

Deconstructing SE constructions: number agreement and post-syntactic variation

Javier Ormazabal^{1 2} & Juan Romero^{2 3}

[Revised version, August 2021]

[*impersonal se*] is a normal transitive structure with no exotic properties,
and we do not have anything interesting to say about it here

(Raposo & Uriagereka 1996)

Abstract

Most analyses of non-paradigmatic SE in Romance are constructed to derive the distribution of their agreement pattern, which forces them to distinguish passive and impersonal SE-constructions structurally, against all evidence. Instead, we uniformly analyze them as regular active sentences where the subject agreeing with T is *SE* itself, an argumental clitic pronoun, always human and lexically specified as [person], and lacking number ϕ features altogether. We show that number agreement in *SE*-constructions does not behave as a genuine syntactic agreement relation and cannot be used as reliable evidence to determine the subjecthood of the overt argument, which uniformly shows syntactic properties of direct objects. A new agreement asymmetry is revealed between postverbal and preverbal/null arguments which conceals two different postsyntactic processes with very distinctive properties: morphological *Mutation* of an object clitic into number agreement in a CLLD-like configuration, similar to other crosslinguistically attested agreement displacement processes; and *Number Harmony* of T with a close DP, a phenomenon not ruled by systematic syntactic or morphological conditions that we link to similar agreement patterns in other contexts. Strong evidence for Clitic Mutation comes from Clitic Climbing contexts and from idiomatic expressions containing clitics.

Keywords: non-paradigmatic *se*-constructions, clitic mutation, number harmony, person/number agreement, quirky subjects, Feature geometry, Agree, Match

1. Introduction

Non-paradigmatic *SE* in Spanish distinguishes two different constructions, traditionally referred to as *passive* (1a), where the verb agrees with its complement as in analytic passives (P-SECs), and *impersonal* (1b), where the verb shows a default third person agreement, (I-SECs):⁴

- (1) a. Se censuraron los documentos
SE censored.pl the documents.pl
'The documents were censored'
- b. Se censuró a los oponentes políticos
SE censored.sg DOM the opponents political.pl
'The political opponents were censored'

1 University of the Basque Country (UPV/EHU), <javier.ormazabal@ehu.eus>

2 Basque Group for Theoretical Linguistics (HiTT)

3 University of Extremadura (UEx), <juantxo.romero@gmail.com>

4 The only analysis we are acquainted of where passive and impersonal *se* are treated truly as the same syntactic construction is the one developed in a series of papers by Pujalte and Saab (especially see Pujalte & Saab 2012, 2014, Saab 2018, Pujalte 2018). There are important empirical and theoretical differences between their analysis and the one we propose here, but we share with them the leading idea of eliminating construction-based stipulative distinctions. See discussion below.

Distributionally, I-SECs (1b) are resorted to in all contexts except in those transitive configurations in which the object does not receive Differential Object Marking (DOM) (1a). Theoretical approaches overwhelmingly assume this state of affairs and complete the paradigm in (1) with the opposite pattern in (2), generally considered ungrammatical or dialectal in the literature:

- (2) a. * Se censuró los documentos
SE censored.sg the documents.pl
 b. *Se censuraron a los oponentes políticos
SE censored.pl DOM the opponents politcal.pl

This divergence is derived by assuming the existence of two constructions with different structural and case-assignment properties (3) (there are a few exceptions, most explicitly Oca 1914 and Otero 1972, 1973; also see references in footnote 4):

- (3) “If agreement is taken to be a diagnostic for nominative marking (Chomsky 1981, 1995), then [1a] clearly exhibits nominative case. Less clear is the status of [1b].” [Ordóñez & Treviño 2016 p. 238]

However, as observed by MacKenzie (2013), the paradigm in (1)-(2) constitutes a spurious simplification of the facts, a projection of academicist prescription that happens to fit very conveniently with the theoretical dichotomy illustrated in (3). Consequently, as a preliminary task, a thorough description of subject agreement facts is needed. In section 2 we show that variation on agreement in SECs is far more extended than assumed in previous analyses, which makes most theoretical approaches to SECs empirically unsound.

Furthermore, we uncover a structural distinction that has gone unnoticed in the literature so far:⁵ While number agreement with *postverbal* objects exhibits a high and random degree of variation internal to every dialect, variation with *preverbal* and *null* arguments is systematic and mostly dialectically determined. Agreement with preverbal and null arguments splits dialects between those where the null/preverbal argument covaries with number agreement (A-varieties), as in (4b), and those where both object clitic and subject number agreement are possible outputs in the formal representation of the verbal complement (C-varieties), as in (4a) (see section 2.2 for details).

- (4) a. [Los documentos] se los censuró
 The documents.**pl SE 3O.pl.msc censored.sg**
 b. [Los documentos] se censuraron
 The documents.**pl SE censored.pl**
 ‘The documents were censored’

⁵ DeMello (1995:71-72) observes this asymmetry with respect to the “educated language”, and Ordóñez & Treviño (2016) make a brief reference to it with respect to the Mexican dialect.

Our proposal is that (4a) and (4b) are two instances of the same Clitic Left Dislocation Structure (CLLD) where in (4b) post-syntactic morphological operations reshape the realization of the object clitic as number agreement. As we argue in section 4.2 in detail, our analysis is supported, among other evidence, by the behavior of SECs in clitic climbing configurations. In these configurations, the embedded clitic may appear attached to the right of the infinitival verb (5a) or ‘climb’ to the left of the matrix modal verb (5b):

- (5) a. (Ella) intentó censurar**las**
 (She) tried.sg to.censore.**Clfem.pl.**
- b. (Ella) **las** intentó censurar
 (she) **Clfem.pl.** tried.sg to.censore
 ‘She tried to censure them(fem)’

When SECs appear in a clitic climbing configuration, the clitic appears as an enclitic attached to the infinitive in all dialects (6a). But when it climbs, in C-varieties the clitic may manifest as such (6b), while in A-varieties it must show up as subject agreement (6c).⁶ In other words, when the clitic climbs, it *mutates* into subject agreement (6c).

- (6) a. se intentó censurar**las**
 SE tried.sg to. censor.**Clfem.pl.**
- b. se **las** intentó censurar
 SE**Clfem.pl.** tried.sg to.censor
- c. se intentaron censurar
 SE tried.**pl** to.censor
 ‘Somebody or other tried to censor them’

In this paper we argue that the change of the clitic into number agreement (henceforth *clitic mutation*) with preverbal and null objects exemplified in (4), and the phenomenon of number agreement with postverbal objects (1)-(2) (what we refer to as *number harmony*) are two independent post-syntactic processes, neither of which constitutes a genuine subject agreement relation with T in the syntactic component. As a consequence, we conclude that agreement phenomena cannot motivate a passive-type analysis of SECs.

Our second objective is to present a unified analysis of SECs as regular sentences that do not require any construction-specific provision. We argue that properties attributed to these constructions follow from the lexical features of the clitic *se*, an argumental pronoun in

⁶ MacDonald & Melgares (2018, ex. (19)) present a similar argument to argue that the clitic *le* that appears in non-leísta dialects is a repair strategy to solve a problematic object clitic *lo(s)/la(s)* in SECs (see footnote 14 below):

- (i) a. En Navidad, se suele poder abrazar **{lo/*le}**.
 In Christmas, se tends can hug him.ACC/DAT
 ‘In Christmas, one tends to be able to hug him.’
- b. En Navidad, se suele poder **{lo/*le}** abrazar.
- c. En Navidad, se **{le/*lo}** suele poder abrazar.

subject position that triggers agreement with T, together with the application of post-syntactic processes at the interface levels. We argue that these post-syntactic processes are triggered when the subject is not specified for number, as in the case of SE.

As mentioned, an immediate advantage of our proposal is that it dispenses with passive-like analyses that attribute to these structures emerging properties such as the removal of accusative case required to make the agreement pattern distinctions between \pm DOM objects in Spanish or, more generally, between the agreement/agreementless alternations in transitive SECs. Getting rid of the distinction between P-SEC (7a), I-SEC (7b), and Clitic SEC (7c-d) is consistent with the fact that there is no meaning difference between them (see Mendikoetxea 1999 and references therein) in spite of the fact that, with varying distribution, the three possibilities exist in all Spanish dialects.

- (7) a. Se censuraron los documentos [= (1a)]
SE censored.pl the documents.pl
 ‘The documents were censored’
- b. Se censuró a los oponentes políticos [= (1b)]
SE censored.sg DOM the opponents political.pl
 ‘The political opponents were censored’
- c. [Los documentos] se los censuró [= (4a)]
The documents.pl SE 3O.pl.msc censored.sg
 ‘The documents were censored’
- d. [A nosotros] se nos censuró
DOM us.1pl SE 1O.pl censored.sg
 ‘The documents were censored’

Evidence presented in section 2 shows that the alleged patterns underlying that distinction lack empirical support, depriving passive-based analyses of their *raison d’être*. An immediate consequence of our proposal and its supporting evidence is that agreement cannot be used in these structures as a reliable test to uncover underlying syntactic relations. This poses a general issue about the meaning of overt evidence, and questions the foundations of construction based approaches. As we show in section 2.1, in order to formulate a P-SEC in an empirically coherent way, a structure including a transitive verb and a [-DOM] object, i.e., a construction, is required. On the other hand, a pure derivational approach is perfectly equipped to explain the complete absence of semantic differences and provides a syntax able to accurately relate to the whole range of PF representations.⁷

The structure of the paper is as follows. Section 2 describes number agreement in

⁷ Similar arguments are presented for Dative constructions in O&R (2013a; b), where the same “construction” is reached via different syntactic paths which impose different restrictions. That accounts for why while doubled datives are non compatible with [+DOM] objects, non-doubled ones are fine in those contexts (also see discussion in section 5.1.3 regarding the relation between clitic mutation and ergative displacement).

SECs and shows that it does not behave as a genuine agreement relation. In section 3 we briefly summarize Ormazabal & Romero's [O&R] (2019a) logic and arguments to analyze SECs as regular active sentences where the complement of the verb is also its syntactic direct object and the clitic *se* is a pronominal subject. We also detail the ϕ -features specified in the lexical entry of *se*, which will trigger the agreement with T and, ultimately, the agreement patterns. In sections 4-5 we present the details of a derivation that integrates all the syntactic properties discussed in section 3 as well as the agreement facts discussed in section 2. Section 4 presents the syntactic derivation, common to all SECs: In a nutshell, transitive SECs are regular transitive sentences where *se* merges as the highest argument and behaves as a subject pronoun in the syntax, maintaining ϕ -agreement and nominative case checking-relations with T. An excursus in section 4.3 briefly discusses our proposal on Agree in O&R (2021) as applied to SECs. We also present our analysis of the asymmetry between null/preverbal arguments and postverbal ones. Structures containing preverbal or null arguments are argued to be cases of CLLD where the DO clitic *lo(s)/la(s)* ('him/her/it/them') in most cases mutate into number agreement at PF. We detail the syntactic derivation and define the domain of application of the clitic mutation rule, supporting the analysis with new evidence related to the clitic cluster paradigm discussed in (5)-(6) and to idioms that contain clitics. As for postverbal complements, we argue that subject agreement is a post-syntactic, extragrammatical phenomenon not ruled by systematic syntactic conditions, but amenable to a sociolinguistic approach whose domain of application is tightly circumscribed by the syntactic derivation (Otero 1973). Although the analysis of crosslinguistic--and, in particular, of cross-Romance-- variation is way beyond the scope of the paper, in section 6 we present the general lines of how it could be approached and some specifics about the differences observed in SECs in the Romance family. We close the paper with a general conclusions section.

2. Number agreement in detail

The regularity of number agreement in *se*-structures has been grossly overestimated, possibly for reasons of normative pressure (MacKenzie 2013). In this section we review all the attested agreement patterns in transitive SE structures and we conclude that number agreement cannot be used as a criterion to distinguish between P-SECs and I-SECs. However, a new asymmetry, formerly unnoticed in the literature, is observed between postverbal objects on the one hand, and preverbal/null ones on the other. This section is organized accordingly: we first analyze the distribution of agreement with postverbal objects (section 2.1), and then with preverbal and null ones (section 2.2). Finally, in section 2.3 we

briefly analyze agreement patterns with first and second person.

For each agreement-type discussed in this section, we have conducted an analysis of the data in the Real Academia Española's *Corpus del Español del Siglo XXI (CORPES)*, and added eventual searches in Google that corroborate our point (see appendix for details).⁸ We have also included other sources that we specify in place. Given the characteristics of the database, our main samples are mostly from written sources, which are more prone to be conservative and match the standard. We expect oral speech to depart from the norm to a larger extent, supporting our conclusions further, but the variation observed in the written samples is rich enough for our purposes.

2.1. Dysfunctional agreement with postverbal elements

According to traditional descriptions, the only syntactic elements that should trigger number agreement on the verb in SECs are canonical 3rd person objects that are not marked for DOM (1)-(2). However, these descriptions are not accurate. In this section we show that subject number agreement with postverbal elements cannot be accounted for in syntactic terms (see Gallego 2016, Planells 2018, and Pujalte 2018 for some observations in the same direction).

2.1.1 [-DOM] objects

Apart from some scattered observations now and then pointing out cases where agreement between the verb and [-DOM] objects fails (see especially Lemus 2014, Ordoñez & Treviño 2011, 2016, and Pujalte 2018), DeMello (1995), although restricted to educated Spanish speakers, is the only systematic study that makes a complete description on the range of variation of agreement patterns in SECs. Our analysis of CORPES suggests that Central American and Northern South-American dialects show a stronger tendency than Southern American and Peninsular Spanish for the verb in SECs to show in a default 3rd person singular form, with no subject agreement whatsoever, as exemplified in (8)-(9).

(8) a. Se censuró los documentos [cfr. (2a)]
SE V.sg DP.pl
'The documents were censored'

(9) a. se recuerda las versiones de Francesco Salviati, del Tintoretto...
SE remember.3sg the versions.pl of Francesco Salviati, of Tintoretto
'F. Salviati's, Tintoretto's... versions are remembered'

⁸ Except when indicated, all the examples were retrieved from CORPES between February and July, 2018, and in March 2020. The complete examples and source references are listed in the appendix. Each example there is identified with its corresponding number in the main text. The appendix also contains a short explanation of our data retrieval and analysis of CORPES and Google for each structure. Introductory examples are created to illustrate the basic pattern at each point of the discussion.

- b. ... donde **se** establecía *las bases del nuevo gobierno* [Venezuela].
 ... where **SE** established.**sg** the basis.**pl** of the new government
 ‘... where the foundations of the new government were established.’
- c. **Se** señalaba *los serios problemas...*
SE pointed.out.**sg** the serious problems.**pl** [Cuba]
 ‘The serious problems were signaled...’
- d. [...] que **se** haga *los procesos correspondientes* [Bolivia]
 that **SE** make.**sg** the processes.**pl** corresponding.**pl**
 ‘[...][be the corresponding processes made’

Lemus (2014), from El Salvador, observes that in his own dialect both options, agreement and agreementless, are acceptable and are used interchangeably, underlining the complete absence of semantic effects in this alternation. But this agreement alternation is much more general. First, it is well known that, independently of the dialect, bare plural objects quite commonly do not trigger agreement (Sánchez 2002 and references). A search on the Internet gives us 843.000 results for the agreeing form, *se venden_{PL} casas* (‘houses are on sale’), compared to 718.000 for the non agreeing one, *se vende_{SG} casas*. In the same direction, the collocation *dar las gracias* ‘to thank’ with the plural object *las gracias* (lit. ‘the thanks’) produces more hits with the non agreeing version (*se da_{SG} las gracias*: 150.000 results) than with the prescriptive agreeing one (*se dan_{PL} las gracias*: 119.000 hits) (Google 10/2/2018). Data are even more striking when a singular dative clitic is inserted between *se* and the verb: 288.000 results do not agree (*se le da_{SG} las gracias*) against 63.500 agreeing results (*se le dan_{PL} las gracias*); hardly an 18% follow the Academie’s rule.

