Leonetti Escandell, Victoria & Torregrossa, Jacopo. 2024. The interpretation of null and overt subject pronouns in Spanish compared to Greek and Italian. *Glossa: a journal of general linguistics* 9(1). pp. 1–25. DOI: https://doi.org/10.16995/glossa.9725

Open Library of Humanities

The interpretation of null and overt subject pronouns in Spanish compared to Greek and Italian

Victoria Leonetti Escandell, Complutense University, Madrid, ES, vleonett@ucm.es

Jacopo Torregrossa, Goethe University, Frankfurt, DE, torregrossa@lingua.uni-frankfurt.de

In recent years, several studies have suggested that Null Subject Languages (NSLs) do not exhibit uniform behaviour in all respects. This paper aims to determine whether the interpretation of null and overt subject pronouns in Spanish differs from that in two other NSLs: Italian and Greek. To this end, we replicated the sentence interpretation experiment conducted by Torregrossa et al. (2020) on Greek and Italian to obtain a fully comparable set of data.

The results indicate that Spanish differs significantly from both Italian and Greek in the interpretation of null and overt pronouns. Spanish participants accept null subjects referring to both subject and object antecedents to the same extent. The same pattern is observed for overt subject pronouns. To explain the observed differences, we propose that two properties of its grammar influence the interpretation of null subjects and overt pronouns in Spanish: i) a flexible word order, allowing for SVO, VSO, and VOS; and ii) the existence of differential object marking (DOM). We hypothesise that these two factors may affect antecedent selection, thereby partially accounting for the differences observed among the languages under consideration.

1 Introduction

Most research on null subject languages (NSLs) concentrates on the syntactic properties that these languages share, such as the obligatory availability of null expletives, the presence of free subject inversion and the lack of *that*-trace effects (Rizzi 2013; D'Alessandro 2015, for a review). However, more recent studies have examined the extent to which NSLs differ from each other, with a particular focus on the subtle differences in the meaning and use of null subjects (NSs) and overt subject pronouns (OSPs) (e.g., Filiaci et al. 2013 on Italian and Spanish; Torregrossa et al. 2015 on Greek and Italian), to understand what the grammatical basis for these differences is.

In a recent paper, Torregrossa et al. (2020) conducted a sentence interpretation experiment and demonstrated that Greek and Italian differ in the selection of an antecedent for NSs in a local context. The authors attribute this difference to the varying word order possibilities in these languages. They propose that the availability of different word orders in a language influences the strength of the preverbal position as a cue for identifying the intended antecedent.

The present paper aims to contribute to the understanding of differences among NSLs by investigating an additional language, Spanish. Our main research question is how Spanish patterns in the selection of antecedents for NSs (and OSPs), as compared to Greek and Italian. Based on a systematic comparison of the behaviour of NSs across the three languages, we will be able to analyse whether the results align with the proposal by Torregrossa et al (2020). More specifically, since the differences between Greek and Italian correlate with other syntactic differences—particularly, word order possibilities—we hypothesise that the behaviour of Spanish should also correlate with syntactic differences related to the properties of subjects and objects. Our results can offer a solid basis for examining the validity of the original conclusions.

To ensure that our Spanish data are comparable, we replicated the experiment originally conducted by Torregrossa et al. (2020) in Greek and Italian. By closely following their experimental design and methodology, we aimed to achieve consistency in the data collection process. This involved using the same types of stimuli, maintaining similar testing conditions, and applying identical analytical criteria. This allows a direct comparison of the interpretation patterns of NSs across the three languages, ensuring that any observed differences or similarities are due to linguistic factors rather than variations in experimental procedure.

The choice of Spanish is motivated by two main considerations:

i. Spanish allows different word orders in broad-focus sentences, like Greek and unlike Italian.
ii. Spanish, like Italian and unlike Greek, does not have morphological case-marking, although it has Differential Object Marking (DOM), where certain objects receive a specific marker based on features such as definiteness, animacy, specificity, and topicality.

The paper is structured as follows. In Section 2, we summarise the proposal by Torregrossa et al. (2020) on the relationship between the interpretation of NSs and word-order possibilities in Greek and Italian, as well as the grammatical properties of Spanish that may affect the interpretation of NSs, namely the availability of VSO. In Section 3, we present our study on the interpretation of NSs (and OSPs) by adult native speakers of Spanish. Since this study uses the same experimental materials as Torregrossa et al. (2020), we can systematically compare the behaviour of NSs across the three languages. Finally, in Section 4, we discuss the results and the differences in NSs and OSPs between Spanish and Greek and Italian. Section 5 presents our conclusions and thoughts for future research.

2 Interpreting NSs. Word order and the hierarchical height principle

2.1 A view on Greek and Italian (Torregrossa et al. 2020)

Research on the interpretation of NSs across various languages has shown that they tend to refer to the subject of a preceding clause in both intra- and inter-sentential contexts. This pattern has been attested for languages such as Italian (Carminati 2002; Bader et al. 2023), Spanish (Alonso Ovalle et al. 2002), Catalan (Mayol & Clark 2010), and Greek (Papadopoulou et al. 2015).

Carminati (2002) argues that this tendency results from the relatively high syntactic position of the subject constituent, which in Italian occupies the Specifier of the Inflectional Phrase (SpecIP). Thus, as shown in Carminati (2002: 45), Italian speakers tend to interpret the NS in the second clause of (1) as co-referential with the subject in SpecIP, *Marta*. Conversely, an OSP such as *lei* ('she'), typically refers to a constituent in a syntactically lower position than SpecIP, such as the object *Piera* in (1b).

(1) a. Marta scriveva frequentemente a Piera, quando (pro) era
Marta write.IND.IPFV.3SG frequently to Piera, when pro be.IND.IPFV.3SG
negli Stati Uniti.
in the United States.
'Marta frequently wrote to Piera, when she (=Marta) was in the United States.'

b. Marta scriveva frequentemente a Piera, quando lei era
Marta write. IND.IPFV.3SG frequently to Piera, when she be.IND.IPFV.3SG
negli Stati Uniti.
in the United States.
'Marta frequently wrote to Piera, when she (=Piera) was in the United States.'

However, this generalisation is not equally robust across all NSLs. Different studies provide different results in terms of frequency (Alonso Ovalle et al. 2002; Carminati 2002; Papadopoulou et al. 2015).[1] For example, when compared to Italian NSs, Greek NSs seem to be less sensitive to

[1] The fact that these studies employ various methodologies (e.g., questionnaire studies versus picture-matching tasks) and examine different types of items (e.g., intra- versus intersentential contexts) makes it difficult to compare the data and draw general conclusions.

the function (i.e., reference to a subject vs. object antecedent). The contrast between Greek and Italian is shown in (2) and (3) – taken from Torregrossa et al. (2020). The two sentences have been constructed to force that reference of the NS to *Mario* is the only possible interpretation in both languages, due to gender-matching between the object constituent and the masculine singular adjective *imbarazzato/ntropiasménos* 'embarrassed'. However, while Greek speakers accept this use quite naturally, Italian speakers may have difficulty deriving the same interpretation, due to the strong subject bias associated with Italian NSs. [2]

(2) *Greek*
Epeidí i Caterina epékrine afstirá to Mario, (pro)
Because C.NOM.FEM.SG criticize.IND.PFV.3SG severely M.ACC.MASC.SG pro
ítan ntropiasménos.
be.IND.IPFV.3SG embarrassed.NOM.MASC.SG
'Because Caterina has severely criticized Mario, (he) was rather embarrassed.'

(3) *Italian*
Poiché Caterina ha severamente criticato Mario, (pro)
because C.FEM.SG AUX.IND.PRS.3SG severely criticize.PST.PTCP M.MASC.SG pro
era imbarazzato.
be.IND.IPFV.3SG embarrassed.MASC.SG
'Because Caterina has severely criticized Mario, (he) was rather embarrassed.'

