2<sup>nd</sup> and 3<sup>rd</sup> person subjects.<sup>10</sup> De Prada Pérez (2015) investigates 1sg verb forms in her study of Spanish and Menorcan Catalan bilinguals. In her Catalan L1 bilingual group, 1sg subjects were produced in 21.3% of the 274 cases and in her Catalan control group (with Catalan-dominant speakers) the percentage is 20.7%. With the Spanish bilingual L1 speakers, overt 1sg subject rate is 19.9% and in the Spanish control group with speakers from Valladolid, it is 19.8% (cf. de Prada Pérez 2015: 126). Furthermore, the author does not observe any significant differences between the language groups. In the present study, 1sg subject expression rates are considerably higher (= 43%). However, further investigation is needed to determine whether these differences are due to variety-specific or corpus-specific features (such as discourse type and speech situation).

With respect to Peninsular Spanish, Enríquez (1984: 348) observes in her pioneering study of spoken Spanish in Madrid a subject expression rate of 31.89% in 1sg in the total sample. In Travis & Torres Cacoullos’ (2012) study of 1sg subject expression in Colombian Spanish, 49% of 1020 tokens were overt. However, it has frequently been observed that some Latin American Spanish varieties have higher subject pronoun frequencies than Peninsular Spanish (cf. e.g., Morales 1997).

In the present study of Valencian Catalan, 1sg subject expression is, thus, higher than in previous studies on Peninsular Spanish (Enríquez 1984) as well as

10. In Casanova’s (1998: 112) study, there are 242 overt subjects (= 28%). The author states that 103 of these are pronominal, out of which 55 are 1sg *jo* (Casanova 1998: 117). However, the total number of 1sg verb forms is not indicated so that this number cannot be directly compared with the results of the study presented here.

Menorcan Catalan (de Prada Pérez 2015). However, a word of caution is necessary before postulating cross-linguistic differences: even though some variety-specific factors might intervene (such as syncretism between 1sg and 3sg in the present perfect), it cannot be the only factor because subject expression has a high percentage of 41% also in the indicative present in our study and the most frequent verb forms like *(jo) dic* ‘(I) say’ (57% overt subjects) or *(jo) recorde* ‘(I) remember’ (45%) do not have any syncretism in the verb inflection. Furthermore, different studies rely on a different design and corpus data. Thus, I have shown that the type of corpus data and speech situation (monologue vs. conversation) can significantly affect 1sg subject expression.

### 7.2. Verb type and frequent verb forms: verbs of saying and verbs of cognition

With respect to verb type and verb form, it has been shown that verbs of communication have the highest overt pronoun rate in our sample of Valencian Catalan varieties. Verbs of cognition have average subject expression frequencies.

Interestingly, in Posio’s (2013, 2014) study of European Portuguese and Peninsular Spanish, differences between the two languages in 1sg subject pronoun frequencies with verbs of cognition and communication are linked to the overall frequency of the verb forms: the author argues that the combination of *yo* plus *creo* might have conventionalized in Spanish due to high frequency of use (cf. Posio 2014: 16). This would imply that the combination of 1sg *yo* with this cognitive verb has acquired a higher degree of fixation correlating with the more advanced grammaticalization of constructions with *creo* in Spanish (cf. Posio 2014: 16) than the corresponding combination with *achar* in EP.

In our Valencian Catalan data, *(jo) dic* ‘say’ is frequent in the sample and high expression rates of *jo* seem to be related to its frequent use in contexts of direct speech in the speakers’ narrations. In the sample, direct style is a very a common mechanism to narrate past events:

- (6) *I jo el llisc i dic “Sí” dic “És de l’ auelo”*  
 and I it read.1SG and say.1SG yes say.1SG (it)is of the grandfather  
*en sia diu “Tu la vols?” I jo dic “Jo sí”*  
 then says.3SG you it want.2SG and I say I yes  
 ‘And I read it and (I) say “yes” (I) say “It’s from grandfather” and then he says  
 “Do you want to have it?” And I say “Me, yes”.’ (Parlars, Sollana)

In (6), there is switch reference between two protagonists of a past narration: the speaker and her daughter, who is introduced in the immediately preceding discourse. When reference switches to the speaker after introduction of 3<sup>rd</sup> person *diu* ‘(she) says’, an overt 1sg pronoun is used in *I jo dic* ‘and I say’. Apart from alternating direct speech, a further factor is subjectivity because the speaker’s volition is in focus. In fact, the additional factor of subjectivity might be crucial because alternating speech is neither a necessary nor a sufficient condition for subject expression, as the 3<sup>rd</sup> person singular null subject with *diu* in (6) shows.