Consider also a more general case. In recipes SECs are commonly used for describing cooking actions. Although recipes are a clear instance of highly conventionalized written texts, around 20% of the examples do not exhibit the expected agreement pattern (Google 28-6-2018):

- (10) a. *se fríe las patatas* (1.660) - *se fríen las patatas* (9.820)
 potatoes is/are fried
- b. *se cuece las patatas* (2.780) - *se cuecen las patatas* (11.500)
 potatoes is/are boiled
- c. *se pela las patatas* (1.210) - *se pelan las patatas* (13.600)
 potatoes is/are peeled off
- d. *se corta las patatas* (2.890) - *se cortan las patatas* (24.000)
 potatoes is/are cut
- (11) a. *se fríe los huevos* (3.370) - *se fríen los huevos* (9.920)
 eggs is/are fried
- b. *se cuece los huevos* (3.000) - *se cuecen los huevos* (11.100)
 eggs is/are boiled

It seems, therefore, that in addition to the mentioned dialectal trends, there is also a more general sociolinguistic tendency, so that, when the academicist pressure instilled by formal education diminishes, lack of agreement increases.⁹

Two LI reviewers suggest, following analyses for Italian (Cinque 1988; D’Alessandro 2007), that these facts might correspond to an alternation between a P-SEC and an I-SEC: when the internal argument agrees, there is a P-SEC, when it does not, it is an I-SEC (see Ordoñez & Treviño 2016 and references therein). However, this way of rephrasing the distribution is completely inconsistent with the facts in A-varieties. First, according to standard analyses, in I-SECs accusative case is not suppressed, but assigned to the object, as the presence of object clitics in these constructions clearly indicates:

- (12) a. A tu hermana se **la** vio en malas compañías
 DOM your sister SE 3Osgf saw in bad companies
 ‘Your sister was seen in bad company’
- b. Se **nos** obligó a hablar en público
 SE 1Opl forced to speak in public
 ‘We were forced to speak in public’

If non-agreeing forms in (9)-(11) were impersonal, we would expect parallel examples with clitics to be fully available. In other words, if I-SECs are allowed with postverbal objects (no number agreement), standard analyses predict that the accusative clitic should be available when the argument is left dislocated or null. That prediction is not borne out: Examples like (13b) with an object clitic are not acceptable in general (Ordoñez & Treviño 2016, MacDonald & Melgares 2018; sections 2.2. and 2.3). Contrasting with postverbal arguments (13a), preverbal ones always show number agreement (13c) and the clitic alternative is not possible (13b).¹⁰

9 A clear indication that agreement in this context is perceived as unclear by many Spanish speakers, is the numerous queries regular people make on this issue to normative linguistic institutions (Real Academia de la Lengua, Instituto Cervantes, FUNDEU, etc.). Otero’s (1972: 238) comment that while “educated people” exclusively use agreement forms, unagreeing forms are common “on the other side of the tracks”, goes in the same direction, although we do not think the sociolectal cut is as radical as he suggests.

10 A reviewer observes that for Mexican Spanish Ordoñez & Treviño (2016) do report the existence of a doubling strategy:

- (i) A estos terrenos se les vendió a un buen precio
 DOM these lands SE 3plDat sold.sg at a good price
 ‘These lands were sold at a good price’

Notice first that this is not a general pattern in dialectological terms: the only reference to it we know of is Ordoñez & Treviño’s discussion on Mexican Spanish, and it is not clear to us how extensive the phenomenon is even in that area. It thus cannot be considered a counterargument to our objection in the text, since the clitic strategy should be available each time the agreementless structure is an option. Second, as observed by Ordoñez & Treviño, the clitic that appears in that context is not the predicted accusative *lo* but the dative *le*. And, most importantly, the context is also exceptional in that the clitic doubles an inanimate DOM argument.

Altogether, these properties suggest that there are two independent processes involved in the case under discussion. On the one hand, the appearance in otherwise non-*leísta* dialects of a clitic *le*, instead of *lo*, in contexts where the doubled objects are marked with DOM, a very general phenomenon in American Spanish,

- (13) a. Se recibió tarde las invitaciones
SE received late the invitations
'The invitations were received late'
- b. *se (las) recibió tarde
SE 3Oplf received.sg late
'They were received late'
- c. Las invitaciones se recibieron/*recibió tarde
The invitations SE received.pl./received.sg late
'The invitations, they were received late'.

On the other hand, if P-SECs were like analytical passive in that Case is not assigned to the object, DOM would not be assigned and animate arguments would agree with the verb. However, that option is not attested at all (14), neither with 3rd person animate arguments nor with 1st/2nd person ones.

- (14) a. *Se censuraron los oponentes políticos
SE censored.pl the opponents political.**pl**
'The political opponents were censored'
- b. *Se censuraron/censuramos nosotros
SE censored.pl/censored.1pl we1st.**pl**
'We were censored'

The only way to allow Case assignment in (13a) and to block P-SEC in (14) would be to resort to some mechanism where DOM objects are excluded *a priori*. Given the situation in analytic passives, we cannot see how that could be captured derivationally in a principled way and, in fact, we know of no attempt in the literature to explain this distribution of facts. In that sense, traditional analyses, including generative ones, are constructionist in nature.

To compound matters, finding a coherent P-SEC/I-SEC proposal for the cases discussed so far would only cover one corner of the problem since, as we show next, the rule often fails to apply also on the opposite side of the equation. We find many cases where impossible syntactic agreement holds between the verb and any noun in the sentence (long distance agreement 2.1.3, complements of prepositions, and even nominal adjuncts 2.1.4).

2.1.2. [+DOM] objects

According to traditional descriptions, [+DOM] objects do not trigger agreement. As in the cases analyzed in the previous subsection, the description does not seem to meet the data in

except in the Southern Cone, with very deep historical roots (see footnotes 13 and 38, and references there). On the other hand, a geographically more limited phenomenon--most probably a linguistic innovation--in which DOM extends to inanimate NPs in certain contexts, very much in the same direction as the situation mentioned in footnote 18; see references there for extensive discussion. Obviously, in a variety where DOM has extended to mark inanimate arguments, contexts where *le* surfaces are expected also to include those cases where the doubled argument is a DOM inanimate, as in (i).

these contexts either, and the sentence rendered as ungrammatical in (2b), repeated here as (15), is much more common than usually assumed (see DeMello 1995; Planells 2017).

- (15) Se censuraron a los oponentes políticos [cfr. (2b)]
SE V.pl DOM DP.pl
 ‘The political opponents were censored’

As a matter of fact, examples in written texts are readily available in every Spanish dialect (16), and much more frequently in oral Spanish.

- (16) a. En 1996 se eligen a las primeras autoridades municipales [Nicaragua]
 In 1996 **SE elect.pl DOM the first authorities.pl municipal**
- b. ... y se rastreaban a los miembros del Partido [Mexico]
 and **SE track.pl DOM the members.pl of the Party**
- c. Al iniciarse la menstruación se aislaban a las jóvenes [Spain]
 When beginning menstruation **SE isolated.pl DOM the young.women.pl**

MacKenzie (2013) studies frequencies of sg./pl. in ‘se+verb+a+plural object’ sequences in CREA corpus for *elegir* (‘choose’), *invitar* (‘invite’), *nombrar* (‘name’) and *matar* (‘kill’). Agreement results range from 9.65% with *invitar* to 38.89% with *elegir*.

In sum, the distribution of agreement between [+DOM] and [-DOM] objects in Spanish hardly reflects anything more than a prescriptive choice. The weaker the influence of the Academy, both geographically and sociolinguistically, the greater the asystematicity of the agreement patterns in SECs: Agreement between V and DOM objects permeate the Spanish language, and it is only consistently rejected by speakers that show a highly prescriptive behavior in this regard.

2.1.3. Long-distance agreement

A context where this dysfunctional agreement shows up frequently is in long distance agreement relations, where the complement of an infinitival clause agrees with matrix V (Etxepare 2006; Bhatt 2005; Boeckx 2004, among others).

- (17) Se decidieron [censurar los documentos]
SE V.pl [Tenseless clause V infinitive DP.pl]
 ‘It was decided to censor the documents’

CORPES yields many such results for all geographical areas (see observations in the appendix), and a Google search for *se requieren hacer* (‘it is required to make’) yields 17.400 hits. Among the examples we observe not only long-distance agreement with [-DOM] objects (18a-b) (Gallego 2019) but, in combination with verbs as *nombrar* (‘to appoint’) also with [+DOM] ones (18c).

- (18) a. En esta profesión **se requieren hacer evaluaciones** [Spain]
 In this job **SE require.pl to.make evaluations.pl**
 ‘In this job you must conduct evaluations’
- b. Se valoran reducir las superiores a 1300 euros [Spain]
SE evaluate.pl to.reduce the superior to 1300 euros
 ‘Reducing those (pensions) higher than 1300 euros will be considered’
- c. Se requieren nombrar a los responsables [Paraguay]
SE require.pl to.appoint DOM the leaders
 ‘I is required to appoint the leaders’

Note that these structures disallow analytic passives (19a), and Clitic Climbing (19b).

- (19) a. *Las evaluaciones fueron requeridas hacer
 The evaluations were.pl required to do
- b. *las requirió hacer
 3Oplf required to do
 ‘She required to do it’

Therefore, agreement between the matrix verb and the complement of the embedded infinitive is completely unexpected in this syntactic context as well.

Moreover, irregular agreement is found even in configurations parallel to hyper-raising contexts (20).

- (20) a. Aquí se consideran que los hechos del 1 de octubre son hechos gravísimos
 Here SE consider.pl that the facts.pl of.the 1st of october are facts grave.superlat
 ‘Here it is considered that the October 1st facts are extremely serious facts’
- b. Se consideran que estamos frente a un renacimiento de la humanidad
 SE consider.pl that are.1pl in.front of a renaissance of the humanity
 ‘It is considered that we are facing a renaissance of the humanity’
- c. Se consideran que hay personas con categorías superiores e inferiores
 SE consider.pl that there.are persons with category superior and inferior
 ‘It is considered that there are people superior to others’

In (20) the matrix verb agrees with the subject of the finite subordinate clause. The number of hits in Google for this structure rises up to 8.260.000 (30-6-2021), including, as (20b) illustrates, many cases of number agreement with embedded first and second person subjects.

2.1.4. Other dysfunctional agreement patterns: P-Complements and adjuncts

Agreement in SECs is even messier than described so far. The verb may actually show agreement with different postverbal elements that in general syntactic contexts would never trigger agreement. It may agree with complements of lexical prepositions (Gallego 2016; all the examples come from peninsular Sp.).

- (21) Se hablaron [de los documentos]
 SE V.pl [pp P DP.pl]

- (22) a. Dijo que se hablaron [pp con las autoridades]
said that SE talked-3.pl with the authorities
'He said that the authorities were talked to'
- b. Se llegaron [pp a los 74,3 millones de operaciones]
SE arrived-3.pl to the 74,3 million of operations
'74,3 million of operations were reached'
- c. Es bueno reconocer cuando las cosas se hacen bien y se piensan en nosotros
is good to.acknowledge when the things SE make.pl well and SE think.pl in us
'It is good to acknowledge when things are well done and people think about us'

Although marginal compared with some of the phenomena previously presented, it is not an isolated unexpected match: A Google search for “se hablaron de temas” (‘talk about issues’) produces 6.350 results of plural agreement between the verb *hablar* and P complement *temas* in Spain, Argentina, Mexico, Colombia, etc. (23). These figures, in many cases from formal sources, are high enough not to dismiss them.

- (23) [...] aclaró que tampoco "se hablaron de temas de la farándula" [Argentina]
(he) explained that neither SE talk.pl about show business issues.pl
'He explained that show business issues were not mentioned either'

This kind of agreement can even be found in long distance relations: For *se deben recurrir* (‘it must be resorted to’) Google returns 26.900 hits (24).¹¹

- (24) Indicó que se deben recurrir a otras figuras [Paraguay]
He pointed out that SE must.pl resort to other criminal definitions.pl

Finally, it is very common to find examples where agreement is not even triggered by arguments, but by temporal DP-modifiers.:

- (25) Se bailan los lunes
SE V.pl DP_{Adjunct}.pl
'People dance on Mondays'
- (26) a. ... donde no se abren los domingos, ...?
where no SE open.pl the sundays.pl, ...?
- b. Se trabajan los fines de semana y festivos
SE work.pl the weekends.pl, and holidays

In the case of *se abre los domingos* (‘it is opened on sundays’), for instance, in a Google search (2018/06/05) we find 27.400 hits for the irregular agreement pattern *se abren los domingos*, while 27.300 for the expected one, *se abre los domingos*. Although data in this subsection do not fit with standard P-SECs (verbs are not transitive), their morphological behavior mimics it. In the next section we argue that these agreement patterns cannot be understood as subject agreement in any syntactically coherent way.

¹¹ Although the sequence “se pueden recurrir a” is ambiguous, and agreement may be regular meaning *appeal* (*esas sentencias se pueden recurrir al Supremo*, ‘those sentences may be appealed to the Supreme Court’), the first 50 examples are all instances of irregular agreement with the complement of the preposition and meaning *resort to*.

2.1.5. Number agreement is not subject agreement

If we consider the whole range of data, it is clear that something unexpected is happening from the point of view of subject agreement. As said, figures are big enough not to dismiss them as performance errors. When we turn to true subject agreement, it is much more regular and predictable: neither it appears with unexpected elements nor it suddenly disappears with agreeing ones. Thus, in contrast to (26), temporal nominal adjuncts never ever trigger agreement in other contexts, not even with impersonal verbs:

- (27) llueve/*llueven todas las tardes/ los domingos
rain.sg/rain.pl every the afternoons/the Sundays
'It rains every afternoon/on Sundays'

And, so-called *attraction errors* aside, true subject agreement is never affected, not even in the most favorable contexts. Thus, for instance, we showed in section 2.1 that bare plurals fail easily to agree in SECs (Mendikoetxea 1999; Sánchez 2002). In clear contrast, bare plural subjects of unaccusative sentences cannot avoid agreement (28).

- (28) cayeron/*cayó almohadillas
fell.pl/fell.sg small pillows
'Small pillows fell'

Consider now the causative-inchoative alternation. As in non paradigmatic SECs, in inchoative sentences (i) the verb ends up agreeing with its complement, and (ii) there is a SE intransitivizing the verb (29). In spite of its similarity, agreement failure is unattested.

- (29) se durmieron/*durmió los niños
SE_{AC} slept.pl/slept.sg the children
'Children fell asleep'

Finally, consider analytic passive sentences. In contrast to what we have seen in section 2.1.1 for P-SECs, lack of agreement between the verb and the complement is clearly ungrammatical.

- (30) a. fueron/*fue asesinados muchos opositores al régimen
were/was killed many opponents to the regime
'Many opponents to the regime were killed'
b. fueron/*fue rescatados los cuerpos
were/was recovered the bodies
'The bodies were recovered'

The consequence of these results is clear: agreement facts in non paradigmatic SECs do not work as predicted by theories based on syntactic agreement. In order to explain their agreement patterns, we cannot simply treat them as subjects, because this is not the way subject agreement behaves. As shown in (27), we cannot resort to a trivial default syntactic agreement notion either, because default agreement is systematic, in clear contrast to subject agreement behavior in these constructions (D'Alessandro 2007; López 2020). Further evidence in the same direction is presented in the next section, where we show that there is a

clear asymmetry regarding agreement between postverbal and preverbal arguments, an asymmetry that, again, is not present in regular subject agreement relations.

2.2 Agreement with preverbal/null arguments

When the internal argument is null¹² or appears preverbally, the situation is radically different. That argument covaries with an object clitic in some cases and with a verbal plural agreement in others:

- (31) a. DP/∅ SE **Object.Clitic** + V
 b. DP/∅ SE **V.Number.agreement**

The properties of (31a-b) are very different from plural agreement with postverbal arguments. Both structures in (31) exist in all dialects of Spanish, but their distribution diverges slightly, yielding what we have called A(greement) and C(litic)-varieties.