To explain the contrast between (2) and (3), Torregrossa et al. (2020) present a two-part argument. First, they propose that the retrieval of the antecedent for NSs is influenced by the hierarchical structure of the preceding sentence. Specifically, NSs are more likely to refer to constituents that are hierarchically higher within the structure. This explains the tendency for the Italian NS in (1a) to refer to the subject *Caterina*, which occupies a higher structural position than the constituent *a Piera* in the previous sentence. Second, they argue that the proximity of the two antecedents in linear terms influences the likelihood that the hierarchically higher constituent will be selected by the NS. In other words, native speakers find the interpretation of the NS in

[2] It should be noted that the sentences in (2) and (3) differ from (1) because the subordinate clause precedes the matrix clause. It has been shown that the linear order between matrix and subordinate clause may affect pronoun interpretation (see, e.g., Fedele & Kaiser 2014 for Italian). Crucially, Italian and Greek native speakers have the same intuitions related to the facility of linking the NS to the object antecedent if the two clauses in (2) and (3) are two independent clauses, as in (i) or (ii):

(i) Caterina ha severamente criticato Mario. Era imbarazzato.
(ii) I Caterina epékrine afstirá to Mario. Ítan ntropiasménos.
'Caterina criticized Mario severely. (He) was embarrassed.'

(1a) relatively straightforward because the linear distance between the two constituents, *Caterina* and *a Piera*, is significant, with a verb and an adverb intervening between them.

However, when the linear distance between two potential antecedents of an NS is reduced, the preference for the hierarchically superior constituent (e.g., the one in SpecIP) tends to diminish.[3] This is illustrated in (4), where the linear distance between the subject and object constituents is reduced, as no intervening constituent separates them. Notably, the NS in (4) no longer exhibits a strong interpretative bias; it can easily refer to either the clitic-left-dislocated indirect object in the sentential left periphery, *a Piera* ('to Piera'), or to the subject constituent in Spec,IP, *Marta* (see also Rizzi 2018; Torregrossa et al. 2020 for similar examples).[4]

(4) A Piera, Marta le scriveva frequentemente, quando (pro) era negli Stati Uniti.
To Piera Marta CL.DAT.FEM.SG write.IND.IPFV.3SG frequently when pro be.IND.IPFV.3SG in the States United
'To Piera Marta frequently wrote (to her), when (she) was in the United States.'

Torregrossa et al (2020) suggest the principle of hierarchical height to account for this behaviour.

(5) A DP is more prominent than another DP if the former is hierarchically higher than the latter.

Therefore, in the case of NSs, the greater the difference in hierarchical height between constituents, the stronger the bias toward the higher constituent. Specifically, as the distance between two constituents increases, speakers are more likely to interpret NSs as referring to the hierarchically higher constituent.

The study by Torregrossa et al. (2020) revealed a significant difference between Italian and Greek. Their results show that in Greek, the SpecIP position is not as strong a cue for identifying the antecedent of an NS as it is in Italian. The difference seems to be related to word order possibilities, among other factors.[5] In broad-focus contexts, Greek allows both SVO and VSO orders, whereas Italian only allows SVO. VSO in Italian is possible only with a narrow or contrastive focus interpretation for the subject (Belletti 2004) – a possibility also for Greek, as

[3] It is important to note that the difference in linear distance should not be considered a grammatical principle but rather a principle of parsing.

[4] However, it must be noted that this intuition is not as clear as it is with SVO-sentences such as the one in (1) (see also Rizzi 2018 for similar judgments). This may be because the different positions of the subject and the object are marked more prominently in (1), with the verb intervening between the two constituents, than in (6), where no constituent intervenes between them.

[5] Hierarchical height may then interact with other principles (e.g., topicality, coherence relations, etc.), which makes the interpretation of referring expressions in discourse a multifactorial process (Torregrossa et al. 2019 for discussion).

shown in (6) and (7). In the examples, the narrow or contrastive focus interpretation is signalled by uppercase.

(6) *Greek*

Plirose	{o giatros/O GIATROS}	ton	architektona.
pay.IND.PFV.3SG	the doctor.NOM.MASC.SG	the	architect.ACC.MASC.SG

(7) *Italian*

Ha	pagato	{?? il dottore/IL DOTTORE}	l'architetto.
AUX.IND.PRES.3SG	pay.PPT	the doctor.MASC.SG	the architect.MASC.SG

'The doctor paid the architect.'

According to Roussou & Tsimpli (2006), the subject and the object can occupy the same clausal domain in Greek but not in Italian. This is due to differences in case-marking features. In Italian, nominal phrases lexicalise the same set of features since they do not receive overt case marking. As a result, the subject and the object must appear in different clausal domains: while the object usually remains in the VP-domain, the subject must occupy a higher projection (i.e., SpecIP). In Greek, by contrast, subjects are overtly case-marked with nominative (*o giatros* 'the doctor'), whereas objects receive accusative marking (*ton architektona* 'the architect'). The fact that they lexicalise different features makes it possible for them to co-occur in the same domain.

Building on these assumptions, Torregrossa et al. (2020) argue that in Greek, the linear distance between subject and object constituents—used as a cue for interpreting NSs—is not as strong as it is in Italian. This is because, unlike in Italian, subjects and objects in Greek do not necessarily occupy different clausal domains. Consequently, the difference in linear distance between subjects and objects in broad-focus sentences is not as pronounced in Greek as it is in Italian.

As a result, Greek NSs exhibit a lower tendency to refer to antecedents in the preverbal position and, as a result, a greater tendency to refer to antecedents in the postverbal position compared to Italian NSs, as demonstrated by Torregrossa et al. (2020). This finding suggests that the reliability of a constituent's hierarchical height as a cue for interpreting NSs varies across languages, depending on how the hierarchical height principle interacts with other grammatical properties specific to each language.

2.2 Extending the view to Spanish

Several studies have observed that Spanish allows the VSO order in broad-focus sentences, distinguishing it from Italian and other Romance languages such as Catalan and French. This distinction is illustrated by comparing the Italian sentence (8) with the Spanish sentence (9), both cited from Leonetti (2014: 37):

(8) *Ha comprato Maria il giornale
AUX.IND.PRES.3SG buy.PPT Maria the newspaper

(9) Ha comprado María el periódico.
AUX.IND.PRES.3SG buy.PPT Maria the newspaper
'Maria bought the newspaper.'

As mentioned before, the VSO order is possible in Italian only if the subject is a narrow or contrastive focus. Additionally, it is worth noting that some Spanish speakers are more likely to accept sentences like (9) if they are preceded by a constituent in the left periphery, such as an adverb, as demonstrated in (10) from Leonetti (2014: 48) and supported by Zubizarreta (1998) and Roussou & Tsimpli (2006).

(10) Ayer ganó Juan la lotería.
Yesterday win.IND.PFV.3SG Juan the lottery
'Yesterday Juan won the lottery.'

This observation suggests that while the VSO order is generally acceptable in Spanish, its naturalness may be influenced by the discourse context or the presence of additional elements that help frame the sentence. Consequently, Greek and Spanish are expected to exhibit similar patterns in selecting antecedents for NSs. These patterns are different from those found in Italian.

3 The study

We aim to understand to what extent NSs are interpreted in Spanish in a different way than in Greek and Italian. For the present study, we created a Spanish version of the pronoun interpretation task used in Torregrossa et al. (2020). Seventy-two Spanish native speakers took part in the study. The presentation of the results will be based on the comparison between the data presented in Torregrossa et al. (2020) and the Spanish data collected for this study. To compare the interpretation of NSs across the three languages, we will use a linear mixed effects model, which improves on the ANOVA analysis used in Torregrossa et al. (2020) by modelling random effects.