Furthermore, *jo* with *dic* signals a change of perspective rather than switch reference in several cases, the overt pronoun marking a shift from the *external* narrator to the speaker *inside* the narration:

- (7) *i aní al metge i li diguí a a A. /jo dic ... me*  
 and went.1SG to.the doctor and him said.1SG to to A. I say me  
*passa açò [...]*  
 happens that  
 ‘and I went to the doctor and I said to A....I say “This and that is happening to me...”’ (Parlars, Catarroja)

First, the speaker uses a simple past tense form of *dir* with a null subject – *diguí* ‘(I) said’ – and then the change to direct style is accompanied by a change from a morphologically marked past to narrative (or historic) present, i.e., to a morphologically marked present which nevertheless narrates a past event or state (cf. also Posio & Pešková 2020 for discussion of *digo* ‘(I) say’ in spoken Spanish).<sup>11</sup> This way, the narrator reports the events from the perspective inside the narration. This change from external to internal perspective is accompanied by the overt strong pronoun *jo* ‘I’.

This can be observed in various interviews, frequently in settings of alternating direct speech between two ‘protagonists’ of a narration, one of them being the speaker=narrator:

- (8) *em va dir “jo vaig a apuntar-me a música” I jo dic*  
 to.me PST.3SG say I go.1SG to sign.up-REFL.1SG to music and I say  
 “ah! pos jo també”  
 ah well I too (Parlars, Alcoi)  
 ‘and (he) said to me “I will sign up for music” And I say “Ah! Well me too”’

In (8), the 3<sup>rd</sup> person periphrastic past tense form of *dir* is used to introduce direct speech. When changing to the speaker=narrator perspective, there is a change to the narrative present and, at the same time, overt *jo* and the discourse marker *i* ‘and’ is used.

Thus, it seems to be the case that perspectival (and not only referential) shifts are crucial to account for the high frequency of *jo* ‘I’ with the 1sg verb forms of *dir* in the spoken Valencian Catalan data examined here: the change from external narrator to internal speaker within the narration is accompanied in several cases by a change from morphologically marked past to narrative present and overt expression of the 1sg subject.

The observation that the sequence (*i*) *jo dic* ‘(and) I say’ appears in very specific discourse contexts with a high frequency might have as a consequence that it has a degree of ‘fixation’ or, as Posio (2013, 2014) argues for several

11. For discussion of the historic present in a different context in Catalan – the evolution of the periphrastic past – see Pérez Saldanya & Hualde (2003).

subject+verb sequences in Spanish and Portuguese, it might have become a “formulaic sequence”. Routinization of 1sg pronoun + verb sequences and specialization for determined discourse functions is a typical phenomenon with verbs of cognition and of communication (cf. also SAY+comp structures in different Romance languages; Cruschina & Remberger 2008).

The relevance of the contextual factors of (alternating) direct speech and shift from external narrator to internal speaker perspective receives further support from the verb *fer* ‘make’, which can introduce direct speech in the spoken Valencian Catalan varieties examined here:

- (9) [...] *me'n tornava i ensia fea ma mare* “*ja has*  
REFL.1SG returned.1SG and then made my mother already has  
*vingut?*” *I jo fea* “He vingut?” [...]  
come.PRT and I made have.1SG come.PRT (Parlars, Sollana)  
‘[...] I returned and then my mother said [lit. ‘my mother made’]: “have you  
already come back?” And I said [lit. ‘I made’]: “If I have already come back?”  
[...]’

In this context, an overt 1sg pronoun frequently arises: In the sample, there are 59 forms of *fer* ‘make’ in 1sg, out of which 17 are used as a verb of saying introducing direct speech (as in (9)). Out of these, 11/17 (= 65%) appear with overt *jo*, compared to 16/42 (= 38%) with other types of *fer* ‘make’.