First, preverbal [+DOM] arguments, including first and second person ones, exhibit a completely uniform pattern cross-dialectally: they are always doubled by a pronominal clitic (32a) and never trigger subject agreement (32b) (cfr. (15)).¹³

- (32) a. Ayer (a las oponentes políticas) se *(las) censuró
 Yesterday **DOM** **DPpl.** **SE Cl.ac.pl.** **V.sg**
 ‘The political opponents were censored’
 b. *Ayer (a las oponentes políticas) se censuraron
 Yesterday **DOM** **DPpl.** **SE** **V.pl**
 ‘The political opponents were censored yesterday’

Second, verbal agreement is strictly circumscribed to null/preverbal [-DOM] arguments in all varieties; no significant exceptions are found in corpora nor in oral elicitations in which any other preverbal element might trigger number agreement.¹⁴

- (33) a. Finalmente se censuraron/ *censuró [referring to the documents]
 Finally **SE** **V.pl** **V.sg**
 ‘Finally, they were censored’
 b. Los documentos se censuraron/ *censuró
DP.pl **SE** **V.pl** **V.sg**
 ‘The documents were censored’

¹² Following a reviewer’s suggestion, we use “null argument” as a term for contexts where the referential DP agreeing with the verb or doubled by the clitic is not overt; no theoretical content should be attributed to it. In fact, we do not think there is a null argument strictly speaking. In section 4 we argue that in both cases, the argumental DP is the clitic itself, whose reference is a definite description previously introduced in the discourse (in the case of null topic contexts) or by the left dislocated DP (when that preverbal topic is present). In section 4 we will accommodate the terminology (null/left dislocated topics) accordingly.

¹³ There is interesting variation regarding which specific clitic is used and what the extent of DOM is in each speaking area, but it seems to be due to dialectal differences on DOM itself (Fernández Lagunilla 1975, Mendikoetxea 1999, 2008, Ordóñez 2018, De Benito 2013, Macdonald & Melgares 2018), rather than different conditions on SECs. The only case that does seem to be sensitive to SECs is the presence of *leísmo* in non-*leísta* dialects (see Mendikoetxea & Battye 1990, Fernández Ordoñez 1999, Rigau & Picallo 1999; Ordoñez & Treviño 2016, MacDonald & Melgares 2018). See footnote 38 in section 5.1 for discussion.

¹⁴ We will come back to the special case of relative clauses and A’-dependencies in section 5.2.

In C-varieties the clitic strategy extends to [-DOM] objects, so that the described agreement option in (33) coexist with the clitic one in (34):

- (34) a. Los documentos se *(los) censuró *C-Varieties*
 DP.pl.**pl** SE Cl.m.**pl** V.**sg**
 ‘The documents, they were censored’
- b. Se *(los) censuró
 SE Cl.m.**pl** V.**sg**
 ‘They (the documents) were censored’

Third, as we note in the examples (33)-(34), with preverbal/null [-DOM] objects agreement is mandatory; there are no relevant cases where neither agreement nor object clitic would appear (cfr. (13)). Variation is strictly established on the agreement/clitic strategy distinction, and it does not affect number.

The clitic strategy is reported in the literature (Santiago 1975, Sánchez 2002, and references), but it has not been exhaustively described, and standard theories on SECs tend to ignore it. Essentially, the clitic strategy coexists with the agreement one in the southern half of South America (*NGLE*, García Negroni 2002). In those areas we commonly find them both in the CORPES, and all our (linguist) informants from the Buenos Aires area agree that (i) they are equally available, and (ii) there is no semantic difference between them (also see appendix):^{15 16}

- (35) a. Se las puede sujetar en postes [las orquídeas]
 SE 3Opl.fm can fasten to wooden poles [the orchids]
- b. Se pueden sujetar en postes [las orquídeas]
 SE can.3pl fasten to wooden poles [the orchids]
 ‘The orchids can be fastened to wooden poles’

15 We are very thankful to Carlos Muñoz, Mercedes Pujalte, Andrés Saab and Pablo Zdrojewski for patient discussion and important clarifications.

16 There is a subtle split between null and overt preverbal objects. For null arguments, most of our informants observe a slight preference for (36b), also noting some tendency to use the clitic strategy in colloquial contexts and the agreement one in written or formal ones. With preverbal objects, the situation is reversed and our informants (from Argentina) agree that they highly prefer the agreement pattern over the clitic one. In fact, they find sentences like (ia), not ungrammatical but “strange” or even “somehow degraded”, in contrast to (ib), which they consider completely natural:

- (i) a. ??Las orquídeas se las puede sujetar en postes de madera
 The orchids **SE 3Opl.fm** can fasten to wooden poles
- b. Las orquídeas se pueden sujetar en postes de madera
 The orchids **SE can.3pl** fasten to wooden poles

However, it may be a particularity of the Rioplatense dialect. This is confirmed by our results in CORPES, where we find almost no clitic example with preposed objects in Argentina and Uruguay, but numbers increase as we go North, and examples are extremely common in Ecuador.

Our Argentinian informants tend to “save” these examples marking the preverbal object with DOM, which suggests that in Rioplatense the conditions under which DOM applies are gradually changing. According to Di Tullio & Zdrojewski (2013), Di Tullio, Saab & Zdrojewski (2019), SECs constitute one of the contexts where this phenomenon is most evident, together with ECM and causative contexts, but it is general, and very productive in left dislocated contexts. Also see Liman Spanish (Sánchez 2006, 2010; Sánchez & Zdrojewski 2013 and references) and Santiago de Chile’s dialect (Silva-Corvalán 1980) for a similar process.

- (36) a. *estas ponderaciones se las realiza* por sectores [Bolivia]
 those considerations.**pl.fm** SE 3Opl.fm make.sg by sectors
- b. *estas ponderaciones se realizan* por sectores
 those considerations.**pl.fm** SE make.**pl** by sectors
 ‘Those considerations are made by sectors’

In clear contrast, in C varieties this strategy is impossible with [-DOM] postverbal objects:

- (37) **se las puede sujetar las orquídeas en postes de madera*
SE 3Opl.fm can fasten **the orchids** to wooden poles
 ‘The orchids can be fastened to wooden poles’

In the rest of the dialects object clitics with preposed and null [-DOM] objects are sometimes attested, but in general terms subject number agreement (33) is the sole strategy available with preverbal/null [-DOM] arguments for all speakers in A-varieties,¹⁷ including dialects mentioned in section 2.1.1 where postverbal arguments tend not to agree.¹⁸ The result is that SECs may be agreementless with postverbal arguments but they show obligatory agreement with preverbal ones:

- (38) a. Se censuró/censuraron los documentos
 SE V.sg/ V.pl DP.pl
- b. Los documentos se censuraron/ *censuró
 DP.pl SE V.pl V.sg
 ‘The documents were censored’

Standard descriptions fall short regarding the importance of the clitic structure and its alternation with the agreement one. Note that, inasmuch as our description is right,

17 Examples from a variety of geographic areas are occasionally elicited (e.g. (i)) but they are uncommon outside the area described in the text as belonging to the C-variety, where examples are legion.

- (i) a. Cuando **se la** desenchaja, *la luz de la lámpara* alumbr... [Mexico]
 When **SE 3Osg.fm** disarrange.sg, the light of the lamp illuminates
- b. En cuanto a *las incorpóreas*, no **se las** veía como cosas
 With respect to the incorporeal, do not **SE 3Opl.fm** saw.sg as things
- c. ... *las palabras del bautizo*, **se las** degrada. [Dominican Rep.]
 ... baptism words, **SE 3Oplfm** degrade.sg
- d. Por el contrario, si **se los** considera variables, [C. Rica]
 On the contrary, if **SE 3Oplms** consider.sg variable

This clitic/agreement alternation has a strange status outside that area: even for people that do not elicit the clitic form (including the authors), sentences in (i) do not strike as ungrammatical, they often go unnoticed if elicited by other speakers, they do not induce interpretive issues and, as said, some examples appear in the corpus now and then (see appendix), but speakers clearly differentiate the status of the clitic strategy as extraneous to their own grammar and would not make use of it.

18 Ordoñez & Treviño (2016) and MacDonald & Melgares (2018) report that in Mexico and Honduras respectively the clitic strategy is impossible with [-DOM] objects. Our CORPES data include some cases from both areas, but they are very infrequent (see previous footnote):

- (i) a. [... las estrellas fugaces], **se las puede ver** con mayor facilidad.
 the shooting stars SE cl.fpl may.sg see with beager easiness
 ‘[... the shooting stars], you may see them more easily’
- b. [Sus funciones...] cuando se las quiere simular por esquemas clásicos
 Their functions when SE cl.fpl want simulate by schemes classics
 ‘[...] when you want to simulate them by classical schemes’

theoretical approaches based on constructions are forced to postulate three different syntactic constructions --passive, impersonal and clitic SECs-- that show no meaning or use differences and, quite tellingly, extend to all dialects in one way or another.

2.3. *First and second person objects*

To complete the picture, first and second person objects show no dialectal or idiolectal variation whatsoever: They never ever allow number or person subject agreement, and always manifest themselves through an obligatory object clitic, as in any transitive clause.

- (39) a. * (nosotros) se censuramos /censuraron (nosotros)
 us SE censored.1pl/censored.pl us
- b. (A nosotros) se nos censuró
 ‘(Us), we were censored’

Examples like (39a) are completely unconceivable for any speaker of Spanish: there is no exception reported in the literature, no [*se + Verb(2nd pers. pl.)*] result in CORPES and in the rare [*se + Verb(1st pers. pl.)*] combinations it is clear that the first person plural does not refer to some agreeing internal argument of the verb, but they are non-corrected typos or anacoluthons (e.g. *si lo que se pretendemos es...* ‘if what SE we pretend is that...’).¹⁹

Once again, this sharply contrasts with analytical passives, where first and second person objects are legitimate candidates for passivization and, obviously, cannot surface as object clitics:

- (40) a. Nosotros fuimos censurados
 we were censored
- b. * A nosotros nos fue censurados
 DOM us us was censored
 ‘We were censored’

Given the availability of first and second person subjects in other constructions across the board and, especially, in analytical passive (40a), its failure in structures like (39) comes as a surprise in the P-SEC/I-SEC analysis.

2.4. *Summary*

The behavior of number agreement in *SECs* is completely exceptional among Spanish subject agreement relations in different respects:

- (i) With preverbal/null arguments, subject agreement alternates with object clitics in some dialects, systematically occurs with [-DOM] objects only, and never fails.

¹⁹ First and second person objects have been analyzed as object agreement markers shaped like clitic doubling, as opposed to third person clitics *lo(s)/la(s)*. See especially Roca 1996, O&R 2013a, Odria 2017, Rodríguez-Ordoñez 2016, Alcaraz 2021, and references; also see Baker & Kramer 2018 and references for a different view on clitic doubling. We briefly come back to these cases in section 4.2.

- (ii) With first and second person as well as with 3rd person [+DOM] objects, no agreement is possible; they are uniformly represented as object clitics.
- (iii) With postverbal elements agreement is often erratic; it frequently fails with [-DOM] objects and unexpectedly occurs with other DPs : [+DOM] DOs, long distance [±DOM] items--even hyper-raising--or complements of Ps and adjuncts.

This distribution sharply contrasts with analytical passives, where subject agreement is robust and systematic.

In the remaining of this paper we present an integral analysis of non-paradigmatic SECs that derives the distribution in (i)-(ii); we also argue that although (iii) should be analyzed by extra-grammatical mechanisms yet to be determined—thus, its distribution is not predictable in syntactic or morphological terms—the properties of the syntactic derivation do constraint their domain of application. With that goal in mind, we first summarize the gist of O&R’s (2019a) argumentation supporting a unified analysis of I-SECs and P-Secs, and present the lexical properties of *se* that make it possible. In sections 4 and 5 we present the details of the derivation and discuss the post-syntactic operations responsible for “subject” number agreement with the object and, more specifically, the distribution of facts presented in this section.

3. The properties of *SE* constructions in a nutshell

The primary focus of most analyses on *SECs* has been to justify P-SECs as a “selective passive” where accusative case is removed, but only for [-DOM] complements, so that subject agreement may target them.²⁰ These analyses not only require mechanisms *ad hoc* for these constructions, they are also based on two empirical claims that have been rejected in section 2. First, under the right structural conditions [-DOM] objects may be realized as object clitics, which dismantles the idea that it may not show up marked with accusative case.²¹ Second, preverbal and postverbal arguments --allegedly, equally genuine subjects under standard analyses-- present a radically different behavior, which questions their subjecthood. More generally, its behavior in *SECs* invalidates number agreement as evidence for nominative Case or subjecthood. This alone already indicates that the passive

20 Proposals can be broadly divided between those which argue for two different *se* tokens, and those which argue that *se* is inserted in different designated positions. For a good state-of-the art on non-paradigmatic *se* and throughout discussion of the issues involved see, especially, Mendikoetxea (1999) for Spanish, D’Alessandro (2007) for Italian and Sánchez (2002) for Romance in general with special attention to Spanish. We set aside the case of Romanian, which may require an independent analysis (Dobrovie-Sorin 1998).

21 Mendikoetxea (1992), Ordoñez (2004) and Torrego (2008) argue that Italian and Spanish differ in that the former allows the clitic strategy but Spanish does not. As the discussion in the previous section shows, that is not the right generalization. See also the previous footnote.

approach is not on the right track but, in addition, there are independent positive arguments that confirm the same conclusion.

3.1. A regular active sentence with *SE* as the subject

Our proposal (O&R 2019a)²², developing a suggestion by Oca (1914),²³ is that P-SECs are just active transitive sentences, where the internal argument of V is a regular direct object and the subject is the clitic *se*. The peculiarities of these constructions --especially those regarding agreement-- are due to the lexical properties of *se* as the subject agreeing with T, with no difference between passive and impersonal SECs:

A) P-SECs are active sentences in all morphological or syntactic senses: far from having different voice, impersonal and passive *se* have the same active morphology (Raposo & Uriagereka 1997; Rivero 2001; O&R 2019a) and their syntax is that of an active sentence. First, P-SECs may be formed with all transitive predicates allowing [-DOM] complements, (with the only restrictions imposed in C below). That includes, crucially, predicates completely reluctant to analytic passivization, where it is not even clear that the alleged accusative case to be erased existed in the first place (e.g. measure complements; Mendikoetxea 1999). Second, they do not accept *by*-phrases (Pujalte 2013, Saab 2014, *Nueva Gramática de la Lengua Española [NGLE]* 41.6.1, and references). Finally, they neither “passivize” 1st/2nd person objects nor animate ones (see section 4).

B) Consequently, the internal argument is the syntactic object of the sentence, not its subject, contrary to what number agreement might suggest. That is shown, among other things, by the fact that it can be modified by pseudorelatives (Aldama 2016, O&R 2019a), and that, unlike subjects of analytical passives, it may not be controlled from outside (MacKenzie 2013). Furthermore, idiomatic lectures of fixed object idioms are not lost in P-SECs, showing that grammatical relations have not been altered (O&R 2019 also see section 5.1).

C) The subject of the sentence is the clitic *se* (Oca 1914). As such, it shows properties typical of overt pronominal subjects in Spanish, including its animacy interpretation: the structure is only possible when the subject is argumental, and it is disallowed by verbs that do not select animate subjects (Mendikoetxea 1999). *Se* must appear in a Case-position, and

²² In many respects, the present paper and O&R 2019a complement each other and may be seen as two sides of the same proposal. Space limitations and the very dynamic of the paper advise against extending section 3 further by repeating the arguments there. The reader is encouraged to go to the original paper, to Mendikoetxea (1999), and references there for extensive discussion on the properties listed in this section as well as for some additional ones.

²³ Otero (1972, 1973) makes a similar proposal; also see Raposo & Uriagereka 1996 for some suggestions in the same direction, although they end up proposing a mechanism of accusative-absorption in the direction of standard passive-like analyses.

thus, is incompatible with infinitives that do not license nominative subjects, but allowed in tensed clauses in general and also in infinitival constructions where overt pronominal subjects are licensed (Cinque 1988; Mendikoetxea 1999; O&R 2019a). It also raises in raising-to-subject contexts, as required by case considerations (Mendikoetxea 1999, Martins & Nunes 2016 and O&R 2019a, section 3.2 for extensive discussion).

D) *Se* is syntactically active, and it may host secondary predication (Demonte 1986; Rivero 2001; Martins & Nunes 2016; O&R 2013b and references there; but see Pujalte & Saab 2012 for different judgments; also see arguments in Collins 2017 for a related discussion with respect to short passives in English). In addition, it shows the same obligatory control/disjoint reference distribution in infinitive/subjunctive alternations as any other active sentence, no matter whether it is interpreted as a generic or an indefinite, and may bind anaphoric elements vague enough not to conflict in features with it, as we argue next.