Crucially, Spanish, like Greek, allows for VSO order, which could affect the interpretation patterns of NSs and OSPs. Therefore, we aim to verify the hypothesis that the preverbal position is a less reliable cue for interpreting NSs (or OSPs) in languages allowing for alternation between SVO and VSO in broad-focus contexts. Consequently, we will use exclusively SVO sentences in broad-focus contexts and not examine how changes in word order between the subject and the object affect the interpretation of NSs or overt subject pronouns.

3.1 Participants

Seventy-two adult native speakers took part in the study (Mean age: 26;8, Range: 18–78). Before conducting the study, each participant had to complete a short questionnaire, related to background information (age, level of education, region of origin), knowledge of a second or third language/dialect and level of proficiency in this language: 78.1% of participants were speakers of the Central variety (Madrid, Castilla-La Mancha, Castilla y León), while 12.2% were speakers of the Levante variety (Cataluña, Comunidad Valenciana, Islas Baleares); 9.7% were speakers from other areas. All participants who declared to be bilingual or to have a self-reported level of proficiency in a second language higher than C1 of the European Framework for Language Competence were excluded from the analysis.

The study by Torregrossa et al. (2020) included 62 adult native speakers (mean age: 30;7, range: 22–54) from both Northern Italy (Lombardy and Emilia Romagna, 51%) and Central Italy (Marche, Tuscany, Umbria, 49%), with 50% of them being speakers of both Italian and a local dialect as well as 62 adult native speakers of Greek (mean age: 21;6, range: 18–48), mainly born and raised in Northern Greece (e.g., areas around Thessaloniki), but also in Central (19.35%) and Southern Greece (17.75%), with 11.3% of them being speakers of a local dialect.

3.2 Materials and procedure

We administered an interpretation task in which participants had to indicate to which extent they interpreted a NS vs. an OSP as referring to a subject vs. object antecedent, based on a 5-point Likert-scale. Two factors were manipulated: *Type of Pronoun* (NS vs. PRON) and *Pronoun Interpretation* (reference to the subject vs. reference to the object), resulting in four different conditions.

Each sentence consisted of an SVO-main clause introducing two same-gender referents, one in subject and one in object position, followed by a subordinate clause containing either a NS or an overt pronoun in subject position, whose reference was potentially ambiguous between the subject and the object constituent in the main clause (see (11)). The stimuli were adapted into Spanish from the stimuli in Greek and Italian used in Torregrossa et al. (2020) – see (12) and (13). Crucially, the sentences were kept as ambiguous as possible, to prevent the action described in the subordinate clause from being prototypically associated with one of the two referents. This way, the effect of world knowledge on reference assignment was controlled for, or kept to a minimum at least. Furthermore, we kept the type of subordinate clause and the tense and aspect of the verb constant across stimuli: the subordinate clause was always a temporal one, corresponding to a *while*-clause in English. The verb was always the *imperfecto* expressing imperfective, unbounded aspect in the past. This was done for two reasons. First, we wanted to ensure comparability with the stimuli used in Torregrossa et al. (2020), since the imperfective

aspect was used in the Greek and Italian stimuli, too. Second, varying the type of subordinate clause and the aspectual properties of the verb may lead to variation in the interpretation of the pronoun occurring in the subordinate clause. For example, based on sentences similar to (13), Martín-Villena (2023) noticed that the use of *cuando* ('when') instead of *mientras* ('while') is associated with a less pronounced tendency of NSs to be interpreted as referring to a subject antecedent. Likewise, using the conjunction *cuando* and substituting the imperfective aspect in the subordinate clause with the perfective one may lead to a different interpretation of the event structure, whereby the events described in the main and subordinate clause are interpreted as sequential (rather than simultaneous) to each other (Keller-Cohen 1981). In fact, pronoun interpretation depends on a comprehender's representation of the event structure (Kehler & Rohde 2019).

(11) *Spanish*
   a. El doctor pagó al arquitecto mientras Ø cerraba la cartera.
   b. El doctor pagó al arquitecto mientras *él* cerraba la cartera.

(12) *Italian*
   a. Il dottore pagò l'architetto mentre Ø chiudeva la cartella.
   b. Il dottore pagò l'architetto mentre *lui* chiudeva la cartella.

(13) *Greek*
   a. O giatros plirose ton architektona, eno Ø ekleine to fakelo.
   b. O giatros plirose ton architektona, eno *aftos* ekleine to fakelo.
      'The doctor paid the architect, while pro/he was closing the folder.'

To manipulate *Pronoun Interpretation* (reference to the subject vs. object), each sentence was followed by a question asking how likely it was that the action expressed by the verb in the subordinate clause was being performed by the subject or the object of the main clause, as in (14):

(14) a. ¿Cómo de probable es que fuera el doctor/el arquitecto el que cerraba la cartera? *(Sp)*
   b. Quanto probabile pensi che fosse il dottore/l'architetto a chiudere la cartella? *(It)*
   c. Poso pithano einai oti o giatros/o architektonas ekleine to fakelo; *(Gr)*
      'How likely do you think it was that the doctor/the architect closed the folder?'

Participants had to answer this question based on a 5-point Likert scale, where 1 indicated that it was not likely at all (that the subject or object antecedent performed the action expressed in the subordinate clause), while 5 indicated that it was very likely. The experimental material included 30 items appearing in four versions that differed with respect to *Type of Pronoun* (null vs. overt) and target question (i.e., *Pronoun Interpretation*, referring to the subject or the object

antecedent). Following a Latin Square Design, we created four lists, each containing 30 critical sentences and 15 fillers.

The fillers were created by manipulating the type of referring expression (a full DP instead of a NS or an overt pronoun in the subordinate clause), the gender of PRON in the subordinate clause together with the gender of one of the two antecedents (e.g., a female character used as an object, followed by a gender-matched overt pronoun, see (15)), the type of subordinate clause (e.g., a temporal clause indicating precedence in time instead of simultaneity) and the linear order between the main and the subordinate clause (Torregrossa et al. 2020 for a similar methodology).

(15) Antes de que el jefe de estación parase a la ladrona, ella tiró la cartera que había robado. ¿Cómo de probable es que fuera el jefe de estación el que tirase la cartera?
'Before the station master stopped the (female) thief, she threw the wallet that she had stolen. How likely do you think it was that the female thief threw the wallet?'

Other sentences including questions related to same-gender characters were also included as fillers.[6]

(16) Ana dice: "Silvia está muy joven y muy guapa."
¿Cómo de probable es que Ana haya visto a Silvia?
'Ann says: "Silvia looks very young and pretty."
How likely is it for Ann to have seen Silvia?'

Each participant was assigned only one experimental list: 18 participants completed List 1, 18 List 2, 19 List 3 and 17 List 4. Before conducting the experiment, participants had to read the following instructions in Spanish (the same as in the Italian and Greek experiments from Torregrossa et al. 2020): "The following questionnaire consists in reading and interpreting some sentences. Specifically, you will read a sentence that mentions two characters, A and B (for example, "A greets B [...]") and a sentence that acts as a continuation and describes an action performed by only one of the two characters (for example, "[...] while A/B runs"). After reading the two sentences, you will be asked how likely you think it is for character A or B to perform the action described in the continuation. Remember that 1 indicates a low likelihood and 5 indicates a high likelihood."

3.3 Results

3.3.1 Null subjects in Italian, Greek and Spanish

Figure 1 reports the means and standard errors (+/– 1.5) for the Likert-scores associated with null subjects referring to a subject antecedent as opposed to an object antecedent across the three

[6] These sentences correspond to a study on copular sentences (Escandell-Vidal 2023).

languages, Italian, Greek and Spanish. Italian speakers accept the interpretation of NSs referring to a subject antecedent to the greatest extent, followed by Greek speakers and Spanish speakers, respectively (Mean in Italian: 3.88, SD: 1.16; Mean in Greek: 3.55; SD: 1.36; Mean in Spanish: 3.30, SD: 1.53). By contrast, Italian accepts the interpretation of NSs as referring to an object antecedent to the lowest extent, followed by Greek speakers and Spanish speakers, respectively (Mean in Italian: 2.38, SD: 1.23; Mean in Greek: 2.64; SD: 1.39; Mean in Spanish: 3.23, SD: 1.46).