Turning to verbs of knowledge, different verb forms vary considerably with respect to subject expression frequencies: the most frequent verb form (*jo sé* ‘(I) know’ appears with an overt pronoun in 99/140 = 29% cases. This is a lower frequency than the average 1sg pronoun rate in the study of Valencian Catalan presented here. Note, however, that several cases of 1sg *saber* with a null subject contain semi-fixed sequences of the type *no (ho) sé* ‘(I) don’t know (it)’, which appear in answers to indicate the epistemic stance of the speaker:

- (10) *Bueno això em van dir no ho sé*  
well that to.me PAST.3PL say not it know.1SG  
‘Well, that’s what they told me, I don’t know’ (Parlars, Agost)

The second most frequent verb of knowledge in 1SG is (*jo recorde* ‘(I) remember’, which has an overt subject pronoun in 45% of 89 cases. 1sg subject expression seems to be related in several cases to evidentiality and/or epistemicity with this verb form (see Herbeck 2021 for Spanish (*yo creo* ‘I think’ and (*yo sé* ‘(I) know’). The concept of evidentiality depends on the relation of the speaker to the transmitted information - in particular, to whether the source of information is personal experience, direct (sensory), indirect (non-sensory) evidence, or hearsay (see Willett 1988; De Haan 2005; Rooryck 2001; and references). The concept of epistemicity relates to the degree of (in-)security, (un-)certainty or commitment the speaker has with respect to the truth of the transmitted information (see e.g. Givón 1982). With overt *jo*, the verb form *recorde* often stresses the speaker as the

source of information, i.e., it indicates that the transmitted information relies on personal experience, but also on subjective perception. At the same time, it relates to the speaker's epistemic stance in that it attenuates the speaker's commitment to the truth of what is being reported, as an anonymous reviewer points out:

- (11) *pues... ací... hi havien jo recorde que havien allà a l'eixida de*  
 well here there were I remember that were there at the-exit of  
*de la carretera allí havien unes paraes que ficaen els... sacs una*  
 of the road there were some stands that put.3PL the sacks a  
*mntanya aixina*  
 mountain like.that  
 'well ... here ... there were I remember that there, at the exit of the road, there  
 were some places where they put a lot of sacks...'  
 (Parlars, Alginet)

In this part of the discourse, the speaker starts objectively with existential *hi havien* 'there were' but changes to the personal, subjective form *jo recorde que* 'I remember that', indicating that the truth of the transmitted information depends on the speaker's perspective. In these examples, overt 1sg subject pronoun expression is not plausibly related to topicality, emphasis or contrast – it shifts from an objective description to a personal report from the perspective of the speaker's perception.

An interesting case is the form (*jo*) *h[a] conegut* '(I) have (got to) known', which is not among the most frequent verb forms in 1sg, but appears with an overt pronoun in most cases 13/16 (= 81%). In several instances, the 1sg present perfect form of *conèixer* with an overt pronoun appears together with the demonstrative *això* 'that':

- (12) *això jo no ho he conegut* (Parlars, Algemesí)  
 this I not it have.1SG known
- (13) *Això no ho he conegut jo* (Parlars, Agost)  
 this not it have.1SG known I  
 'I didn't know that.'

Even though 1sg *he* [a] cannot be morpho-phonologically distinguished from 3sg *ha* here, the high subject expression rate with this form of *conèixer* rather seems to be related to its frequent use in contexts of contrast: the demonstrative refers back to a statement of another interlocutor and, thus, the speaker indicates that s/he, in contrast to the interlocutor, does not have knowledge or is not acquainted with the described state or event.

However, the results indicate that syncretism between 1sg and 3sg does play a role in certain contexts. This also seems to be the case with different verb forms of *anar* 'go' which, as shown in Table 7, has highest overt subject frequencies with present perfect (17/23 = 74%) and imperfective '(I) went' (29/52 = 55%) and lower with the periphrastic past (39%) and with present tense (29%). Furthermore,

in some cases of the 1sg perfect of *anar*, the auxiliary is completely omitted if the strong pronoun is realized.<sup>12</sup> In these cases, person marking and reference is fully carried by the strong pronoun:

- (14) *Jo anava al jo no anat aci a les escoles no anat*  
 I went.1SG/3SG to.the I not gone here to the schools not gone  
*Jo anat a l- a les mongetes*  
 I gone to th- to the nuns.DIMINUTIVE (Parlars, Sollana)  
 ‘I haven’t gone to the schools here, I went to the ‘nuns’

Thus, syncretism seems to play a role in several examples of overt 1sg subject pronoun expression. However, as could be very clearly detected in the explorative study of concrete verb forms, frequently occurring subject+verb sequences have to be considered together with the discourse contexts in which they are used, leading to a routinization of the function, not only of the strong pronoun, but of the pronoun+verb combination.