The idea that *se* is the missing argument required in SECs has already been proposed for I-SECs (Cinque 1988, Mendikoetxea 1999, D'Alessandro 2007, and references therein). Moreover, the reasons to discard Oca's idea that *se* is a subject clitic are theory-internal pre-minimalist arguments that do not hold anymore (see O&R 2019a, section 3 for discussion). In sum, a derivation of I-SECs and P-SECs where *se* itself is the sentential subject and participates in a regular nominative case relation with T straightforwardly derives all the structural properties of non-paradigmatic SECs listed above in a natural and unified way. In the next subsection we present the details of *se*'s lexical entry. Extending MacDonald (2017), we consider *se*'s binding properties as particularly indicative of its syntactic status, giving us a clear indication about its formal feature specifications.

3.2 Interpretable features in *se*

Saab (2014) observes that unlike other impersonal elements such as *uno* 'one' or *la gente* 'people', the subject of SECs does not allow pronominal binding in (41), and concludes that there is no syntactic subject.²⁴

²⁴ Also see Otero (1986) and Burzio (1986). Burzio discusses the contrast in (i) and notes, attributing the observation to Cinque (p.c.), that *si* cannot bind a possessive (we adapted Burzio's glosses and translations):

- (i) b. * SI ama i suoi-loro eroi
SE loves the his/their heroes
'One loves his heroes'
- b. Si loda spesso se stessi
SE praises often themselves
'One often praises oneself'

Burzio interprets this as indicating that SI has number features but lacks person features, and third person reflexive *se stessi* is an impersonal form, while *suoi/loro* are specified for third person; but see discussion in the text for a different interpretation. We are very grateful to Andrés Saab, who brought Burzio's observation to our attention and contributed with very valuable discussion.

- (41) Aquí uno/la gente/*se puede dejar su saco y marcharse
 here one/the people/SE can leave.INF **his/her** coat and leave
 ‘Here one/people may leave their coat and leave’

However, the inability to bind the pronoun may also follow as a consequence of a feature mismatch between *su* (‘his/her’), a third person pronoun, and its antecedent *se*. Thus, it has been observed that determiners may act as bound variables in Romance, but unlike the pronoun *su*, which is specified for third person, this bound determiner gets its person interpretation from the binder.²⁵ Contrasting with the binding failure in (41), *se* may bind the less specified determiner variable (example from MacDonald 2017):

- (42) Aquí **se** puede dejar **el** saco.
 here **SE** can leave.INF **the** coat
 ‘One_i can leave his/her_i coat here.’

If, as we propose, the divergence is derived from the feature specifications of *su* and the determiner,²⁶ two conclusions may be driven from the contrast in (41)-(42). First, *se* is a syntactically active subject and, as such, has the ability to bind a c-commanded anaphoric element; in that sense, non-paradigmatic *se* clearly contrasts with other *se*-clitics that cannot enter into binding relations (Alcaraz 2021). Second, given the contrast between minimally specified determiners and fully specified 3rd person pronouns, we must conclude that *se* is not 3rd person. This conclusion is supported by the fact, largely discussed in the literature, that SEC interpretation is not confined to third person, but it may refer to any animate argument, including first (43a) and second person (43b) (examples from Oca 1914; also see Cinque 1988, Menuzzi 1999, Mendikoetxea 1999; D’Alessandro & Alexiadou 2003, O&R 2019a, among others).

- (43) a. ¿se puede? [knocking on the door]
 SE can.3s
 ‘May I?’
 b. aquí no se habla [a father looking at his son]
 here no SE speak.3s
 ‘You cannot speak here’

25 We frame the discussion in a weak definite approach (e.g. Guéron 1983, 2006 and much work thereafter; also see Espinal & Cyrino 2017 for recent extensions) because it is based on the morphological features of the determiner and maps most directly to the observation that determiners and possessive pronouns behave differently with respect to binding by *se*. However, in virtually all approaches (e.g. Vergnaud & Zubizarreta’s 1992) the non-denoting determiner and, importantly, its difference with possessive pronouns plays a role and our point may be equally raised.

26 Eric Reuland (p.c.) observes to us that in general, underspecification does not create an offending feature mismatch, and he mentions two such cases: *zich*, underspecified for number and gender, is generally bound by a full DP; similarly, in Russian a full DP may bind *sebja*, underspecified for person, gender and number, etc. In fact, that is a very general pattern with anaphors and bound pronouns. However, observe that in all those cases, the most specified object is the binder (*Jan*, *Ivan*) and the least specified one is the bindee. In contrast, the binding problems with SECs arise in the reverse situation, where the binder *se* (underspecified for person features and lacking number) fails to bind *su*, which is specified for person and number. In that respect, the contrast with *uno* as a potential binder of *su*, mentioned by Saab, and the possibility of weak determiners in the bindee position are quite revealing.

The specific interpretation *se* adopts is mostly determined on pragmatic grounds. Notice, also, that first and second person subjects in the next clause may corefer with the subject of the impersonal construction in these contexts:²⁷

- (44) a. ¿*se* puede?... Si te viene mejor, vengo en otro momento...
 SE can.3s If O2sg come better, come.1sg in other moment
 ‘May I?... If it suits you better, I will come some other time...’
- b. ¡Aquí *se* cumplen las reglas! Si no, te vas a vivir por tu cuenta.
 Here SE comply.pl the rules! If not, cl2sg go.2sg to live by your own
 ‘Here you must comply with the rules! Otherwise, you (addressee)
 better go live on your own’

Since, as said, *se* is always interpreted as animate, following Richards (2014), we assume that animacy in pronouns is encoded as a person feature. We thus interpret its ability to bind anaphoric determiners but not 3rd person pronouns as indicating that the person feature is underspecified, in the sense that it lacks a specific 1st, 2nd or 3rd value. Let us see how these results may be put together.

Φ-features, as well as their internal organization, play an important role at different grammatical levels and their interfaces (also see Ackema & Neeleman 2018 and references for extensive, though slightly different, discussion):

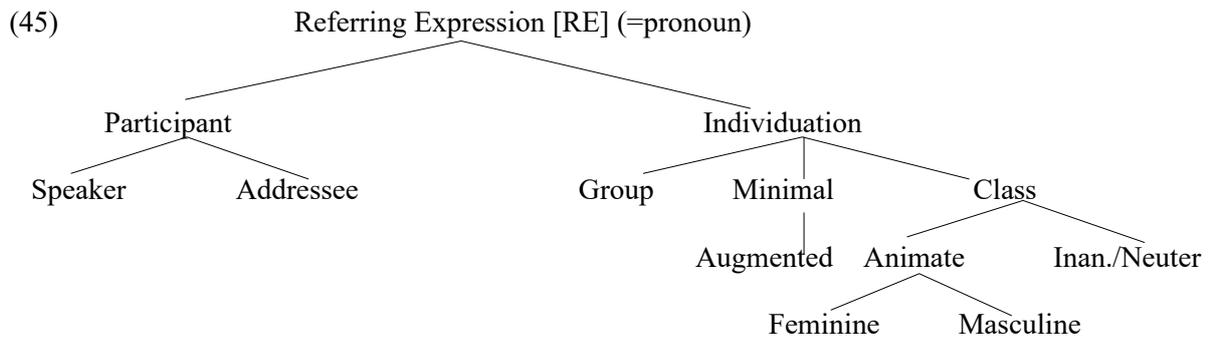
A) At the computational level, interpretable features determine the syntactic behavior of lexical items and the corresponding feature values of the functional projections they agree with.

B) At the semantic level the feature structure of these syntactic objects—in the cases we are discussing, the pronouns--must be interpreted.

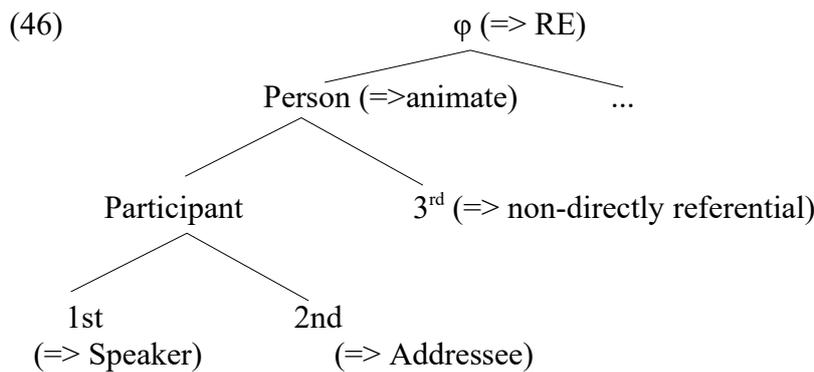
C) The morphology component operates on the features of the pronoun as well as on the non-interpretable features valued in the agreeing functional heads.

Restricting our attention to its person features, semantically *se* must be treated as truly underspecified for any person value and interpreted contextually. In Spanish, it must be therefore distinguished from third person animate pronouns (*ella/él* ‘she/he’) on the one hand, and from no-person arguments such as sentential subjects or the object clitics *lo(s)/la(s)* (‘it’/‘he’/‘she’/‘them’), as we discuss below (also see Trommer 2008 for morphological arguments in the same direction). Most feature hierarchies stemming from Harley & Ritter’s (2002) proposal on pronouns are morphology-oriented and do not leave enough room for such fine-grained distinctions in the syntactic derivation, far less in the semantic component. In particular, both their hierarchy in (44) and their conception of underspecification are too narrow (see Bianchi 2003, D’Alessandro & Alexiadou 2003, McGinnis 2005, D’Alessandro 2007, among others, for discussion).

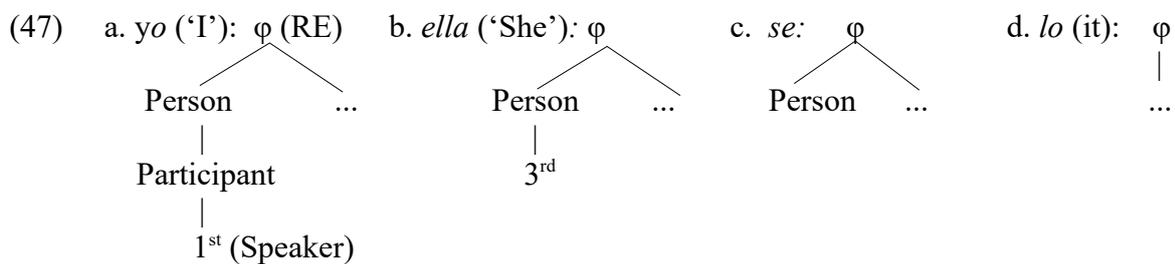
²⁷ We are grateful to Eric Reuland for bringing those facts to our attention.



Assigning a 3rd person value by default to underspecified [person] arguments, as in Harley & Ritter’s morphological hierarchy --or Béjar’s (2003) syntactic version, for that matter-- makes the wrong interpretive predictions even if an additional overruling mechanism is allowed. In order to capture the distinction observed, we first need a more articulated feature geometry system for [person] that incorporates Harley & Ritter’s unmarked values into the hierarchy (see, e.g., McGinnis 2005; also see D’Alessandro 2007:26ff for a similar discussion also involving *se/si* and a different solution):²⁸



The resulting feature specification of the relevant forms are distinguished in (45) (we also include a 1st person pronoun for illustration):



To complete the picture, there is independent evidence and general consensus that *se* lacks number feature specifications (Mendikoetxea 2012; Harris & Halle 2005 and references).²⁹ In sum, its lexical entry contains the feature information in (48):

²⁸ The place Harley & Ritter attribute to the animate/inanimate distinction as hierarchical subdivisions under the Class node must also be reconsidered accordingly; see O&R (2021) for discussion.

²⁹ In the Italian tradition, there is more discussion on the number features of *se*. Napoli 1976, Belletti 1982, Burzio 1986, among others, propose that *se* is singular while Chierchia 1995 argues that it is semantically

- (48) $se: \quad \varphi$
 |
 Person

With these ingredients, in the next section we present a formal analysis of *SECs* as regular sentences that derives all the properties discussed through the paper. That includes the agreement distribution in section 2 as well as the structural properties summarized in section 3.1 and developed in detail in O&R (2019a).

4. The derivation

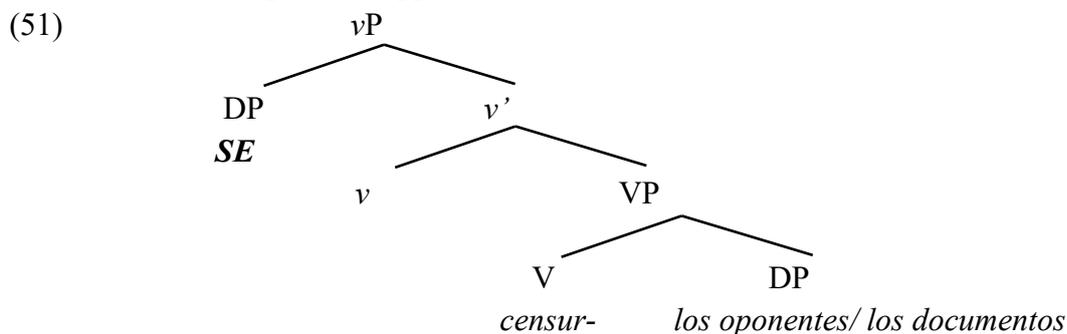
4.1. The syntax of *se-constructions*

From the point of view of the syntactic derivation, the initial quote heading this article, from Raposo & Uriagereka, may be extended to all non-paradigmatic *SECs*:

- (49) P-SECs and I-SECs are normal syntactic structures with no exotic properties.

That is, no selective Case-absorption or argument suppression operation is involved in P-SECs, and they both, P-SECs and I-SECs, follow the same derivation. Derivationally speaking that is the null hypothesis. First, *se* is introduced by external merge in the argument position determined by the predicate: the external argument position for transitive active sentences (50) and unergatives, and the internal argument position for unaccusatives and passive sentences. We illustrate the derivation with transitive structures because they cover all the relevant structural properties under discussion.

- (50) a. Se censuraron los documentos [=1a]
SE censored.**pl** the documents.**pl**
 ‘The documents were censored’
 b. Se censuró a los oponentes políticos [=1b]
SE censored.**sg** **DOM** the opponent political.**pl**
 ‘The political opponents were censored’

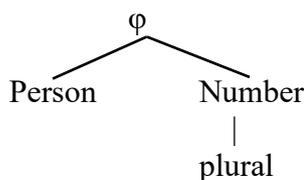


The difference in the derivation between [+DOM] and [-DOM] objects is the same difference observed in other transitive constructions, and it is therefore independent of the properties of *se*. We follow previous work arguing that DOM and non-DOM objects in

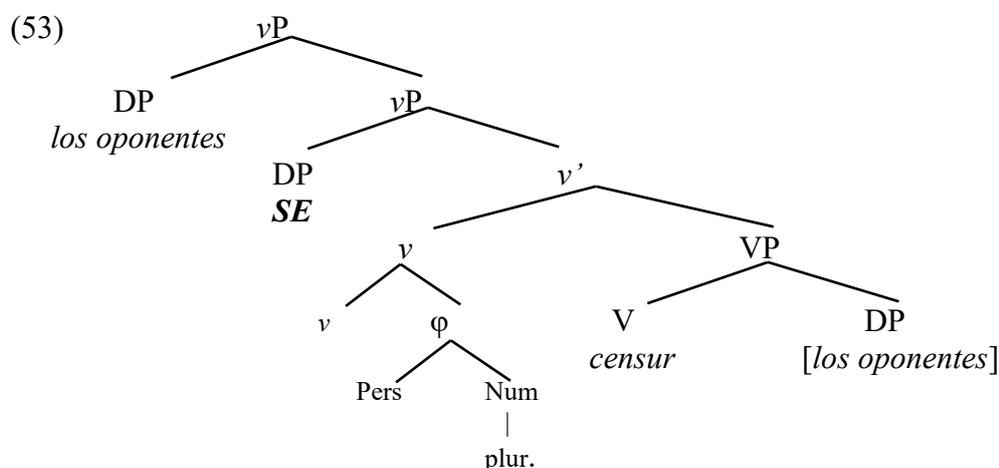
plural. See also D’Alessandro 2007 and references for discussion.

Spanish are structurally different (Torrego 1998, López 2012, O&R 2007, 2013a,b, 2016, 2019b and references there): object agreement is triggered only in the presence of an argument encoding person in the sense of Richards (2014) (see previous section). Thus, [-DOM] direct object DPs like *los documentos* ('the documents') in (50a) lack person features and do not agree with *v*. The *a* in [+DOM] DPs (50b), on the other hand, is a morphological reflex of a syntactic agreement/Case-checking relation in the *v*-domain.³⁰ The DP has the following ϕ -feature representation:

(52) (a) *los oponentes políticos* ('the political opponents'):



Simplifying things a bit, let us assume that the Agree relation is realized in (Spec, *v*P). We assume that *v* enters the derivation specified as [*uφ*] and it is valued via Agree, which copies the feature specifications of its controller, the DOM object (see section 4.3. for details, discussion and references on these assumptions). The resulting structure after the application of Agree between the probe *v* and the animate argument *los oponentes* is as in (53).