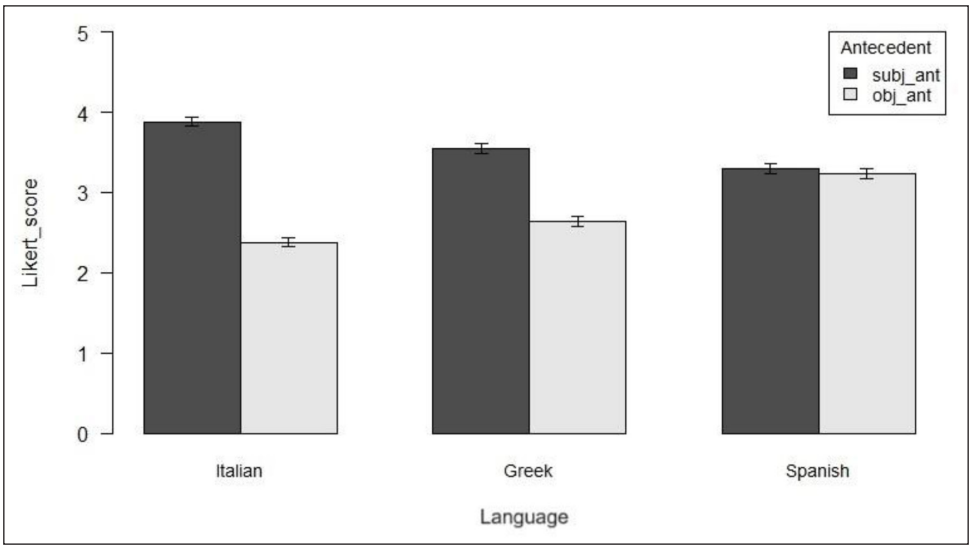

Figure 1: Means (bar plots) and standard errors (+/– 1.5) for Likert-scores associated with null subjects as referring to subject or object antecedents across the three languages, Italian, Greek and Spanish.

We used R (R Core Team 2012) and lme4 (Bates & Maechler & Bolker 2012) to conduct a linear mixed-effects analysis, considering all Likert scores associated with NSs as dependent variable.[7] As fixed effects, we used *Pronoun Interpretation* (reference to the subject vs. reference to the object) and *Language* (Greek, Italian and Spanish) as well as their interaction. We specified random slopes for the effect of *Pronoun Interpretation* by participant (ID).[8] We chose the Likert-scores associated with reference to the subject in Italian as the reference level, because Italian is the language that most clearly uses NSs to refer to subject antecedents (see Section 2). The LME-analysis reveals a significant lower-order effect of *Pronoun Interpretation*, indicating that in

[7] Likert-scale scores correspond to an ordinal variable, and not a continuous one. Therefore, there is some debate as to whether it is advisable to run linear mixed-effects models with a variable of this type. Here we follow Norman (2010), who recommends using linear mixed-effects models with Likert-scale scores.

[8] The resulting model was:

m0 <- lmer (Likert_scores ~ 1 + pronoun interpretation * language + (1 + pronoun interpretation|ID), data = null, control = lmerControl(calc.derivs = FALSE)).

Italian, participants tend to assign lower scores to NSs referring back to object antecedents as opposed to subject antecedents, as reflected by the negative estimate (Table 1). We also found a significant lower-order effect of *Language*, both with Greek and Spanish, showing that participants tend to assign lower scores to NSs referring to subject antecedents in both languages, as revealed by the fact that in both cases, the estimate is negative. Finally, we found a significant interaction between *Pronoun Interpretation* and *Language*, which holds for both languages. This suggests that the difference between the Likert-scores related to NSs referring back to a subject antecedent and NSs referring back to an object antecedent is less marked in both Greek and Spanish as opposed to Italian. This interpretation is based on the observation of a positive estimate.[9]

Fixed effects	***Estimate***	***SE***	***t***	***p***
Intercept	3.88	0.10	39.15	< .001
Pronoun Interpretation (object)	–1.50	0.17	–8.80	< .001
Language (Greek)	–0.33	0.14	–2.33	.02
Language (Spanish)	–0.59	0.13	–4.36	< .001
Pronoun Interpretation (object) Language (Greek)	0.58	0.24	2.43	.02
Pronoun Interpretation (object) Language (Spanish)	1.42	0.23	6.14	< .001

Table 1: Parameters of the linear mixed effects analysis concerning the Likert-scores associated with null subjects as referring to a subject antecedent or an object antecedent across languages (Italian, Greek and Spanish). The fixed effects, their estimates, standard errors (SE), *t*-scores and *p*-values are given.

We used *emmeans* in R (Lenth 2020) to compute pairwise comparisons. The analysis shows that whereas in Greek, the Likert scores associated with NSs referring to a subject antecedent differ from the ones associated with an object antecedent ($\beta = .91$, $t = 5.36$, $p < .001$), this is not the case in Spanish ($\beta = .08$, $t = .50$, $p = .99$). Furthermore, the Likert scores associated

[9] We checked for normality and homoskedasticity of the model's residuals, mainly because the Spanish participants had a higher age range than the Italian and Greek participants. For normality, we relied on the inspection of a Q-Q plot. This did not indicate any violation of the normality assumption: the data followed a straight line. For homoskedasticity, we conducted a Levene's test. The test was significant, meaning that the assumption of equal variance was not met. A visual inspection of the boxplot of the residuals indicated that whereas Greek and Spanish showed a similar distribution, this was not the case for Italian. Therefore, the age range of Spanish speakers did not seem to affect the variance of their responses in comparison to Italian and Greek speakers. By contrast, the Italian participants exhibited a greater amount of variance in their responses, probably because of dialectal variation (participants from both the North and the centre of Italy participated in the study).

with NSs referring to an object antecedent in Spanish differ significantly from the ones in Italian ($\beta = -.83$, $t = -6.40$, $p < .001$) and Greek ($\beta = -.58$, $t = -4.42$, $p < .001$). By contrast, the *emmeans* analysis reveals no difference between Greek and Italian in the Likert scores associated with NSs referring to an object antecedent ($\beta = -.26$, $t = -1.90$, $p = .41$).

3.3.2 Overt subject pronouns in Italian, Greek and Spanish

Figure 2 reports the means and standard errors (+/– 1.5) for the Likert-scores associated with OSPs referring to a subject antecedent as opposed to an object antecedent across the three languages, Italian, Greek and Spanish. Italian speakers accept the interpretation of OSPs as referring to a subject antecedent to the lowest extent, followed by Greek speakers and Spanish speakers, respectively (Mean in Italian: 1.98; SD: 1.18; Mean in Greek: 2.23; SD: 1.27; Mean in Spanish: 3.02, SD: 1.55). By contrast, Italian accepts the interpretation of an overt pronoun as referring to an object antecedent to the greatest extent, followed by Greek speakers and Spanish speakers, respectively (Mean in Italian: 4.09, SD: 1.15; Mean in Greek: 3.68; SD: 1.32; Mean in Spanish: 3.41, SD: 1.51).