### 7.3. Theoretical consequences: Perspective and null vs. strong speaker pronouns

I have argued that several instances of overt 1sg subject pronoun expression with epistemic verb forms (e.g. *(jo) recorde* ‘(I) remember’) and verbs of communication (e.g. *(i) (jo) dic* ‘(and) I say’) in spoken Valencian Catalan are the result of a perspectival, rather than referential or topic shift. With *dir* ‘say’, overt *jo* often indicates a shift from the external narrator to the internal speaker perspective inside a narration. This shift is accompanied by a shift from simple or periphrastic past to the narrative (or historic) present in several cases.

The question arises how this is encoded in the syntactic structure. It has been assumed in the literature that preverbal strong pronouns are (left-dislocated) topics (Ordóñez & Treviño 1999; Frascarelli 2007), similarly to clitic left dislocated elements. In fact, strong pronouns have often been argued to be shifting (Frascarelli 2007) or contrastive (Mayol 2010) topics. However, as argued in Herbeck (2021) for Spanish *creer* ‘think’ and *saber* ‘know’, several 1sg and 2sg strong subject pronouns seem to be *perspectival markers*, rather than topics, above all with verbs of cognition. In this approach, 1sg and 2sg subject pronouns, rather than being sanctioned by a [+shift] or [+contrast] feature in the specifier of a topic phrase (see (15)), are licensed by such a feature in the specifier of a perspectival functional projection – Cinque’s (1999) and Speas & Tenny’s (2003) Epist(emic) or Evid(entail) Phrase with verbs of cognition (see (16)):<sup>13</sup>

12. The similarity between the vowel of the auxiliary *h[a]* and the first vowel of the participle *anat* might play a role in these cases.
13. In Herbeck (2021), a clear-cut distinction between EpistP and EvidP is not made for the licensing of strong pronouns as perspectival markers, but the latter are located in one perspectival functional category (in the vein of Uriagereka’s 1995 FP), which can have different perspectival values or ‘flavors’.

(15) [<sub>TopP</sub> *Ell* Top<sub>[+contrast]</sub> [<sub>TP</sub> T-*Dorm* [<sub>vP</sub> *pro* v=*dorm* ... *molt*]]]

(16) [<sub>EpistP</sub> *Jo* Epist<sub>[+contrast]</sub> -*recorde* [<sub>TP</sub> T-*recorde* [<sub>vP</sub> ... *recorde* [<sub>CP</sub> *que* ...]]]]]

I suggest that the same reasoning accounts for 1sg subject expression with *dir* ‘say’ in spoken Valencian Catalan. As shown in the previous sections, 1sg *jo* with this verb form indicates a shift from external to internal perspective in several cases. This can be implemented with Cinque’s (1999) and Speas & Tenny’s (2003) Speech Act (SA) phrase above TP. In fact, Speas & Tenny (2003: 320) argue that the SAP projects a ‘speaker’ argument in its specifier, which is “the agent of the speech act” (*ibid.*):

(17) [<sub>SAP</sub> SPEAKER SA ...

I argue that the feature [+shift] can be assigned to the Spec of a SA phrase, just as it can be associated with a Spec,TopP. If the speaker argument is associated with [+shift], an overt, 1sg pronoun arises, which indicates a shift to internal speaker perspective within the narration:

(18) I [<sub>SAP</sub> *jo* SA<sub>[+shift]</sub> -*dic* [<sub>CP</sub> C ...

Due to high frequency of use, *dic* ‘say’ as a lexical V might have become a more functional element – a marker of direct speech – being directly merged in the category SA with the 1sg pronoun *jo* in its specifier. In this derivation, (*jo*) *dic* can be generated by external merge into the functional category encoding perspective, apart from a derivation where the verb and the subject pronoun are generated in the vP domain.