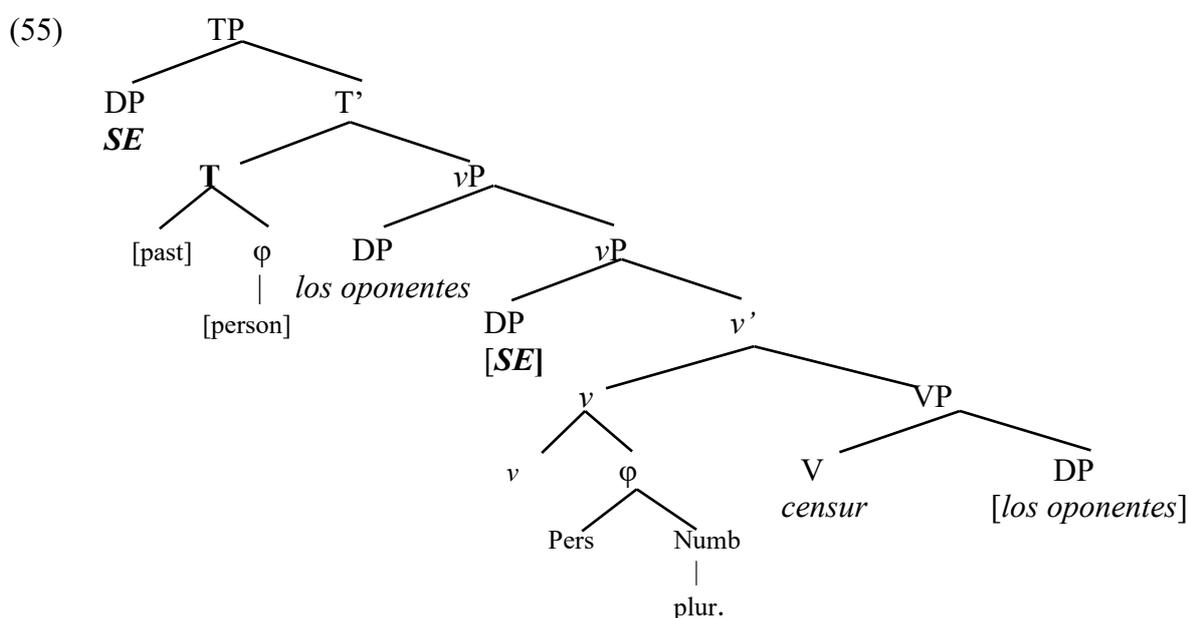
The analysis extends to first and second person objects, for which we have argued elsewhere that clitics are actually agreement markers and have person and number ϕ features.³¹

30 See O&R (2013a;b; 2021) for extensive discussion, including dialectal variation, and for some differences between Case and agreement relations; also see next footnote. The “inert nature” from the point of view of agreement of inanimate objects explains, for instance, why these elements may incorporate in polysynthetic languages, why they do not trigger PCC effects or why in Spanish 3rd person inanimate object clitics behave as determiners, and not like agreement markers (Roca 1996; Bleam 1999, O&R 2013a, Alcaraz 2021).

31 See O&R (2013a) and references. Our general analysis of SECs works equally well if 1st and 2nd person clitics are not agreement markers but pronominal clitics (e.g. Baker & Kramer 2018 and references). It does not depend on our specific analysis of DOM and dative doubling, and alternative analyses are compatible with it. The crucial factor that will make the difference in our derivation in section 5 is the fact that 1st and 2nd clitics, and also animate arguments, are specified for person while inanimate *lo(s)/la(s)* are not.

- (54) Se os/nos castiga (a los mismos de siempre) sin razón
 se 2Opl/1Opl punish.sg (DOM the same of always) without reason
 ‘You/We, the usual suspects, are punished without a reason’

At this point of the derivation (53), accusative Case has been assigned to the object, both in transitive P-SECs and in I-SECs. From that point on, the derivation is the same for (50a) and (50b). T is merged and triggers an agreement relation with *se*, as with any other regular subject. On the one hand, T enters the derivation specified as [$u\phi$] and it is valued by the subject, *se*, via Agree, which copies the feature specification of its controller (see section 4.3.). On the other hand, *se* moves to (Spec, TP) where it checks the EPP feature in T, and receives nominative Case:



This part of the derivation is the same independently of whether *se* is generated in the external argument position, or within VP in the position of internal arguments, as in unaccusatives and passives:

- (56) En este país, cuando se nace opositor, se es asesinado
 In this country, when SE born opposing, SE is assassinated
 'In this country, when you are born a dissident you are assassinated'

In sum, no *ad hoc* mechanism is postulated, and consequently, the syntactic distribution of SECs is correctly predicted to be the same as that of regular sentences where the subject has a more articulated ϕ -feature structure. The derivation is restricted exactly by the same general conditions constraining any other syntactic derivation and by nothing else, and their characteristic features are only determined by the specific properties of the lexical elements involved, as in any other case. The only remaining issue is the agreement pattern, to which we will return in section 5.

4.2. Clitic Left Dislocation and dialectal variation

Consider now the derivation of sentences containing a preverbal/null argument. Remember that, as discussed in section 2.2, this structure is subject to the dialectal variation illustrated in (57a-b).

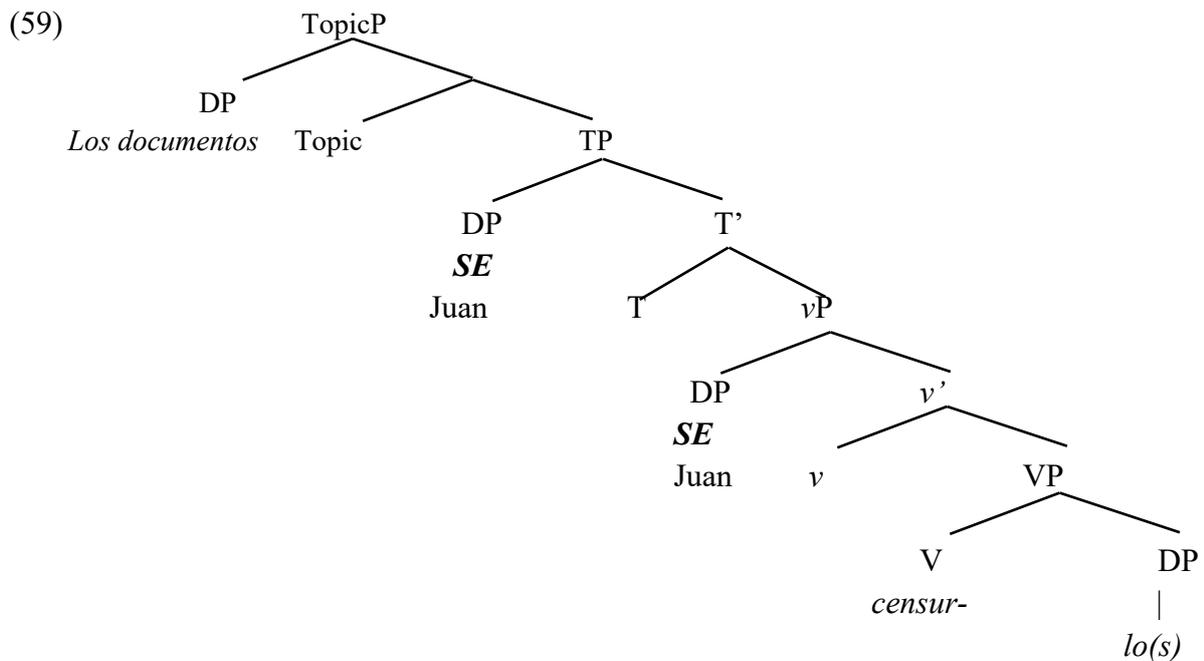
- (57) a. (Los libro-s) se vendieron
(The books.pl) SE sold.**pl.**
- b. (Los libros) se **los** vendió
(The books.pl) SE Cl.masc.**pl.** sold.**sg.**
'The books, they were sold'

(57b) corresponds broadly to Spanish dialects of South America from Argentina to Ecuador, while (57a) is regularly used in other dialects of Spanish with the qualifications discussed in section 2.2. Raposo & Uriagereka (1996) and Martins & Nunes (2016) argue that the preverbal argument in sentences like (57a) is not in subject position nor associated to it, but left dislocated.³² Considered from our 'active' approach to P-SECs, the structure of (58b) is exactly the same as the structure we find in CLLD in Spanish and other Romance languages (58a). We assume that the object DP is base-generated in a non-argumental position (van Riemsdijk 1997 and references therein),³³ and that D-clitics *lo(s)/la(s)* are merged in the internal argument position from where they cliticize to the verb (Uriagereka 1997, Roca 1996, O&R 2013a, a.o.). The derivation is represented in (59):

- (58) a. Los documentos (Juan) los censuró ____
The documents Juan Cl.masc.**pl** censored.**sg**
'The documents, John censored them'
- b. Los documentos se los censuró ____
The documents SE Cl.masc.**pl** censored.**sg**
'The documents somebody censored them'

32 They present a battery of arguments concerning Galaic-Portuguese, but many of their arguments extend to Spanish and Romance in general.

33 We represent the dislocated element in the specifier of TopicP, but nothing hinges on this decision as long as the pronoun cliticizes from the argumental position.



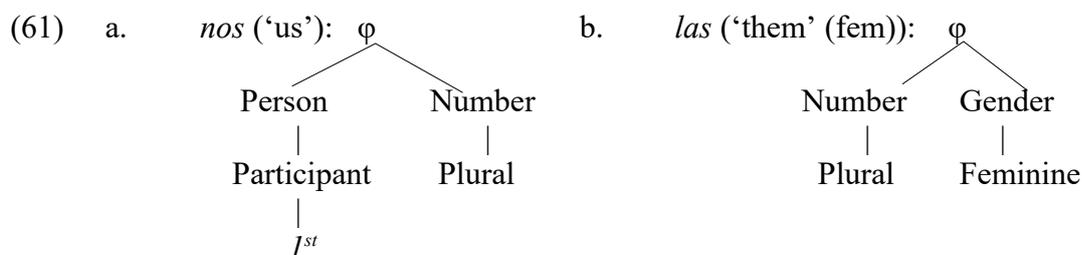
As we show in section 5.1, there are strong reasons to assume that (59) is the structural description for both (57b), with an overt clitic, and (57a), where the clitic mutates into number agreement. Except for the presence of the dislocated element, the derivation extends to cases where the argument is not overt: in our terms, these are null topic contexts where, as in (59), the DP *los* ('them') merges in the argumental position.

The same analysis automatically extends to 1st and 2nd person objects, and to animate ones. In this case, however, there is no variation: all dialects and speakers follow the object clitic strategy:

- (60) a. A vosotros (Juan) os censuró
DOM you (Juan) Cl.2.pl. censured.sg
 'As for you, John censured you'
- b. A vosotros se os censuró
DOM you SE Cl.2.pl. censured.sg
 'As for you, you were censured'
- c. *A vosotros se censurasteis
DOM you SE censured.2pl

We will come back to the ungrammaticality of (60c) and related issues in section 5.1, but at this point it is important to stress that the main difference between (57) and (60) is that 1st/2nd person pronouns are fully specified for person and number, while the clitic *lo(s)* has no person features (Roca 1996; O&R 2013a,b; Alcaraz 2021, and references; in (61) we abstract away from issues concerning the feature geometry of number and gender):

From this point on, the derivation continues as in (55), where the subject SE agrees with T.



4.3. Excursus on formal features, Agree and Match

In section 3 we concluded that *se* is a subject pronoun (section 3.1) specified for person and lacking number features (section 3.2). This is not the only context where the subject is not fully specified for person and/or number features (see section 5 for short discussion of other cases), and any approach to T-agreement must account for them. The Agree mechanism as developed in Chomsky (2000, 2001) is subject to a completeness condition on Match, applied under feature identity; however, contexts where the goal is not ϕ -complete and the derivation converges show that it cannot be a general condition for agreement. To illustrate, consider SECs: even if our analysis of P-SECs (62b) were not on the right track, I-SECs (62a) still constitute one case among a myriad of cases where the subject —be it *se*, a null pronoun, etc.— could never qualify to ‘completely’ satisfy T ϕ -features.

- (62) a. Se corre mejor en pista
 SE run.sg best on track
 ‘One runs best on track’
- b. Se censuraron los documentos [=1a]
SE censored.pl the documents.pl
 ‘The documents were censored’

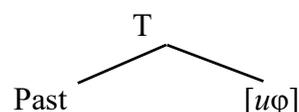
The mere existence of a default verbal agreement form in all impersonal contexts, where no possible number agreement-checking argument is available, may be taken as evidence that number agreement does not need to be valued. And the same extends to person features; see, e.g., Béjar’s (2006:148) analysis of Subject [3rd person]-Object [3rd person] combinations in Mordvinian. By the same token, given that (62) is grammatical, it would be a stipulation to impose a second cycle on (62b) for the derivation to converge and, in fact, we have seen in section 2 that that is not what happens. If there were a second cycle (Rezac 2003, Anagnostopoulou 2003, 2005, Béjar & Rezac 2009, among others), we would expect number features that remain unvalued in the first cycle to be available for Agree to be established with the object, showing a consistent pattern, contrary to facts. If, as we have argued, number agreement in transitive contexts is not syntactic, not only second cyclic agreement cannot be imposed over the derivation, but in fact, it must be prevented.³⁴

³⁴ This conclusion has important consequences for the treatment of agreement restrictions, especially those phenomena associated with the Object Agreement Constraint (the Person-Case Constraint and its extensions).

In more general terms, Match lacks explanatory power, as observed by Baker (2008: 44). While probably appropriate to feed the morphological component (see section 5.1), the idea of a set of pre-arranged, arbitrarily defined, feature slots to be valued by the goal looks like a conspiratorial prerequisite that fails to capture the systematic properties of agreement in natural language. It is, in some sense, a very short-sighted “look ahead” from the lexicon into the computation.

Notice, in addition, that when combined with a feature hierarchy system where features do not have intrinsic values but depend on their place in the geometry, the valuation operation as established standardly is hard to codify from a strictly formal point of view. As Béjar (2003, 2008) observes, each node in the hierarchy entails the presence of all the superordinate nodes, so that when a pronoun is specified for, say, [1st person], it is entailed by the hierarchical organization that it is also specified for [participant], etc. up to the root of the hierarchy. Consequently, given those entailment relations and assuming that the feature hierarchy of uninterpretable features in T is identical to its interpretable ones in pronouns (the null hypothesis), a [*u*person] probe would be general enough to target all goals that have person features. In particular, if we follow our proposal in section 3, bearing person features would include *se* [person], as well as [3rd person] and all [participant] goals. Generalizing even further, the simplest system would be one where the probe in T is specified as [*u* ϕ]:

(63) T as a syntactic probe [before Agree]:



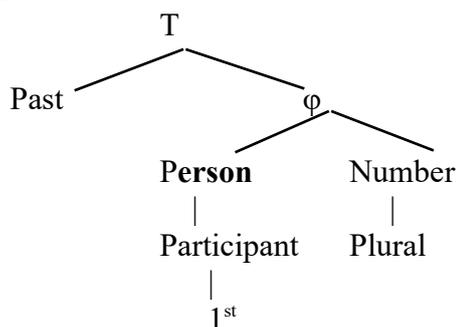
We then may assume that the Agree relation copies the entire hierarchical structure of the goal (the “controller” of the agreement relation) into the probe (the “target”).³⁵ Since T ϕ -features are uninterpretable, the idea that Match copies whatever the controller brings to the picture is both suitable for its morphological analysis and in compliance with standard views on agreement in natural languages. The resulting complex head will be different depending on the set of features the subject contributes to the agreement relation, as desired:

See O&R 2007, 2013b, 2019b for some evidence that facts similar to those described here are also present in PCC contexts.