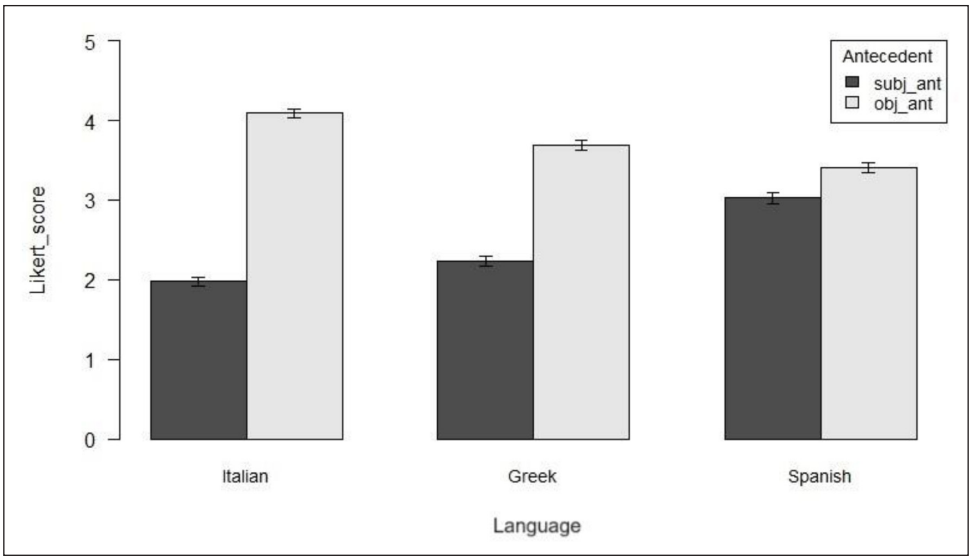

Figure 2: Means (bar plots) and standard errors (+/–1.5) for Likert-scores associated with overt subject pronouns as referring to subject or object antecedents across the three languages, Italian, Greek and Spanish.

We performed an additional linear mixed effects analysis, considering the Likert scores associated with OSPs as dependent variable. As fixed effects, we used *Pronoun Interpretation* (reference to the subject vs. reference to the object) and *Language* (Greek, Italian and Spanish) as well as their interaction. We specified random slopes for the effect of *Pronoun Interpretation*

by participant (ID). In line with what has been done in the previous analysis (Section 4.1), we chose the Likert-scores associated with reference to the subject in Italian as the reference level. The LME-analysis reveals a significant lower-order effect of *Pronoun Interpretation*, indicating that in Italian, participants tend to assign higher scores to OSPs referring to an object antecedent as opposed to a subject antecedent, as reflected by the positive estimate (Table 2). We also found a significant lower-order effect of *Language*, only in association with Spanish, showing that participants tend to assign higher scores to OSPs referring to subject antecedents than in Italian, as revealed by the fact that the estimate is positive. Finally, we found a significant interaction between *Pronoun Interpretation* and *Language*, which holds for both languages. This suggests that the difference between the Likert-scores associated with OSPs referring to a subject antecedent and OSPs referring to an object antecedent is less marked in both Greek and Spanish as opposed to Italian. This interpretation is based on the observation that the estimate is negative.[10]

Fixed effects	***Estimate***	***SE***	***t***	***p***
Intercept	1.98	0.10	19.54	<.001
Pronoun interpretation (object)	2.12	0.19	11.18	<.001
Language (Greek)	0.25	0.14	1.76	.08
Language (Spanish)	1.04	0.14	7.53	<.001
Pronoun Interpretation (object) Language (Greek)	–0.67	0.27	–2.49	.01
Pronoun Interpretation (object) Language (Spanish)	–1.75	0.26	–6.77	<.001

Table 2: Parameters of the linear mixed effects analysis concerning the Likert-scores associated with overt subject pronouns as referring back to a subject antecedent or an object antecedent across languages (Italian, Greek and Spanish). The fixed effects, their estimates, standard errors (SE), *t*-scores and *p*-values are given.

As in the case of NSs, we used *emmeans* in R (Lenth 2020) to compute pairwise comparisons. The analysis shows that in Greek, the Likert scores associated with NSs referring to a subject antecedent differ from the ones associated with an object antecedent ($\beta = -1.45$, $t = -7.65$, $p < .001$). However, this is not the case in Spanish ($\beta = -.37$, $t = -2.13$, $p = .28$). Furthermore, the

[10] The analysis of normality and homoskedasticity of the model's residuals showed a similar pattern as the one observed in footnote 6 for the model related to null subjects. The inspection of the Q-Q plot did not indicate any violation of the normality assumption. The Levene's test was significant and the observation of the boxplot of the residuals showed that the Italian participants exhibited a greater amount of variance in their responses than the Greek and Spanish participants, which may again be related to dialectal variation.

Likert scores associated with OSPs referring to a subject antecedent in Spanish differ significantly from the ones in Italian ($\beta = -1.04$, $t = -7.53$, $p < .001$) and Greek ($\beta = -.79$, $t = -5.70$, $p < .001$). By contrast, the *emmeans* analysis reveals no difference between Greek and Italian in the Likert scores associated with OSPs referring to a subject antecedent ($\beta = -.25$, $t = -1.76$, $p = .50$).

4 Discussion

This study aimed to collect a new set of data on the interpretation of NSs and OSPs in Spanish to determine how the antecedents are chosen. Our results reveal that Spanish exhibits significant differences from both Italian and Greek in the interpretation of null and overt pronouns. These differences point to the importance of considering cross-linguistic variation when studying pronoun interpretation, as the principles governing pronoun use can vary considerably even among closely related languages.

This finding contrasts with the general assumption that the morphosyntactic form of a referring expression encodes the activation/accessibility of its antecedent (see Ariel 1990 as a main reference). Based on this theory, referring expressions have a different function (e.g., reference to an antecedent in preverbal vs. postverbal position) depending on their morphosyntactic form. For example, referring expressions which are not realized phonologically (like NSs) may refer to antecedents in preverbal position (i.e., subjects in our experiment), while phonologically realized referring expressions (like OSPs) refer to antecedents in postverbal position (i.e., objects in our experiment). Crucially, this should apply to all languages considered in this study. Since all of them exhibit the distinction between NSs and OSPs, these two forms should be assigned two different discourse functions (i.e., reference to a constituent in preverbal vs. postverbal position, respectively). However, the results of this study reveal that the extent to which NSs and OSPs refer to a subject or an object antecedent varies across Italian, Greek and Spanish.

Italian features the clearest division of labour between NSs and OSPs. Among all participants, Italian speakers are most likely to accept a NS as referring to an antecedent in preverbal position and an OSP as referring to an antecedent in postverbal position. Incidentally, it should be noted that the mean Likert score associated with reference of an OSP to an antecedent in postverbal position (object) is the highest one among all conditions and language groups (see Section 3.3.2). Following Carminati (2002) and Torregrossa et al. (2020), we argued that in Italian, the preverbal/postverbal position of an antecedent is a very reliable cue for the interpretation of NSs or OSPs. This was confirmed by the Italian data observed in this study: NSs were most likely to refer to an antecedent in preverbal position, which corresponds to the subject constituent in our study, whereas OSPs tended to refer to an antecedent in postverbal position, corresponding to an object (see Section 2).

Greek also features a clear division of labour between NSs and OSPs, just like Italian. Greek speakers tend to accept a NS referring to an antecedent in preverbal position (the subject) to a greater extent than a NS referring to an antecedent in postverbal position (the object), and an OSP referring to an object antecedent to a greater extent than an OSP referring to a subject antecedent. However, the significant interaction between *Pronoun Interpretation* and *Language* shown in Table 1 for NSs reveals that in Greek, the difference between the two discourse functions of NSs (i.e., reference to the subject vs. to the object) is less marked than in Italian. The results related to OSPs show the same tendency – with the difference between the two discourse functions being less marked in Greek than in Italian – even if the interaction is not significant. The same results related to the interpretation of null and OSPs in Greek emerged from the study by Torregrossa et al. (2020), although the data have been modelled differently in the present study, i.e., by considering the random effects structure (see Section 3.3). As discussed in Section 2, the preverbal position of an antecedent in Greek is not as relevant as a cue for the interpretation of NSs as it is in Italian. This has been traced back to the availability of VSO in broad focus contexts in Greek. In these types of sentences, the two possible antecedents (the subject and the object) appear in postverbal position.