## 8. Conclusion

In this paper, I have investigated 1sg subject pronoun expression in spoken Valencian Catalan varieties through an examination of 7 interviews from the spoken Valencian Catalan corpus *Parlars*. The results show that morphological syncretism (1sg=3sg) significantly influences 1sg subject expression in the examined sample. With respect to verb type, cognitive verbs do not have significantly higher 1sg subject expression frequencies in the sample, but verbs of saying do. This is similar to what Posio (2013, 2014) observes in his study on European Portuguese. This seems to show that different verb types and verb forms have particularly high or low 1sg subject expression frequencies across languages and varieties. However, it has also been shown that further research is necessary to determine whether we are dealing with cross-linguistic patterns, given that the type of discourse situation influences subject pronoun expression. Thus, even though some variety specific properties – such as the existence of a further paradigm with syncretism between 1sg and 3sg in the present perfect – might influence subject expression, the type of discourse situation and the degree of interactivity is a further factor, as the comparison between monologues and conversations indicates.

In the sample of spoken Valencian Catalan, one interesting 1sg subject plus verb sequence is the combination of *(i +) jo + dic* ‘(and) I say’, which is frequently used with an overt pronoun. This is related to a common use of direct speech in past narrations in the data examined here. I have argued that strong pronoun expression with 1sg *dir* ‘say’ corresponds to a shift between two levels of narration: *I jo dic* ‘and I say’ is often used in past narrations in which the external narrator changes to the internal perspective of the speaker within the narration, which is commonly accompanied by a change from the (morphological) past tense to the narrative/historic present. High 1sg subject pronoun expression in the investigated sample (43% in total) might thus also partly be influenced by the overall high frequency of determined narrative strategies in spoken discourse.

The results of the study of spoken Valencian Catalan show that perspectival, and not only referential or topic shifts have to be considered for 1sg subject expression. In syntax, an overt strong pronoun can be the result of a [+shift] feature assigned to the Spec of a Speech Act Phrase with the result of overt realization of a ‘speaker’ argument in a perspectival functional projection above TP.

## References

- Aijón Oliva, Miguel Ángel & Serrano, María José. 2010. El hablante en su discurso: Expresión y omisión del sujeto de *creo*. *Oralia* 13: 7-38.
- Aparicio, Juan, Taulé, Mariona & Martí, Ma. Antònia. 2008. AnCora-Verb: A Lexical Resource for the Semantic Annotation of Corpora. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis & D. Tapias (eds.). *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Marocco. European Language Resources Association (ELRA), 797-802. <<https://aclanthology.org/L08-1>>
- Barbosa, Pilar. 1995. *Null subjects* (Doctoral dissertation), Massachusetts Institute of Technology.
- Beltran Calvo, Vicent & Segura-Llopes, Carles. 2019. *Els parlars valencians*. València: Publicacions de la Universitat de València.
- Beltran Calvo, Vicent, Esplà, Miquel, Segura-Llopes, Carles, Sentí, Andreu, Guardiola i Savall, Maria Isabel & Montserrat, Sandra. 2019. *Criteris per a la transcripció del corpus Parlars*. Segona versió. València: Universitat de València. <<https://roderic.uv.es/handle/10550/71244>> (16.06.2022).
- Bentivoglio, Paola. 1983. Topic continuity in discourse: A study of spoken Latin-American Spanish. In Givón, Talmy (ed.). *Topic continuity in discourse*, 255-312. Amsterdam: John Benjamins. <<https://doi.org/10.1075/tsl.3.06ben>>
- Brugman, Hennie & Russel, Albert. 2004. Annotating Multi-media/Multi-modal resources with ELAN. In M. Lino, M. Xavier, F. Ferreira, R. Costa & R. Silva (eds.). *Proceedings of the 4th International Conference on Language Resources and Language Evaluation (LREC 2004)*, 2065-2068. Paris: European Language Resources Association.
- Büring, Daniel. 2003. On D-trees, Beans, and B-accents. *Linguistics and Philosophy* 26: 511-545.