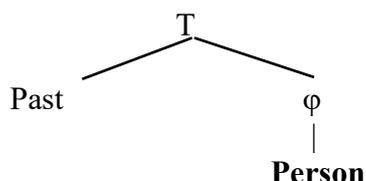
35 The result is close to Baker’s (2008: 44) desideratum, which we fully share, for a design where:

“all Fs are potential agreeers and they agree with whatever features they can find in their environment according to structural principles. If agreeing functional heads are not prespecified as agreeing in particular features, it follows that there cannot be any matching condition of the kind Chomsky envisions. This condition should become superfluous – at least if agreement in ϕ -features is the only relevant kind of agreement in natural languages.”

- (64) Result after application of Agree:
 a. with *nosotros* ('we'):



- b. with *se*:



Under these natural assumptions, the agreement process in SECs is by all criteria a regular one. Consequently, T's [$u\phi$] has been completely valued in the syntax. However, in the morphology the resulting matrix of features in T has the number slot empty, since it cannot be copied from the feature matrix of the agreeing subject *se* whose lexical specification lacks the corresponding interpretable feature. In other words, the agreement problem is not syntactic, but morphological.

In the next section we analyze the agreement behavior of SECs. Given the discussion so far, the asymmetries between postverbal and preverbal positions correspond syntactically to the difference between *in situ* and clitic left dislocated objects in regular transitive clauses. As we discuss next (section 5.2), the postverbal DP pattern is not very different from what we see in other contexts. On the other hand, the clitic's behavior is closer to some morphological agreement displacement processes known in multiple agreement/clitic clusters of various languages. From the discussion so far it should be clear by now that none of these effects may occur in narrow syntax but, as argued, they are determined by it.

5. Post-syntactic processes

In section 2 we showed that agreement patterns behave differently depending on the position of the object. With preverbal topics, its morphological manifestation is obligatory, either as an object clitic or as subject agreement. In contrast, with postverbal ones the only morphological device used is subject number agreement, but it is established in a semi-random fashion. Since their behavior is sharply different, we conclude that each one follows a different path. For preverbal/null topics the A-position is occupied by a clitic (section 4.2), which in some dialects mutates post-syntactically into subject agreement (section 5.1). For

postverbal objects, we propose that the verb may show up marked with a default singular number or it may harmonize post-syntactically with a plural element following it (section 5.2).

5.1 Dialectal variation in the morphological component: Clitic Mutation

The pronoun *se* is the only subject clitic in Spanish. It cliticizes into T just in the same way French *il* ('he') or impersonal *on* ('somebody'/'you' indefinite) do. Similarly, the determiner head *lo(s)/la(s)* ('it/him/her/them') in object position cliticizes into the verbal complex. The result of the computational component's inner workings is a morphological word that includes the two clitics, the verbal root and all the T features.

(65) *se* + *los* + *censur* + ind.past.person
SE + Cl.masc.pl + *censor* + ind.past.person
'They were censored'

Descriptively speaking, the presence of the object clitic in C-varieties is the straight manifestation of the structure in (65), while in A varieties the clitic is spelled out as subject number agreement. Subject agreement is therefore an alternative morphological expression of the very same structural description that is expressed in C varieties by the clitic. We have termed this particular morphophonological process *clitic mutation*, but we consider it a particular instance of a family of similar agreement effects including several clitic cluster phenomena, agreement displacement, and eccentric or omnivorous agreement (Bonet 1991, Hale 2001, Rezac 2008, a.o.). In other words, in this clitic/number agreement alternation there is no passive SE construction, but a regular active structure whose object clitic is morphologically camouflaged as subject agreement.

In what follows, we first present direct evidence supporting this analysis (5.1.1); second, we detail the conditions under which clitic mutation will apply in the morphological component and those where it will be blocked (5.1.2). To finish, we provide independent crosslinguistic evidence in favor of our proposal, which suggests that the morphological process itself belongs to a well known family of morphological processes (5.1.3).

5.1.1 Evidence for clitic mutation

There are three different contexts which constitute strong evidence for the proposal that in these contexts subject number agreement is a mutated clitic. Consider first idiomatic expressions containing a non-referential clitic, such as *liarla parda* ('to make a complete mess'), *cagarla* ('to fuck it up'), *matarlas callando* ('to go about things slyly'), *pasarlas canutas* ('to have a rough time'), *verlas venir* ('to see them coming'), etc. The clitic is part

of the idiomatic expression and must be present under all circumstances (66) (García-Page 2010 and references). As expected, passivization is completely impossible (67):

- (66) a. En esos pueblos, siempre *(la) liamos parda
 In those villages, always Clfm.sg mishandle.we brown.sg
 ‘In those villages we always make a complete mess’
- b. En la guerra, siempre *(las) pasamos canutas
 During the war, always Clfm.pl pass.we canutas
 ‘During the war we always have a rough time’
- (67) a. * En esos pueblos, siempre es liada parda (por nosotros)
 In those villages, always is mishandled.fm.sg brown.sg (by us)
 ‘In those villages we always make a complete mess’
- b. * En la guerra, siempre son pasadas canutas (por nosotros)
 During the war, always Clfm.pl pass.we canutas
 ‘During the war we always have a rough time’

Non-paradigmatic SECs are the only exception where the clitic disappears. In those cases, which sound absolutely natural, if the clitic of the idiom is singular the verbal form shows up in singular (68a), and if plural, it is obligatorily plural (68b).

- (68) a. En esas fiestas, siempre se (*la) lia parda
 In those villages, always SE Clfm.sg mishandle.sg brown.sg
 ‘In those villages people always make a complete mess’
- b. En la guerra, siempre se (*las) pasan canutas
 During the war, always SE Clfm.pl pass.pl canutas.pl
 ‘During the war people always have a rough time’

A second piece of evidence comes from Clitic Climbing contexts. In transitive structures (69), including SECs (70), all dialects of Spanish show enclisis to the embedded infinitive when the clitic appears downstairs (69a) (70a). In regular transitive contexts, when the clitic climbs and reaches the finite verb, it appears as a proclitic to the finite verb (69b). However, in SECs there is a split: in C-varieties the clitic may be preserved also in that position (70b), while in A-varieties it shows up as subject agreement (70c), in contrast to (69c).

- (69) a. intentó censurar**las**
 SE tried.sg to.censor.Cl**fem.pl.**
- b. **las** intentó censurar
 SE Cl**fem.pl** tried.sg to.censor
- c. *intentaron censurar
 SE tried.**pl** to.censor
 ‘Somebody or other tried to censor them’
- (70) a. se intentó censurar**las** *All dialects*
 SE tried.sg to.censor.Cl**fem.pl.**
- b. se **las** intentó censurar *C-varieties*
 SE Cl**fem.pl** tried.sg to.censor
- c. se intentaron censurar *A-varieties*

SE tried.**pl** to.censor
 ‘Somebody or other tried to censor them’

The parallelism is maintained across the board. In long distance ClClimb, A-varieties retain the clitic in all the intermediate positions (71a-c), and mutate it into agreement only when it reaches the *se* clause (71d), exactly what clitic mutation predicts:³⁶

- (71) a. Se tiene que empezar a poder entender **.lo.s**
 SE must.**sg** that begin.infin to can.infin understand.infin.**Clmasc.pl**
- b. se tiene que empezar a poder. **lo.s** entender
 SE must.**sg** that start.infin to can.**Clmasc.pl** understand.infin
- c. ??se tiene que empezar. **lo.s** a poder entender³⁷
 SE must.**sg** that start.infin.**Clmasc.pl** to can.infin understand.infin
- d. se tienen que empezar a poder entender
 SE must.**pl** that start.infin to can.infin understand.infin

Finally, clitic mutation is also subject to clitic clustering effects: When more than one clitic is generated in the embedded domain, the entire cluster must climb together (72a,b) vs. (72c,d). As expected under Clitic Mutation, subject agreement behaves in this respect as a clitic (73). In contrast, if it were true subject agreement, the ungrammaticality of (73c) would be completely unexpected, since subject agreement does not interfere with clitic clusters in other contexts (72a,b):

- (72) a. Intentaron censur**artelas**
 tried.they to.censor.Cl2sg.Clfm.pl
- b. **Te las** intentaron censurar
 Cl2sg Clfm.pl tried.they to.censor

36 This subject agreement is subject to all the constraints clitic climbing shows in parallel contexts: adjunct wh-islands (i)-(ii), intervention effects with negative heads, etc.

- | | | | | |
|------|----|---|----|--|
| (i) | a. | Sabemos leer las
know.we to.read.them | b. | Las sabemos leer
Them know.we to.read |
| | c. | Sabemos cómo leer las
know.we how to.read.them
‘We know how to read them’ | d. | * Las sabemos cómo leer
Them know.we how to.read |
| (ii) | a. | Se sabe leer las
SE know. sg to.read.them | b. | Se saben leer
SE know. pl to.read |
| | c. | se sabe cómo leer las
SE know.sg how to.read.them
‘It is known how to read them’ | d. | *Se saben cómo leer
SE know.pl how to.read |

37 Example (71c) sounds a bit odd to us, and our informants have the same impression. Note however, that the corresponding sentence with a subject fully specified for person sounds as marginal:

- (i) a. ??Tenemos que empezar.**los** a poder entender
 ‘We must begin to be able to understand them’

Whatever the reason for this marginality is, paradigms with *se* and with fully specified subjects behave in the same way even in this respect.

- c. ***Te** intentaron censurar**las**
Cl2sg tried.they to.censor.Clfm.pl
 - d. ***Las** intentaron censurarte
Clfm.pl tried.they to.censor.Cl2sg
'They tried to censor them to you'
- (73)
- a. Se intentó censurar**telas**
SE tried.sg to.censor.Cl2sg.Clfm.pl
 - b. Se **te** intentaron censurar
SE Cl2sg tried.pl to.censor
 - c. *Se intentaron censurarte
SE tried.pl to.censor.Cl2sg
 - d. *Se **te** intentó censurar**las**
SE Cl2sg tried.sg to.censor.Clfm.pl
'Someone tried to censor them to you'

We know of no previous discussion of these facts in the literature, but we see no possible explanation for them in traditional terms. In contrast, this is precisely what the application of Clitic Mutation predicts, since subject number agreement is structurally an object clitic, and it only surfaces as number agreement after morphological manipulation.

5.1.2 Clitic mutation and its domain of application

The purpose of this section is to make explicit the properties of the interface syntax-morphology. By no means do we pretend to lay out a morphological analysis in such an intricate area as clitic-agreement interactions, but thoroughly describe the conditions imposed by syntax to Clitic Mutation. According to our analysis in sections 3-4, syntax supplies Morphology with a Tense head valued for a [person] feature --via Agree with *se--*, but with no number features specified. Since the morphological exponent of the T head in Spanish encodes number, morphology needs to assign it a value. A priori, there are two different options: (i) the number slot may be assigned a default (singular) value; or (ii) it fulfils the morphological requirement by **blending** some element into the morphological word to provide a value for it. C Varieties have both options generally available. In (74a) object features are realized as a clitic and subject agreement is set to a default third person, while in (74b) object clitic features **blend** with subject agreement. In A varieties, only the second option, Clitic Mutation, is available.

- (74) a. (Los documentos) se los censuró > b. se censuraron
DP_{Topic}.pl SE Cl.3O_{masc}.pl V.past.sg > SE V.pl
'The documents were censored'

Following Distributed Morphology postulates, for instance, we can assume that in SECs an

impoverishment rule is applied in order to provide a value for number to subject agreement. The impoverishment rule has the effect of removing any conflicting feature --in this case, gender and accusative-- paving the way to Clitic Mutation. As a matter of fact, a strikingly similar effect is found in the interaction of clitics. Bonet (1991; 1994) observes that in many clitic clusters in Romance, the output is often opaque with respect to the information encoded by the same clitics in isolation. According to her analysis, adopted in DM, in those contexts the clitics may suffer an impoverishment morphological process, as a consequence of which their representation corresponds to the PF realization of another clitic.

Consider now the distribution of the clitic/agreement strategies for pronominal (75) and, in general, for [+DOM] objects (76), vs. [-DOM] objects:

- (75) a. A vosotros (Juan) *(os) censuró [= (60a-c)]
DOM you (Juan) Cl.2.pl. censored.sg
 ‘As for you, John censored you’
- b. A vosotros se *(os) censuró
DOM you SE Cl.2.pl. censored.sg
 ‘As for you, you were censored’
- c. *A vosotros se censurasteis [_{cl} A vosotros se os censuró]
DOM you SE censored. 2.pl
- (76) a. A las oponentes políticas (Juan) *(las) censuró
DOM DPpl. (Juan) Cl.ac.pl. V.sg
- b. A las oponentes políticas se *(las) censuró [= (32a)]
DOM DPpl. SE Cl.ac.pl. V.sg
 ‘The political opponents were censored’
- c. *A las oponentes políticas se censuraron [= (32b)]
DOM DPpl. SE Cl.ac.pl. V.pl
 ‘The political opponents were censored’

Remember that, unlike number, person agreement has been valued between *se* and T (section 4.3). Thus, descriptively speaking, the difference between these contexts and the previous cases where clitic mutation applies is that all clitics in (75)-(76) contain [person] specifications. We may thus assume that these features conflict with the [person] features in T and the clitic cannot mutate. That includes 1st and 2nd person clitics as well as [+DOM] ones (see section 4)³⁸ In more general terms, clitic mutation will then be possible when the

38 Two reviewers raise the question of why many non-*leísta* speakers make use of the clitic *le* in these particular contexts. This generalized use of the clitic *le* poses a problem to virtually all approaches, since it specifically arises in I-SECs, commonly analyzed as regular transitive structures since Cinque (1988). We do not have a full explanation for this fact, but within the logic of our proposal it is most probably linked to the blocking effects on clitic mutation compelled by clitics specified for person features. In this specific context there is evidence that *le* has not the same syntax than in other transitive clauses. Consider the following sentences:

lack of number features in the subject and the lack of person features in the object cooccur. On the contrary, when T agreement fully specifies person and number features, the clitic has no room in T's resulting morphological exponent to mutate into subject number agreement, not even with *lo(s)/la(s)* (75c).

With Clitic Mutation, Spanish morphology deals with the fact that subject agreement is morphologically mandatory at the expenses of object clitics. As we discuss next, with some interesting differences, all these are properties that clitic mutation shares with other morphological phenomena crosslinguistically.

5.1.3. Extensions: clitic mutation in a broader morphological context

Crosslinguistically, Clitic Mutation is not an isolated phenomenon. It may be compared to the family of effects known as *agreement displacement* in the morphology of many languages. To illustrate, in Basque, when the verbal form fulfills certain conditions, ergative agreement takes the form and position of absolutive agreement. Descriptively speaking, in order for agreement displacement to take place, a necessary condition is for the absolutive argument to be third person, an agreement that happens to lack a morphological exponent (77c,d) vs. (77a,b).

- (77) a. Nik zu maite **z**- intu -da -n
 I.erg you.abs love **2abs**-aux-1erg-past
 'I loved you'
- b. Zuk ni maite **n**- indu-zu -n
 You.erg I.abs love **1abs**-aux-2erg-past
 'You loved me'
- c. Zuk bizitza maite **z**- enu -en (vs. *u -**zu**.n)
 You.erg life.abs love **2erg**-aux-past (aux-**2erg**-past)
 'You loved life'
- d. Nik bizitza maite **n**- u -en (vs. *u-**da**-n)
 I.erg life.abs love **1erg**-aux-past aux-**1erg**-past
 'I loved life'

Most analyses interpret that distribution as indicating that the absolutive agreement relation is defective for [-participant], but we have extensively argued that no agreement relation holds between the verbal complex and third person objects in many languages including Basque (O&R 2007). If correct, the ergative element colonizes the absolutive agreement slot

- | | | | |
|-----|--|------|---|
| (i) | *A nadie le viste
DOM anybody 3Osg saw.you
'You did not see anybody' | (ii) | A nadie se le vio
DOM anybody SE 3Osg saw
'No one was seen' |
|-----|--|------|---|

Negative quantifiers cannot be clitic-doubled in Spanish unless the clitic is the morphological exponent of a syntactic agreement node (O&R 2013 and references therein). The contrast between (i) and (ii) suggests that the clitic is realizing different features in each sentence: While in (i) it is a pronominal clitic, in (ii) it is an agreement head.

in (77c,d) when that slot lacks person features altogether. That description of the phenomenon is not very different from diachronic accounts (Gómez & Sáinz 1995, Lakarra 2005, Ariztimuño 2017), where the subject pronoun is argued to occupy the first position precisely in those cases where no object pronoun may occupy it (historically Basque lacks third person pronouns).