The data concerning Spanish correspond to the original empirical contribution of this study, for which we have adapted the task designed for Italian and Greek in Torregrossa et al. (2020). Crucially, we found that Spanish is the language that mostly deviates from a clear division of labour between NSs and OSPs. Spanish speakers seem to accept NSs referring to a subject antecedent and an object antecedent to the same extent. The same holds true for OSPs. In fact, in a sentence like (17), both the null and the OSPs were interpreted as referring to the subject and the object antecedents to the same extent, so the gatekeeper and the mail carrier were both as likely to open the door, while in Greek and Italian the gatekeeper was mostly associated to the NS and the mail carrier was mostly associated to the OSP.

(17) El portero reconoció al cartero mientas Ø/él abría la puerta.
'The gatekeeper recognised the mail carrier while he was opening the door.'

The Likert-scores associated with the interpretation of NSs and OSPs in both conditions (reference to the subject vs. to the object) differ significantly from the respective conditions in Italian and Greek. These differences call for an explanation.

The study by Filiaci et al. (2013) was the first to investigate the extent to which the interpretation of NSs (and overt pronouns) in two NSLs, Italian and Spanish, conform to Carminati's (2002) generalisation. Using two self-paced reading experiments, the authors found differences in the processing costs associated with null and OSPs in these languages. A NS forced to refer to an object antecedent incurred additional processing costs for both Italian and Spanish native speakers. In contrast, processing OSPs referring to a subject antecedent was associated

with extra processing costs in Italian but not in Spanish. To account for this difference, the authors proposed that OSPs in Italian and Spanish are syntactically different. In Italian, *lui/lei* ('he/she') are strong pronouns and, therefore, must refer to a non-prominent antecedent (e.g., an object). In contrast, in Spanish, *él/ella* ('he/she') are weak forms and are less biased towards an object antecedent (see Cardinaletti & Starke 1999 for the distinction between weak and strong pronouns).[11] According to Ariel (1990), the form of referring expressions encodes the degree of accessibility (or prominence) of the associated referents: the less complex a pronoun is syntactically (e.g., clitics vs. weak pronouns or weak pronouns vs. strong pronouns), the greater its tendency to refer to a prominent antecedent.

Though appealing, the idea that the same pronominal form might have a different syntactic structure across languages raises a significant learnability issue. Cardinaletti & Starke (1999) demonstrate that weak and strong pronouns differ in syntactic behaviour. Consequently, children can rely on syntactic evidence to learn whether OSPs in their language are weak (as in Spanish, according to Filiaci et al. 2010) or strong (as in Italian). However, this consideration cannot be extended to NSs. Given that NSs are covert syntactic objects, it is unclear what kind of syntactic evidence children might use to learn whether NSs are clitics or weak pronouns in their language.

Crucially, the proposal outlined by Torregrossa et al. (2020) avoids this problem because it links the interpretation of NSs to other grammatical properties, such as word order. The preverbal position of the antecedent is not a relevant cue for the interpretation of NSs in Spanish, which is due to the availability of VSO in this language, as discussed above for Greek. This suggests that children could use observable syntactic structures, such as word order differences, to infer the interpretation patterns of NSs in their language.

The proposal by Torregrossa et al. (2020) does not account for the differences between Greek and Spanish directly; however, it can inspire a way to find an explanation, based on the idea that other observable syntactic differences related to the prominence of the object can contribute to reduce the subject bias of NSs.

Recall that, in Greek, overt case marking was the crucial factor that determines the flexibility in word order: since the subject and the object lexicalise different case features, they both can co-occur within the same domain (Roussou & Tsimpli 2006). In Spanish (like in Italian), nominal phrases are not marked for case, which means that subject and object constituents do not lexicalise different features and, hence, should not be able to appear in the same sentential domain. However, contrary to Italian, certain object constituents in Spanish exhibit morphological marking under certain conditions. The relevant contrast is given in (18) and (19) from Fábregas (2013).

[11] Incidentally, the idea that different pronominal forms have different syntactic structures across languages might also explain the observed differences between Greek and Italian in the use of NSs. For example, NSs in Italian may function as clitics (Frascarelli 2007), whereas NSs in Greek might be weak pronouns. If this were the case, the NSs in the two languages would represent different stages in a hypothetical null subject cycle (cf. Fuß & Wratil 2013).

(18) Encontré *al* superviviente.
find.IND.PFV.1SG DOM a survivor
'I found the survivor.'

(19) Encontré un problema.
find.IND.PFV.1SG a problema
'I found a problem.'

Spanish does not have case marking; instead, it has Differential Object Marking (DOM) (see Leonetti 2004; 2014; López 2012; Fábregas 2013, for a general overview). DOM is a linguistic phenomenon where the direct object of a verb is marked differently depending on various factors such as animacy, definiteness, specificity, or topicality. In this respect, Spanish differs also from Italian (Belletti 2004). Thus, in languages with DOM, only certain objects receive a specific marker (like a preposition or suffix), based on these semantic or pragmatic features. This marking conveys additional information about the object within the sentence.

The interpretation of DOM is one of the most debated issues in the grammar of Spanish and a detailed account of DOM is beyond the scope of this article. However, if one wants to establish a connection between the availability of VSO and the lexicalization patterns of subjects and objects in a language, the difference between case marking in Greek and differential object marking in Spanish should be taken into account.

Of course, DOM in Spanish is not equivalent to case marking in Greek. Case marking in Greek is obligatory; by contrast, DOM only takes place under certain conditions. In Spanish, DOM is required with animate and definite objects (see (17)), but not allowed with inanimate objects (see (19)).[12] With indefinite animate objects, *a*-marked entities are usually associated with a

[12] As one anonymous reviewer has pointed out, there are some cases (non-specific definite as in (ii), Quine definites, as in (iii)) where DOM is absent:

(i) He encontrado a la niña.
have.1SG found DOM DEF.FEM.SG girl
'I have found the girl.' (the girl we talked about and which is salient in the discourse)

(ii) He encontrado la niña.
have.1SG found DEF.FEM.SG girl
'I have found the type of girl.' (for example, we are directing a movie and need a certain type of girl to act in it).

(iii) Juan busca la mujer perfecta.
Juan seek.3SG DEF.FEM.SG woman perfect.F.SG
'Juan is looking for the perfect woman.'

These expressions, the reviewer writes, can be tracked back to a NS.

All these are cases of non-referential definites, which behave as predicates. The only structures in which the nominal expression can be referred to by a NS are specificational copular sentences (Mikkelsen 2005), where the subject is also intensional, not referential. Compare (iv) and (v):

(iv) He encontrado la niña. (pro) es María.
(v) ¿?He encontrado la niña. (pro) vino ayer.

specific interpretation. For example, speakers tend to use sentences like (20) to express that they are looking for a specific doctor, who is active in discourse and identifiable in the mind of a speaker and/or hearer (cf. Leonetti 2004 for some exceptions to this generalisation[13]). On the contrary, by means of (21), speakers do not usually refer to any specific doctor (Leonetti 2004).

(20) Busca *a* un médico.
look-for.IND.PRES.3SG DOM a doctor

(21) Busca un médico.
look-for.IND.PRES.3SG a doctor
'(S)he looks for a doctor.'

The idea is, therefore, that DOM may affect the interpretation of NSs and OSPs in Spanish (aside from its relationship with the availability of VSO) by enhancing the degree of discourse prominence of the object constituent. In the literature, there have been several attempts to analyse DOM in Spanish as a marker of topicality (Iemmolo 2010; Nikolaeva & Dalrymple 2011; von Heusinger 2018). In particular, Nikolaeva & Dalrymple (2011) propose that DOM serves to mark secondary topics, i.e., entities "[…] such that the utterance is constructed to be ABOUT the relationship between it and the primary topic" (Nikolaeva 2001: 26), with the primary topic corresponding to the subject constituent. In other words, by uttering a sentence like (18) above, the speaker intends to convey information about herself, the survivor and the relation between them (as expressed by the verb 'meet').