- Cameron, Richard. 1994. Ambiguous Agreement, Functional Compensation, and Nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5(3): 305-334.
- Cameron, Richard. 1995. The Scope and Limit of Switch Reference as a Constraint on Pronominal Subject Expression. *Hispanic Linguistics* 6(7): 1-27.
- Casanova, Lourdes. 1998. El sujeto en el catalán coloquial. *Revista Española de Lingüística* 29(1): 105-131.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. Oxford: Oxford University Press.
- Cruschina, Silvio & Remberger, Eva-Maria. 2008. Hearsay and Reported Speech: Evidentiality in Romance. *Rivista di Grammatica Generativa* 33: 95-116.
- Davidson, Brad. 1996. 'Pragmatic weight' and Spanish subject pronouns: The pragmatic and discourse uses of 'tú' and 'yo' in spoken Madrid Spanish. *Journal of Pragmatics* 26: 543-565.  
<[https://doi.org/10.1016/0378-2166\(95\)00063-1](https://doi.org/10.1016/0378-2166(95)00063-1)>
- De Haan, Ferdinand. 2005. Encoding speaker perspective: Evidentials. In Frajzyngier, Zygmunt, Hodges, Adam & Rood, David S. (eds.). *Linguistic diversity and language theories*, 379-397. Amsterdam: John Benjamins.  
<<https://doi.org/10.1075/slcs.72.18haa>>
- De Prada Pérez, Ana. 2010. Variation in subject expression in Western Romance. In Colina, Sonia, Olarrea, Antxon & Carvalho, Ana Maria (eds.). *Romance Linguistics 2009: Selected papers from the 39th Linguistic Symposium on Romance Languages (LSRL), Tucson, Arizona, March 2009*, 267-284. Amsterdam: John Benjamins.
- De Prada Pérez, Ana. 2015. First person singular subject pronoun expression in Spanish in contact with Catalan. In Carvalho, Ana M., Orozco, Rafael & Shin, Naomi L. (eds.). *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*, 121-142. Washington: Georgetown University Press.
- Enríquez, Emilia V. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Instituto Miguel de Cervantes.
- Erker, Daniel & Guy, Gregory R. 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88(3): 526-557.  
<<https://doi.org/10.1353/lan.2012.0050>>
- Esplà-Gomis, Miquel & Sentí, Andreu. In prep. An annotated informal and spoken corpus for dialectal conversations: The *Parlars* Corpus for Valencian Catalan. Ms.
- 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: 691-734.  
<<https://doi.org/10.1007/s11049-007-9025-x>>
- García-Miguel, José M., Vaamonde, Gael & González Domínguez, Fita. 2010. ADESSE. A Database with Syntactic and Semantic Annotation of a Corpus of Spanish. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC)*. Valletta (Malta), 17-23 de mayo.  
<[http://www.lrec-conf.org/proceedings/lrec2010/pdf/859\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2010/pdf/859_Paper.pdf)>
- Geeslin, Kimberly & Gudmestad, Aarnes. 2016. Subject Expression in Spanish. Contrasts between Native and Non-Native Speakers for First and Second-Person Singular Referents. *Spanish in Context* 13(1): 53-79.

- GEIEC. 2018. *Gramàtica essencial de la llengua catalana*. Barcelona: Institut d'Estudis Catalans. En línia: <geiec.iec.cat>
- Givón, Talmy. 1982. Evidentiality and epistemic space. *Studies in Language* 6(1): 23-49.  
<<https://doi.org/10.1075/sl.6.1.03giv>>
- Herbeck, Peter. 2021. Perspectival factors and pro-drop: A corpus study of speaker/addressee pronouns with *creer* 'think/ believe' and *saber* 'know' in spoken Spanish". *Glossa: a journal of general linguistics* 6(1): 122.  
<<https://doi.org/10.16995/glossa.5873>>
- Hochberg, Judith G. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language* 62(3): 609-621.
- Holmberg, Anders. 2005. Is There a Little pro? Evidence from Finnish. *Linguistic Inquiry* 36(4): 533-564.
- Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15(4): 531-574.
- Jaeggli, Osvaldo & Safir, Kenneth J. (eds.). 1989. *The Null Subject Parameter*. Dordrecht: Kluwer.
- Luján, Marta. 1999. Expresión y omisión del pronombre personal. In Bosque, Ignacio & Demonte, Violeta (eds.). *Gramática descriptiva de la lengua española*, 1275-1315. Madrid: Espasa.
- Mayol, Laia. 2010. Contrastive pronouns in null subject Romance languages. *Lingua* 120(10): 2497-2514.
- Montserrat, Sandra & Segura, Carles. 2020. Un corpus col·loquial i dialectal del valencià: PARLARS. *Zeitschrift für Katalanistik* 33: 9-44.
- Morales, Amparo. 1997. La hipótesis funcional y la aparición de sujeto no nominal: el español de Puerto Rico. *Hispania* 80(1): 153-165.
- Ordóñez, Francisco & Treviño, Esthela. 1999. Left dislocated subjects and the pro-drop parameter: A case study of Spanish. *Lingua* 107: 39-68.  
<[https://doi.org/10.1016/S0024-3841\(98\)00020-5](https://doi.org/10.1016/S0024-3841(98)00020-5)>
- Orozco, Rafael. 2015. Pronominal variation in Colombian Costeño Spanish. In Carvalho, Ana M., Orozco, Rafael & Shin, Naomi L. (eds.). *Subject pronoun expression in Spanish. A cross-dialectal perspective*, 17-58. Washington: Georgetown University Press.
- Pérez Saldanya, Manuel & Hualde, José Ignacio. 2003. On the origin and evolution of the Catalan periphrastic preterit. In Pusch, Claus D. & Wesch, Andreas (eds.). *Verbalperiphrasen in den (ibero)romanischen Sprachen*, 47-60. Hamburg: Helmut Buske Verlag.
- Posio, Pekka. 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(3): 777-798.
- Posio, Pekka. 2013. The expression of first-person-singular subjects in spoken Peninsular Spanish and European Portuguese: Semantic roles and formulaic sequences. *Folia Linguistica* 47(1): 253-291.  
<<https://doi.org/10.1515/flin.2013.010>>
- Posio, Pekka. 2014. Subject expression in grammaticalizing constructions: The case of *creo* and *acho* 'I think' in Spanish and Portuguese. *Journal of Pragmatics* 63: 5-18.