As is the case of virtually all proposals concerning the agreement facts in *SECs*, current analyses of agreement displacement phenomena mostly lay out syntactic solutions (Ortiz de Urbina 1989, Rezac 2008, Béjar & Rezac 2009, a.o.). However, all these phenomena present similar contextual restrictions, suggesting a similar morphological analysis. In addition, in the case of ergative displacement, there is a long tradition of morphology-based explanations (Laka 1993, Albizu & Eguren 2000, Fernández & Albizu 2000, Arregi & Nevins 2012, a.o.).

The main difference between Basque ergative displacement and Spanish Clitic Mutation is that while in the later person agreement has been checked between T and *se*, in the case of Basque there is no person agreement relation in *v*. That explains why the entire ϕ features of the ergative can take over the object agreement slot in Basque morphology while only number is involved in Spanish *SECs* and no clitic mutation applies with animate clitics (1st, 2nd and [+DOM]) in the last case. The differences then follow from the different morphological specifications of the host exponents in each case, as desired.

5.2 A postsyntactic phenomenon (II): Number Harmony

Agreement in *SECs* with postverbal DPs exhibits properties that radically differ from contexts with preverbal or null ones, and furthermore its structural description does not include an object clitic, and in consequence, it cannot be an instance of Clitic Mutation. In this section we argue that this agreement is the reflex of a post-PF procedure that we call *Number Harmony* (NH) (see Demonte & Pérez Jiménez (2012) for similar observations regarding number agreement between adjectives and conjunct NPs; see also Benmamoun, Bhatia & Polinsky (2009)).

5.2.1. Conditions for NH

The only aspect in which syntax plays a role in Number Harmony is that the derivation results in a structure where the set of ϕ -features in T includes a [person] value supplied by its agreement with *se* but no [number], and that there is a DP closeby in postverbal position (78). But, in contrast to clitic mutation contexts, the V-object relation is not represented in V's morphological word.

- (78) se + censur + **past.indic.person** los documentos
 SE + censor + **past.indic.person** the documents
 ‘The documents were censored’

Verbal agreement follows two different paths in that context: either it takes a default singular value (79a) or it adopts the value of the closest nominal in the sentence, exhibiting plural harmony (79b). To complete the paradigm, when the closest nominal is singular, as expected, the verb never ever shows up in plural (79c).

- (79) a. Se censuró los documentos
 SE censored.sg the.pl documents
 b. Se censuraron los documentos
 SE censored.pl the.pl documents
 c. * Se censuraron el documento
 SE censored.pl the.sg document

In all cases, “closest” is computed in pure linear proximity terms, as shown by the battery of long distance cases discussed in section 2.1.

In sum, the distribution of agreement in these contexts does not show any of the properties of a well-behaved syntactic checking relation. It does not seem to have a morphological motivation either: there is no possible source for the plural morpheme within the morphological word in (79) and, unlike in Clitic Mutation contexts, the agreement patterns in unpredictable ways. The same conclusion may be attained concerning PF. In particular, the source of the plural number exceeds by far the domain of regular phonological processes, as clearly shown by hyper-raising cases from section 2.1, repeated in (80):

- (80) Se consideran que hay personas con categorías superiores e inferiores
 SE consider.**pl** that there.are persons with category superior and inferior
 ‘It is considered that there are people superior to others’

In (80), the nominal triggering plural marking in the matrix verb is the argument of the embedded existential sentence, far away from any conceivable prosodic unit with it.

As far as we can see, that leaves some mechanism external to the grammar as the only option. In other words, agreement patterns are determined by extra-grammatical factors. For speakers like us, who invariably resort to number harmony with [-DOM] direct objects (P-SECs) and to a singular default in the rest of the cases (I-SECs), we must assume that this additional regularization process is a learned strategy imposed by means of socio-educative pressure.³⁹

In the next section, we briefly review other instances of subjects lacking number

39 That predicts that variation should be more visible in pre-school children even in highly educated environments, a prediction that we cannot test at this point.

specification in Spanish that follow a similar pattern. We show that in some cases, number agreement works essentially in the same way as in SECs, while in other cases it is fully determined, sometimes in completely unexpected ways.

5.2.2. Other agreement patterns that grammar cannot deal with

5.2.2.1 Existential *haber*. In Standard Spanish agreement between the verb and the nominal phrase is not recommended in these contexts (NGLE 41.6b), and for many speakers the only possible combination is (81a), where the verb appears in a default third person form. However, the use of plural agreement, where the verbal form reflects the number of its nominal argument (81b), is increasing considerably, and in some areas it is becoming general.

- (81) a. Había tres sillas en el porche
There.was three chairs in the porch
- b. Habían tres sillas en el porche
There.were three chairs in the porch
'There were three chairs in the porch'

This situation is remarkably similar to SECs, including a high degree of dialect-internal and idiolectal variation. An additional particularity of this case is that while the present form, *hay*, is a specific form for the impersonal --as opposed to 3rd person *ha--* and does not have plural, the past form is the regular 3rd person singular *había*, which does have its corresponding plural form *habían*. Consequently, number harmony is often obtained in the past but almost never shows up in the present, reinforcing the conclusion that we are not dealing with well behaved syntactic or morphological agreement (for details and dialectal characterization see DeMello 1991; Grácia & Roca 2017; O&R 2021).

5.2.2.2. *Hacer*. A similar case is found with impersonal uses of *hacer* (e.g. *hace calor* = 'it is hot'). As with *haber*, in Standard Spanish it appears always in a default 3rd person singular (82a), but when the complement includes a measure phrase, for some speakers it often triggers number agreement in the verb (82b):

- (82) a. Hace frío/treinta grados (fuera)/varios días
Make.sg cold/thirty degrees outside/several days
'It is cold', 'it is thirty degrees', 'it was several days ago'
- b. Vivir sin calefacción cuando hacen dos grados bajo cero es inaceptable
to.live without hitting when make.pl two degrees.pl under zero is unacceptable
'For people to leave without hitting when it is -2° Celsius is unacceptable'

While with *haber* number harmony is quite common, with *hacer* is more marginal, but as with SECs and *haber*, its use is idiolectal, probably subject to cultural pressure.

5.2.2.3. *Neuter subjects*. The situation is different in copulative sentences whose subject is the neuter deictic pronoun *eso* or a propositional subject. Consider the following sentence:

- (83) Luis dice que eso son tonterías pero para mí no lo son
 Luis says that that.neut are nuts but for me not clitic are
 ‘Luis says that those things are nuts, but they are not for me’

In (83) the deictic pronoun in subject position does not encode number, and it is the predicate *tonterías* that determines agreement. In this case, agreement is obligatory, and default singular is rejected. Similarly, with propositional subjects, agreement is determined by the predicate.

- (84) que no lleguen tarde o heridos son preocupaciones de padre (pero para mí no lo son)
 that not arrived late or hurt are worries of father (but for me not Clitic are)
 ‘That they arrive late or hurt are worries of a father (but they are not for me)’

Note that in these cases when the predicate is substituted by the invariant clitic *lo*, plural agreement is still obligatory (*para mí lo son*).⁴⁰

5.2.2.4. *A-bar movement*. When the complement of V A-bar moves to a preposed position, clitic doubling is not possible (85).

- (85) a. ¿Qué documentos (*los) censuraron?
 which documents 3Opl censored.they
 Which documents did they censor?
 b. Los documentos que (*los) censuraron fueron publicados
 The documents that 3Opl censored.they were published
 ‘The documents they censored were published’

In SECs, these structures exhibit the lack of systematicity typical of Number Harmony (86). As a matter of fact, this is the only case where even speakers that otherwise always adhere to the norm show some degree of variation.

- (86) a. ¿Qué documentos se censuró/ censuraron?
 Which documents SE censored.sg / censored.pl
 ‘Which documents were censored?’
 b. Los documentos que se censuró/ censuraron fueron publicados
 The documents that SE censored.sg / censored.pl were published
 ‘The documents they censored were publish’

This exceptional behavior among preverbal arguments is expected in our analysis: while arguments in clitic left dislocation structures are base generated in its left periphery position,

40 Now and then other cases are found. For instance, the verb *oler* ‘to smell’ allows an impersonal construction, as in (i). In this case the noun in the PP may trigger agreement:

- (i) Aquí huele/huelen a celos
 Here smells/smell to jealous
 There is a jealous scent

the tail of the A'-chain in (86) generates in the VP-internal object position. Clitic doubling is thus not possible and, consequently, clitic mutation lacks the structural conditions to apply (73). In contrast, this A'-chain configuration is a potential target for agreement harmony to apply, and the lack of systematicity reappears: (87a) illustrates the lack of number agreement with plural objects, (87b) the case of plural agreement with DOM objects, and (87c) long distance number agreement.⁴¹

- (87) a. [...] *los servicios que se proporciona* a los alumnos [Spain]
the services.**pl.** that SE provide.**sg** to the students
'The services that were provided to students'
- b. Selección de *las personas a las que se entrevistaron*... [Argentina]
Selection of the people.**pl** DOM the that se interviewed.**pl.**
'Selection of the people tha were interviewed'
- c. [...] *las comunicaciones realizadas o que se prevén hacer* [Cost Rica]
the communications.**pl** realized or that SE expect.**pl.** do
'The comunicaciones realized or those that were expected to be realized'

In sum, different pieces of evidence converge and supports our proposal that subject number agreement in SECs conceals two different phenomena with very distinctive properties, both occurring post-syntactically:

Clitic mutation is a rather regular and systematic morphological process by means of which an argumental D-clitic shows up morphologically as number agreement. It coexists with a clitic strategy whose distribution is also systematic and predictable both dialectally and in terms of the configurations where they apply. The dialectal cut lies in whether this morphological rule/process is optionally applied, as in C-varieties, or obligatory, as in A-varieties.

Number harmony, on the other hand, is a less systematic extra-grammatical phenomenon, with a very vague distribution. Agreement harmony in SECs shares many properties with similar agreement "failures" discussed in section 5.2.2. Moreover, its distribution seems to be largely determined by sociolinguistic and stylistic factors such as the degree of penetration of the Academies' norm, social acceptability, etc. that escape the domain of grammar.

6. A few notes on cross-linguistic variation

An LI reviewer requires that we address the issue of cross-Romance variation within non-

⁴¹ Effects of this type have often been reported in research on agreement intervention in language processing (Zawiszewski et al. 2016, Villata 2017, Mancini et al. 2016), and a more general post-syntactic analysis of a large variety of number agreement phenomena in quirky agreement configurations seems to us worth pursuing. However a detailed analysis of Number Harmony exceeds our area of expertise and goes way beyond our goals in this paper.

paradigmatic SECs. As our proposal about Spanish illustrates, such a task requires a very detailed and fine-grained analysis before deciding how each specific language fits into the picture. Quite often the details needed have not even been considered in previous work, and an in-depth cross-Romance analysis is out of the scope of this paper. Given all that, what follows is by force highly speculative.

An important contribution of Cinque's (1988) seminal proposal to the analysis of *se/si* is that it links the differences between P-SECs and I-SECs to the lexical properties of the different *se* items involved. We have argued that for Spanish that line of analysis is not empirically adequate. However, putting the burden of the explanation on the lexical properties seems the right strategy to account for the differences between non-paradigmatic SECs and other constructions language-internally. As shown, the effects of a minimal lexical distinction may be considerable at different levels of the grammar and beyond. Similarly, divergences in the diachronic evolution of the relevant lexical items may be responsible for important synchronic differences in the Romance family. Notice also that if an important source of crosslinguistic variation is in the lexical diversification of the properties of *se*, *a priori* we will not expect to be able to determine the direction of the changes, in the same way researchers cannot determine which direction the different variants of SARS-CoV-2 will take after several mutations take place. Consequently, we have no expectations about what the output will be in a given system. However, what we could expect of such changes is for the resulting lexical item to present internal coherence relative to the general properties of syntactic derivations, as we argued for *se* in Spanish, with clear empirical predictions about its syntactic and morphological behavior.

Diachronic diversification of Latin $SĒ$'s lexical attributes and cliticization patterns are not uniform in Romance. In the case of Iberian Romance languages, although some authors (e.g. Maddox 2021) present a unified approach to the evolution of all *se*-s, some facts suggest that there is a split at least since the XVI century (see Monge 1954 for Spanish; Martins 2005 for Portuguese). If that is correct, as the Latin reflexive $SĒ$ evolved, a variety of lexical *se* pronouns with particular lexical specifications were created in the different Romance languages, resulting in different subsystems both language-internally and crosslinguistically. Taking Spanish as an arbitrary reference point, different possibilities arise with respect to the evolution of non-paradigmatic *se* in Romance:

(i) Some Romance languages may have taken a completely different path, most characteristically French and Brazilian Portuguese, which have developed alternative impersonal structures with *on* (see, Wolfgruber 2017 and references), and *a gente* (Nunes xxx and references) respectively.

(ii) Rumanian might illustrate the case of a language that presents a more conservative evolution (see below).

(iii) Other languages might have gone further in the division of lexical *se* items, yielding further structural differences. The different position impersonal *se* and passive *se* occupy in clitic clusters of Borgomanero (Manzini & Savoia, 2005; Cennammo, 2014), suggested by an anonymous reviewer, might indicate that number harmony has grammaticalized yielding different lexical *se*-s.⁴²

With these cautions in mind, consider the case of Italian. According to D'Alessandro (2007) "there [is a] case distinction that the sole overt DP manifests when it is a pronoun, as a function of whether or not it agrees with the verb", which support an approach distinguishing I-SECs from P-SECs (examples and glosses from D'Alessandro 2004).

- (88) a. In Italia essi/ *li si mangiano
In Italy they-nom 3rd pl them-acc 3rd pl. si eat-3rd pl
- b. In Italia li/ *essi si mangia
In Italy theM-acc 3rd pl they-nom 3rd pl. si eat-3rd sg
'In Italy they eat it'

In principle, this state of affairs does not contradict our approach as long as the existence of two lexically distinct *si*-s in Italian can be independently motivated. As a matter of fact, this is precisely what D'Alessandro tries to do on the basis of some aspectual contrasts. Note, further, that there are at least two important differences between Italian and Spanish that most probably are the source of additional differences. The first one is DOM, a process absent in Standard Italian and pervasive in Spanish syntax where, if our analysis is on the right track, animate objects induce person-agreement in *v* and, consequently, block clitic mutation. The second property is that in Italian the default form for number is plural (see footnote xx and references).

Putting those differences aside, a unified analysis of P-SECs and I-SECs is worth exploring also for Italian. D'Alessandro's explanation of the paradigm in (88) conceals a hidden assumption that weakens her argument: She assumes that the preverbal pronoun is in an A-position, in which case it must be nominative, and that, consequently, there is a Case difference between the nominative *essi* and the accusative *li* in (88). However, unlike first and second pronouns, third person strong pronouns in Italian do not make a morphological distinction between nominative and accusative forms and, *essi* in (88) might be either a subject or a preverbal object pronoun.

42 However, apart from the need of a much more detailed description of Borgomanero along the lines described in this paper, a careful reading of Manzini & Savoia's discussion advises against going that far in the syntactic analysis of dialectal differences in the internal ordering of clitic clusters.

	Subj Sg	Obj Sg	Plural
1st pers.	io (subj.)	me (obj.)	noi
2nd pers.	tu (subj.)	te (obj.)	voi
3rd pers.	esso (masc.), lui (anim.) essa (fem.), lei (anim.)		essi (masc.) esse (fem.)