Furthermore, DOM marks the topic-worthiness of the corresponding constituent (Iemmolo 2010): features such as animacy, definiteness or specificity (see above) enhance the prominence of the associated constituent and make it a good candidate in the discourse space for being a sentence topic (see von Heusinger & Schumacher 2019 for discussion). This is consistent with the findings of von Heusinger & Chiriacescu (2013) about DOM in Romanian and its discourse-semantic effects. The authors used a discourse continuation task, in which participants were given short story fragments to read. The last sentence of these story fragments featured either a differentially marked or a non-marked object. Participants had to write a discourse continuation of at least five sentences. The results of this experiment showed that compared to non-marked objects, differentially marked objects exhibited a greater tendency to be mentioned

[13] For example, *a*-marking is allowed with non-specific indefinites referring to animate object constituents (Leonetti 2004: 82):

(i) Esta buscando a alguien. / No esta buscando *a* nadie.
AUX.IND.PRES.3SG look-for.GERUND DOM somebody NEG. AUX.IND.PRES.3SG look-for.GERUND DOM nobody
'(S)he is looking for someone.' / '(S)he is not looking for anyone.'

again in subsequent discourse and to be promoted as subject/sentence topic in one of the five continuation sentences. These results suggest that DOM enhances the prominence of the associated constituent, which is reflected in its discourse structuring function. Although the syntax of DOM in Romanian is not the same as that of DOM in Spanish (since DOM in Romanian can co-occur with clitic doubling),[14] both systems are similar in other semantic aspects (Tigău 2021), so it is not to be excluded that some of the observations by von Heusinger and Chiriacescu (2013) could also apply to Spanish data. A first piece of evidence in this respect comes from the results of two corpus studies in von Heusinger et al. (2024).[15] The authors show that DOM-marked objects are more likely to become topics of a following sentence and be mentioned more frequently in later discourse than non-marked objects. As in the case of Romanian considered above, these findings suggest that DOM is associated with a specific discourse function, wherein it enhances the prominence of the associated constituent. However, it should be noted that von Heusinger et al. (2024) did not consider whether NSs or OSPs were used to refer back to marked or non-marked constituents. Therefore, their results are not directly relevant for our investigation, but provide some interesting, initial insights on the discourse function of DOM.

The Spanish data suggest that subject and object constituents can appear in the same clausal domain if there is (at least) one feature whose lexicalisation differs between subjects and objects, independently of its nature. Within the Romance family, Romanian provides empirical evidence in favour of this idea. As in the case of Spanish, Romanian has DOM and also allows VSO (cf. Leonetti 2014 on VSO; Irimia 2018 on DOM in Romanian). However, it is not necessary to assume any relationship between the availability of VSO and DOM in Spanish. In fact, Roussou & Tsimpli (2006) claim that VSO in Spanish is derived differently from Greek, whereby the verb is in the complementizer phrase, the subject in the tense phrase and the object in the verb phrase.

The conclusion that subjects and objects occupy two different syntactic domains is not problematic for the proposal sketched here, for which what is crucial is whether the subject is allowed to appear in postverbal position. Furthermore, our analysis relates to the linear order of the subject, verb and object rather than to the underlying syntactic structure. It is the mere possibility for subject constituents to appear in postverbal position in broad-focus contexts that makes the preverbal position in Spanish a less relevant cue for the identification of the antecedent of NSs than in Italian.

---

[14] An observation made by one of the anonymous reviewers, to whom we are very grateful.

[15] However, it should be considered that the results of the paragraph continuation tasks reported in the same contribution do not align fully with the results of these two corpus studies. We refer to von Heusinger et al. (2024) for discussion.

5 Conclusion

In conclusion, the primary aim of this study was to investigate whether the referential possibilities of NSs and OSPs differ across the three NSLs considered: Greek, Italian, and Spanish. By incorporating the Spanish data, this study provides novel empirical evidence that NSLs differ from each other to a greater extent than traditionally assumed. To explain this finding, we proposed a correlation between specific observable grammatical properties of a language—particularly word order, case-marking, and DOM—and the diverse patterns of pronoun interpretation.

For the Spanish data, we have speculated that two interrelated grammatical properties may influence the interpretation of NSs and OSPs. First, the availability of VSO word order potentially affects the extent to which the preverbal position serves as a strong cue for identifying the antecedent of a NS, similar to what Torregrossa et al. (2020) reported for Greek. Consequently, in NSLs where subject constituents can occur both pre- and post-verbally, there may be a reduced tendency to associate NSs with the preverbal constituent in the preceding sentence. Second, the existence of DOM may enhance the prominence of the marked object, thereby increasing the likelihood that Spanish NSs refer to object antecedents, in contrast to both Greek and Italian. As a result, in Spanish, the difference in the tendency of NSs to refer to subject versus object antecedents is even more attenuated than in Greek, as our results convincingly show.

We acknowledge several limitations in the present study that suggest potential lines for future research to further substantiate our conclusions and empirically test new hypotheses, as one of the reviewers has indicated. First, in the case of Greek and Spanish, it is necessary to examine whether NSs preceded by VSO sentences are interpreted differently than those preceded by SVO sentences. Second, for Spanish, it remains to be tested whether the tendency for NSs to refer to object antecedents is influenced by whether the object is DOM-marked. Lastly, non-grammatical factors (such as implicit causality, coherence relationships and topichood) also play a significant role in the interpretation of NSs and OSPs, potentially interacting with syntactic cues (Leonetti-Escandell 2024). This interaction suggests that the interpretation of referring expressions in discourse is a multifactorial process (Torregrossa et al. 2019) at the interfaces of language.

Abbreviations

DOM: differential object marking

NS: null subject

NSL: null subject language

OSP: overt subject pronoun

SpecIP: Specifier of the Inflectional Phrase

Data Availability

Stimuli and data analysis scripts are available at DOI 10.17605/OSF.IO/XG356.

Ethics and consent

For the present study, we received ethics clearance from the ethics board of the Deutsche Gesellschaft für Sprachwissenschaft.

Funding Information

The research for this article has been funded by the Deutsche Forschungsgemeinschaft (German Research Foundation, grant number: Project-ID 281511265 – SFB 1252 “Prominence in Language” in the project C03 “Reference management in bilingual narratives” at the University of Cologne.

Acknowledgements

We would like to thank Philipp Obrist, Albert Wall and Senta Zeugin for their support with the recruitment of the participants. Furthermore, we would like to thank Victoria Escandell-Vidal, the anonymous reviewers and the editor Lyn Tieu for their very useful feedback on the manuscript.

Competing interests

The authors have no competing interests to declare.

References

Alonso-Ovalle, Luis & Fernández-Solera, Susana & Frazier, Lyn & Clifton, Charles. 2002. Null vs. overt pronouns and the Topic-Focus articulation in Spanish. *Italian Journal of Linguistics* 14(2). 151–169.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Bader, Markus & Torregrossa, Jacopo & Rinke, Esther. 2023. Pinning down the interaction between animacy and syntactic function in the interpretation of German and Italian personal and demonstrative pronouns. *Discourse Processes* 60(9). 655–673. DOI: https://doi.org/10.1080/0163853X.2023.2252699

Bates, Douglas & Maechler, Martin & Bolker, Ben. 2012. *lme4: Linear Mixed-Effects Models using S4 classes (R package version 0.999999-0)*. http://cran.r-project.org/web/packages/lme4/index.html

Belletti, Adriana. 2004. Aspects of the Low IP Area. In Rizzi, Luigi (ed.), *The structure of CP and IP*. New York: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780195159486.003.0002

Cardinaletti, Anna & Starke, Michael. 1999. The typology of structural deficiency. A case atudy of the three classes of pronouns. In van Riemsdijk, Henk (ed.), *Clitics in the languages of Europe*, 145–223. Berlin: Mouton De Gruyter. DOI: https://doi.org/10.1515/9783110804010.145

Carminati, Maria Nella. 2002. *The processing of Italian subject pronouns*. Amherst: University of Massachusetts at Amherst, GLSA publications.