- Posio, Pekka. 2018. Properties of pronominal subjects. In Geeslin, Kimberly L. (ed.). *The Cambridge handbook of Spanish linguistics*, 286-306. Cambridge: Cambridge University Press.  
<<https://doi.org/10.1017/9781316779194.014>>
- Posio, Pekka & Pešková, Andrea 2020. *Le dije yo, digo* – Construccionalización de los introductores cuotativos con el verbo decir en español peninsular y argentino. *Spanish in Context* 17(3): 391-414.
- Rigau, Gemma. 1988. Strong pronouns. *Linguistic Inquiry* 19(3): 503-511.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris Publications.
- Rooryck, Johan. 2001. Evidentiality, Part I. *Glott International* 5(4): 125-133.
- Rosselló, Joana. 2000. A minimalist approach to the null subject parameter. *Catalan Working Papers in Linguistics* 8: 97-128.
- Sanchis Guarner, M. 1950. *Gramàtica valenciana*. València: Torre.
- Schmid, Beatrice. 2006. Els usos del verb fer i la seva traducció a l'alemany. *Zeitschrift für Katalanistik* 29: 115-134.
- Silva-Corvalán, Carmen. 2001. *Sociolingüística y pragmática del español*. Washington, DC: Georgetown University Press.
- Solà, Jaume. 1992. *Agreement and Subjects* (Doctoral thesis), Universitat Autònoma de Barcelona. <<https://ddd.uab.cat/pub/tesis/1992/tdx-0719110-134357/TJSP1de3.pdf>>
- Speas, Peggy & Tenny, Carol. 2003. Configurational properties of point of view roles. In Di Sciullo, Anna Maria (ed.). *Asymmetry in grammar*, Vol. 1, 315-344. Amsterdam: John Benjamins.  
<<https://doi.org/10.1075/la.57.15spe>>
- Travis, Catherine E. & Torres Cacoullous, Rena. 2012. What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics* 23(4): 711-748.  
<<https://doi.org/10.1515/cog-2012-0022>>
- Uriagereka, Juan. 1995. An F position in Western Romance. In Kiss, Katalin É. (ed.). *Discourse configurational languages*, 153-175. Oxford: Oxford University Press.
- Veny, Joan. 1984. *Els parlars catalans*. Palma: Editorial Moll.
- Willett, Thomas. 1988. A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language* 12(1): 51-97.  
<<https://doi.org/10.1075/sl.12.1.04wil>>

### Corpus

Parlars = *Parlars. Corpus oral del valencià col·loquial*. València: Universitat de València. <<http://www.uv.es/corvale>>

### Software/Tools

- Anthony, Laurence (2018). AntConc (Version 3.5.2) [Computer Software]. Tokyo, Japan: Waseda University. <<https://www.laurenceanthony.net/software>>
- ELAN (Version 6.2) [Computer software]. (2021). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. <<https://archive.mpi.nl/tla/elan>>

Lüdecke Daniel (2021). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.10. <<https://CRAN.R-project.org/package=sjPlot>>

R Core Team- 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. <<https://www.R-project.org/>>