D'Alessandro, Roberta. 2015. Null Subject. In A. Fábregas, Antonio & Mateu, Jaume & Putnam, Michael (eds.), *Contemporary linguistic parameters*, 201–226. London: Bloomsbury Press.

Escandell-Vidal, Victoria. 2023. Estar + ILP. Testing the experiential commitment. *Spanish in Context* 20(2). 257–281. DOI: https://doi.org/10.1075/sic.00089.esc

Fábregas, Antonio. 2013. Differential object marking in Spanish: state of the art. *Borealis: An International Journal of Hispanic Linguistics* 2(2). 1–80. DOI: https://doi.org/10.7557/1.2.2.2603

Fedele, Emily & Kaiser, Elsi. 2014. Looking Back and Looking Forward: Anaphora and Cataphora in Italian. *University of Pennsylvania Working Papers in Linguistics* 20(1). 10.

Filiaci, Francesca & Sorace, Antonella & Carreiras, Manuel. 2013. Anaphoric biases of null and overt subjects in Italian and Spanish: A cross linguistic comparison. *Language and Cognitive Processes* 29(7). 825–843. DOI: https://doi.org/10.1080/01690965.2013.801502

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. DOI: https://doi.org/10.1007/s11049-007-9025-x

Fuß, Eric & Wratil, Melani. 2013. Der Nullsubjektzyklus: Etablierung und Verlust von Nullargumenten. In Fleischer, Jürg & Simon, Horst (eds.), *Sprachwandelvergleich – comparing diachronies*, 163–196. Berlin/New York: de Gruyter. DOI: https://doi.org/10.1515/9783110313468.163

Iemmolo, Giorgio. 2010. Topicality and differential object marking. *Studies in Language 34*, 239–272. DOI: https://doi.org/10.1075/sl.34.2.01iem

Irimia, Monica Alexandrina. 2018. Differential objects and other structural objects. *Proc Ling Soc Amer, Proceedings of the LSA* 3(50). 1–15. DOI: https://doi.org/10.3765/plsa.v3i1.4345

Kehler, Andrew & Rohde, Hannah. 2019. Prominence and coherence in Bayesian theory of pronoun interpretation. *Journal of Pragmatics* 154. 63–78. DOI: https://doi.org/10.1016/j.pragma.2018.04.006

Keller-Cohen, Deborah. 1981. Elicited imitation in lexical development. Evidence from a study on temporal reference. *Journal of Psycholinguistic Research* 10(3). 273–288. DOI: https://doi.org/10.1007/BF01067508

Lenth, Russel. 2020. *Estimated Marginal Means, aka Least-Squares Means*. R package version 1.4.6. https://cran.r-project.org/packages/emmeans/index.

Leonetti, Manuel. 2004. Specificity and Differential Object Marking in Spanish. *Catalan Journal of Linguistics* 3. 75–114. DOI: https://doi.org/10.5565/rev/catjl.106

Leonetti, Manuel. 2014. Spanish VSX. In *Romance Languages and Linguistic Theory*, vol. 6, 37–64. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/rllt.6.02leo

Leonetti-Escandell, Victoria. 2024. Reference and anaphora in Spanish and Italian: A view from adult monolinguals and bilingual children. PhD Dissertation. Universidad Complutense de Madrid.

López, Luis. 2012. *Indefinite objects: Scrambling, choice functions and differential object marking*. Cambridge: MIT Press. DOI: https://doi.org/10.7551/mitpress/9165.001.0001

Martín-Villena, Fernando. 2023. *L1 morphosyntactic attrition at the early stages: Evidence from production, interpretation and processing of subject referring expressions in L1 Spanish-L2 English instructed and immersed bilinguals*. Doctoral dissertation, University of Granada.

Mayol, Laia & Clark, Robin. 2010. Pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics* 42(3). 781–799. DOI: https://doi.org/10.1016/j.pragma.2009.07.004

Mikkelsen, Line. 2005. *Copular clauses: Specification, predication and equation*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.85

Nikolaeva, Irina. 2001. Secondary topic as a relation in information structure. *Linguistics* 39(1). 1–49. DOI: https://doi.org/10.1515/ling.2001.006

Nikolaeva, Irina & Dalrymple, Mary. 2011. *Objects and information structure*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511993473

Papadopoulou, Despina & Peristeri, Eleni, Plemenou, Evangelia, Tsimpli, Ianthi Maria & Marinis, Theodoros. 2015. Pronoun ambiguity resolution in Greek: Evidence from monolingual adults and children. *Lingua* 155. 98–120. DOI: https://doi.org/10.1016/j.lingua.2014.09.006

R Core Team. 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org/

Rizzi, Luigi. 2013. *Issues in Italian syntax*. Berlin, Boston: De Gruyter Mouton.

Rizzi, Luigi. 2018. Subjects, Topics and the interpretation of pro. In Petrosino, Roberto & Cerrone, Pietro & van der Hulst, Harry (eds.), *From sounds to structures: Beyond the veil of Maya, Studies in Generative Grammar (SGG)*, 510–530. Berlin, Boston: De Gruyter. DOI: https://doi.org/10.1515/9781501506734-019

Roussou, Anna & Tsimpli, Ianthi Maria. 2006. On Greek VSO again! *Journal of Linguistics* 42. 317–354. DOI: https://doi.org/10.1017/S0022226706003914

Tigău, Alina. 2021. Differential object marking in Romanian and Spanish. A contrastive analysis between differentially marked and unmarked direct objects. In Jabatek, Johannes & Obrist, Philipp & Wall, Albert (eds.), *Differential object marking in Romance*. Berlin: De Gruyter. DOI: https://doi.org/10.1515/9783110716207-007

Torregrossa, Jacopo & Andreou, Maria & Bongartz, Christine. 2020. Variation in the use and interpretation of null subjects: A view from Greek and Italian. *Glossa* 5(1). 95. DOI: https://doi.org/10.5334/gjgl.1011

Torregrossa, Jacopo & Bongartz, Christine & Tsimpli, Ianthi Maria. 2015. Testing accessibility: A cross-linguistic comparison of the syntax of referring expressions. *Linguistic Society of America: Annual Meeting*. DOI: https://doi.org/10.3765/exabs.v0i0.3046

Torregrossa, Jacopo & Bongartz, Christine & Tsimpli, Ianthi Maria. 2019. Bilingual reference production: A cognitive-computational account. *Linguistic Approaches to Bilingualism* 9(4–5). 569–599. DOI: https://doi.org/10.1075/lab.17026.tor

von Heusinger, Klaus. 2018. The diachronic development of Differential Object Marking in Spanish ditransitive constructions. In Seržant, Ilja & Witzlack-Makarevich, Alena (eds.), *Diachrony of differential argument marking*, 283–310. Berlin: Language Science Press.

von Heusinger, Klaus & Chiriacescu, Sofiana. 2013. The discourse structuring potential of differential object marking, the case of indefinite and definite direct objects in Romanian. *Revue Roumaine de Linguistique* 58. 439–456.

von Heusinger, Klaus & Duarte, Tiago & García García, Marco. 2024. Differential object marking and discourse prominence in Spanish. *Isogloss. Open Journal of Romance Linguistics* 10(1)/6. 1–37. DOI: https://doi.org/10.5565/rev/isogloss.394

von Heusinger, Klaus & Schumacher, Petra. 2019. Discourse Prominence: Definition and Application. *Journal of Pragmatics* 154. 117–127. DOI: https://doi.org/10.1016/j.pragma.2019.07.025

Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. Cambridge (Massachusetts): MIT Press.