Thus, given our analysis, the distribution of *essi* and *li* in (88) may be reinterpreted as cases of Clitic Left Dislocation and null topic contexts, as in Spanish, with additional clitic mutation in (88a). If so, the entire paradigm follows: In (890a) (D'Alessandro's examples with our glosses), *li* cannot be a dislocated object, because clitics are not allowed in dislocated positions; it cannot be an object clitic either, because agreement indicates that clitic mutation has taken place. In reverse, the lack of plural agreement indicates that clitic mutation has not applied in (89b), so the presence of the clitic is obligatory with null topics, and there is no possible way it could be substituted with a strong pronoun.⁴³

- (89) a. In Italia *essi*/ **li* *si* mangiano
 In Italy strong-pron.3pl. clitic.3pl *si* eat.3pl
- b. In Italia *li*/ **essi* *si* mangia
 In Italy clitic.3pl strong-pron.3pl. *si* eat.3sg
 'In Italy they eat them'

As observed, many details need further confirmation, but if we are right Standard

43 Note that our hypothesis predicts that those speakers that accept the clitic strategy in (89b) would also be able to combine both the strong pronoun in the dislocated position and the clitic in the object position (i). We found such examples in the web (ii), but a more accurate analysis of the data remains for further research:

- (i) In Italia *essi li si* mangia
- (ii) a. *Essi li si* può trovare in commercio in qualunque store apposito della vostra città
 These them se may find in market in any store specialized of.the your city
 'Those, you may find them on sale in any specialized store in your city'
- b. In pratica *esse le si* deve scegliere sulla base di due parametri principali
 In practice, these them se must choose under.the base of two parameters principal
 'In practice, they must be chosen on the basis of two main parameters'

It could be argued that the strong pronoun *essi* are dislocated objects in (i)-(ii) but subjects in (88), but if so the initial argument no longer has any motivation. Finally, examples in (iii) show that in Italian, as in Spanish, when V's complement is preposed, agreement is mandatory; an unexpected result if it is receiving accusative Case.

- (iii) a. Gli spaghetti.pl in Italia *si* mangiano
 the spaghetti, in Italy SE eat.pl
- b. *Gli spaghetti, in Italia *si* mangia
 the spaghetti.pl in Italia SE eats

Italian and Spanish are amenable to a unified analysis with respect to the syntactic and morphological derivation of non-paradigmatic SECs, and their differences may follow from independently motivated divergences.

On the other hand, Romanian constitutes an entirely different case. According to Dobrovie-Sorin (1998), in Romanian there is no I-SEC; i.e., following Cinque (1988), there is no [+arg] SE. In that case, *se* in Romanian reduces the V valence by one as in the case of middle voice. The evidence for this proposal comes from a variety of sources, including, for instance, the ability to represent the external argument by means of a *by* phrase (90), its distribution more restricted than in Spanish or Italian, or its inability to control into complement clauses.

- (90) S-au adus mai multe îmbunătățiri de către specialiști.
se-have.3pl brought several improvements by experts
'A number of improvements have been brought by experts.'

These facts suggest, following MacDonnald & Maddox (2018), Maddox (2021), that diachronically *se* is in a different stage in Romanian, thus a different lexical item in synchronic terms: while *se* is a clitic pronoun in Spanish or Italian, it is a voice marker in Romanian. As a consequence, agreement facts must be derived in an entirely different way, which is consistent with the observation in the literature (Dobrovie Sorin 1998; Cornilescu 1998; Giurgea 2019) that in Romanian there are no agreement mismatches of the kind exemplified in this paper.

7. Final conclusions

In conclusion, previous analyses of non-paradigmatic *se* constructions in Romance get short of capturing the variation on agreement in Spanish, and are forced to make a structural distinction between passive and impersonal constructions that requires *ad hoc* mechanisms to deal with them. Instead, we present a unified analysis that eliminates construction-based particularities and takes into account the great variation in agreement among dialects and speakers. The weight of the explanation is on the lexical specifications of the pronoun *se* as the subject of the sentence, playing together with regular syntactic mechanisms, which ends up with an incomplete agreement specification in Morphology. Variation --cross-dialectal as well as dialect-internal-- comes from the application of post-syntactic operations to the specific configuration created by syntax: the obligatory/optional application of the morphological rule that we have called “clitic mutation”, and some post-computation number attraction effects subsumed under the notion “number harmony”. Our analysis also helped us to pinpoint a new asymmetry between postverbal and preverbal/null objects that

has gone unnoticed in the literature.

If on the right track, the results of this paper open new questions and research ways to deal with old issues on the syntax and morphology, as well as processing, of agreement. An area where much work may be done is in making more precise the syntax and morphology of quirky contexts and the limits of variation at the post-syntactic level. Once we reconsider number agreement facts in those new terms, the possible contexts of defectiveness and its potential to reanalyze many constructions (quirky subjects, long distance agreement, expletives, etc.) are quite broad. Our proposal also has consequences for the analysis of classical agreement restrictions and their extensions. In particular, if number agreement in many of these contexts is treated as a post-syntactic effect, what Baker (2008) calls the “two-and-a-half agreement” becomes an “at-most-two-agreement” factor, reducing the playing field of possible explanations for agreement restrictions considerably. That is particularly clear if, as argued here, syntactic agreement applies in a single cycle.

Related to the previous point, there is a question that arises on the organization of agreement among the different components in the faculty of language in the broad sense (FLB: Hauser et al. 2002, Fitch et al 2005). The issue is whether the syntactic as well as processing differences between person vs. number vs. gender, and similar cases in the feature system reported in the literature reflect a cartographic structure (Shlonsky 1989; Sigurdsson & Holmberg 2008; Mancini et al. 2016, among others) or rather an “all at once” checking/copying syntactic agreement operation (Chomsky 1995, 2000, among others) together with some processes in other components. It is worth noting that the two approaches are not incompatible and, in fact, it might be the case that we are dealing with a non-uniform system (see Dillon et al. 2013, Zawiszewski et al. 2016, among others, for discussion). However, if the line of analysis we propose in this paper may be effectively extended, there is room to reinterpret the evidence for separated probes in a more compact syntactic system. Fine grained theoretical and experimental work on the syntax of agreement restrictions as well as on language processing would help clarify some of these issues.

To finish, as we have mentioned in several places through the discussion, our results on SECs intermingle in a complex way with the analysis of DOM, clitic doubling, and the syntax and morphology of Romance clitics. As we suggested, most of the issues involved are independent of SECs, but these structures are a particularly well fitted tool to reveal new properties of the system of clitics and agreement.

Acknowledgements

The two authors, listed in alphabetical order, are equally responsible for the entire content of the paper. We are grateful to Pablo Albizu, Alejo Alcaraz, Olga Fernández Soriano, Angel Gallego, Silvia Gumiel, Jon MacDonald, Jairo Nunes, Paco Ordóñez, Andrés Saab, and Myriam Uribe-Etxebarria. To Andrés, as well as to Luigi Andriani, Carlos Muñoz, Mercedes Pujalte, and Pablo Zrdojewski, very special thanks for their patient discussion and important clarifications on their respective dialects. Previous versions of this paper were presented at the *II Encuentros sobre dialectos del español*, Ciudad Real, as well as in the *Basque Group of Theoretical Linguistics (HiTT) 2018-2019 Seminar Sessions*. We are thankful to the participants in these events for lively discussion and suggestions. This work was made financially possible in part by the Government of Extremadura's grant **GRI5163** (*Grants to Research Groups Program*) to the second author, and by the institutions supporting the research activities of the Basque Group of Theoretical Linguistics (HiTT): the Basque Government grant number **IT769-13** (*Euskal Unibertsitate Sistemako Ikerketa-taldeak*), the Spanish Government's Ministry of Economy and Innovation (MINECO) grant number **VALAL FFI2014-53675-P**, and Ministry of Science, Research and Universities (MCIU) grant number **PROLE PGC2018-093464-b-100** financed partially by the European Union's FEDER program.

References

- Ackema, Peter & Ad Neeleman. 2018. *Features of person: from the inventory of persons to their morphological realization*. Cambridge, MA: MIT Press.
- Albizu, Pablo & Luis Eguren. 2000. An optimality theoretic account for "Ergative Displacement" in Basque. In *Morphological analysis in comparison*, ed. by W. U. Dressler et al., 1-25. Amsterdam: John Benjamins.
- Alcaraz, Alejo. 2021. *Agreement, Binding and the structure of clitics*. Ph.D. Dissertation, University of the Basque Country (UPV/EHU)
- Aldama, Nuria. 2016. Pseudo-relatives complement of perception predicates. University of the Basque Country, MA dissertation. [<http://hdl.handle.net/10810/21499>].
- Anagnostopoulou, Elena. 2003. *The Syntax of Ditransitives*. Evidence from Clitics. Berlin / New York: Mouton de Gruyter.
- Anagnostopoulou, Elena. 2005. Strong and weak person restrictions: A feature checking analysis. In *Clitic and affix combinations: Theoretical perspectives*, ed. by L. Heggie & F. Ordóñez, 199-235. Amsterdam: John Benjamins.
- Ariztimuño, Borja. 2017. Finite verbal morphology. In M. Martínez-Areta, ed. *Basque and Proto-Basque. Language- internal and typological approaches to linguistic*

- reconstruction*. Frankfurt: Peter Lang ed.
- Arregi, K., & Nevins, A. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout* (Vol. 86). Springer Science & Business Media.
- Baker, Mark. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.
- Baker, Mark & Ruth Kramer 2018. Doubled Clitics are Pronouns: Amharic Objects (and beyond), *Natural Language and Linguistic Theory* 36:1035-1088.
- Béjar, Susana. 2003. *Phi-syntax. A theory of agreement*. Ph.D. Dissertation, University of Toronto.
- Béjar, Susana. 2008. Conditions on Phi-agree. In D. Harbour, D. Adger, and S. Béjar, eds. *Phi theory: Phi-features across modules and interfaces*, Oxford University Press, 130-158.
- Béjar, Susana, & Rezac, Milan. 2009. Cyclic Agree. *Linguistic Inquiry* 40:35-73. [<https://doi.org/10.1162/ling.2009.40.1.35>]
- Belletti, Adriana. 1982. Morphological passive and pro-drop: The impersonal construction in Italian. *Journal of Linguistic Research*, 24:1-34.
- Benmamoun, Elabbas, Archana Bhatia & Maria Polinsky. 2009. Closest conjunct agreement in head final languages. *Linguistic Variation Yearbook* 9, 67-88.
- Bhatt, Rajesh. 2005. Long-distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23:757-807.
- Bianchi, Valentina. 2006. On the syntax of personal arguments. *Lingua* 116, 2023-2067.
- Boeckx, Cedric. 2004. Long-distance agreement in Hindi: Some theoretical implications. *Studia linguistica*, 58:23-36.
- Bleam, Tania. 1999. *Leista Spanish and the syntax of clitic doubling*. University of Delaware dissertation.
- Bonet, Eulalia. 1991. *Morphology after syntax: Pronominal clitics in Romance*. Ph.D. Dissertation, MIT.
- Bonet, Eulalia. 1994. The person-case constraint: A morphological approach. *The Morphology-Syntax Connection. MIT working papers in linguistics*, 22, 33-52.
- Burzio, Luigi. 1986. *Italian syntax: A Government-Binding approach*. Dordrecht: Reidel.
- Cennamo, Michela. 2014. Passive and impersonal reflexives in the Italian dialects. In P. Benincà, A. Ledgeway & N. Vincent, eds. *Diachrony and Dialects: Grammatical Change in the Dialects of Italy*, Oxford: Oxford University Press, pp. 71-95.
- Chierchia, Gennaro. 1995. Variability of impersonal subjects. In *Quantification in Natural Languages*, ed. by Emon Bach et al., 107-144. Dordrecht: Kluwer.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by Step: Essays in Honor of Howard Lasnik*, ed. by R. Martin et al., 89-155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In M. Kenstowicz, ed., *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, 1-52.
- Cinque, Guglielmo. 1988. On *si* constructions and the theory of *arb*. *Linguistic Inquiry*. 19:521-581. [<http://www.jstor.org/stable/4178596>].
- Collins, Chris. 2017. On the implicit argument in the short passive. Ms. NYU.
- Cornilescu, Alexandra. 1998. Remarks on the syntax and the interpretation of Romanian

- middle passive sentences. *Revue roumaine de linguistique* 43, 317-342.
- CORPES XXI = Real Academia de la lengua: databank
- D'Alessandro, Roberta. 2007. *Impersonal Si Constructions*. Berlin: Mouton de Gruyter.
- D'Alessandro, Roberta, & Artemis Alexiadou. 2003. Inclusive and exclusive impersonal pronouns: A feature-geometrical analysis. *Rivista di Grammatica Generativa* 27:31-44. [<http://hdl.handle.net/10278/2452>]
- De Benito, Carlota. 2013. (Esa tela) se la descose: la pronominalización del paciente en las impersonales reflejas del español peninsular. *Borealis: An International Journal of Hispanic Linguistics* 2:129-157. <http://dx.doi.org/10.7557/1.2.2.2676>
- DeMello, George. 1995. Concordancia entre el verbo pronominal de tercera persona y su sustantivo: *se venden flores* vs. *se vende flores*. *Anuario de Letras*, 33:59-82.
- Demonte, Violeta. 1986. Predication and passive. In *Generative studies in Spanish syntax*, ed. by Y. Bordelois et al, 51-66. Dordrecht: Foris.
- Demonte, Violeta & Isabel Pérez Jiménez. 2012. Closest conjunct agreement in Spanish DPs: Syntax and beyond. *Folia Linguistica*, 46:1, 21-73
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. 2013. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69:85-103.
- Di Tullio, Ángela; Andrés Saab & Pablo Zdrojewski. 2019. Citic doubling in a doubling world. The case of Argentinean Spanish reconsidered. In A. Gallego, ed., *The syntactic variation of Spanish dialects*. Oxford University Press, 215-244.
- Di Tullio, Ángela & Pablo Zdrojewski 2006. Notas sobre el doblado de clíticos en el español rioplatense: asimetrías entre objetos humanos y no humanos." *Filología* 1, pp. 13-44
- Dobrovie-Sorin, Carmen. 1998. Impersonal se constructions in Romance and the passivization of unergatives. *Linguistic Inquiry*, 29:399-437.
- Espinal, M.Teresa and Sonia Cyrino. 2017. The definite article in Romance expletives and long weak definites. *Glossa: a journal of general linguistics* 2:23, 1-26. [<https://doi.org/10.5334/gjgl.160>]
- Etxepare, Ricardo. 2006. Number long distance agreement in (substandard) Basque. In *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 40:303-350.
- Fernández, Beatriz. & Pablo Albizu. (2000). Ergative displacement in Basque and the division of labour between morphology and syntax. In *Proceedings of the Chicago linguistic society* 36, ed. by A. Okrent & J.P.Boyle, vol.2:103-118. Chicago, Chicago Linguistic Society.
- Fernández de Lagunilla, Marina. 1975. Acerca de la secuencia "SE impersonal + enclítico de 3ª persona": ¿Una restricción superficial? *Revista Española de Lingüística*, 5:177-195
- Fernández Ordóñez, Inés. 1999. *Leísmo, laísmo y loísmo*. In *Gramática descriptiva de la lengua española*, ed. by Ignacio Bosque & Violeta Demonte, 1317-1398. Madrid: Espasa Calpe.
- Fitch, W Tecumseh; Mark D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: clarifications and implications. *Cognition*, 97:179-210.
- Gallego, Ángel. 2016. Three types of prepositions in Spanish *se*-sentences. Consequences for cross-dialectal studies. *Dialectología* 17:51-70.
- Gallego, Ángel. 2019. Long distance agreement in Spanish dialects. In *Agreement, case and locality in the nominal and verbal domains*, ed. by Ludovico Franco et al., 85-108.

- Berlin: Language Science Press.
- García Negroni, Maria Marta. 2002. La construcción media con *se*. In *Las construcciones con se*, ed. by Cristina Sánchez López, 276-304. Madrid: Visor Libros.
- García-Page, M. 2010. Locuciones verbales con clítico en español del tipo dársele. *Verba hispánica*, 18:135-145.
- Giurgea, Ion. 2019. On the Person Constraint on Romanian *se*-passives. In L Franco, M. Marchis & M. Reeve, eds., *Agreement, case and locality in the nominal and verbal domains*, Berlin: Language Sciences Press, 109-148.
- Gómez, Ricardo & Koldo Sáinz. 1995. On the origin of the finite forms of the Basque verb. In J.I Hualde, J. Lakarra & R. Trask, eds. *Towards a history of the Basque language*, Amsterdam/Philadelphia: John Benjamins, 235-274.
- Gràcia i Solé, Lluïsa & Francesc Roca. 2017. Había (...) más de lo que parecía: una discusión existencial. In *Relaciones sintácticas: homenaje a Josep M. Brucart y M. Lluïsa Hernanz*, ed. by Angel J. Gallego, Yolanda Rodríguez Sellés & Javier Fernández Sánchez, 365-386.
- Guéron, Jacqueline. 1983. L'emploi 'possessif' de l'article défini en français. *Langue Française* 58:23–35. [<https://doi.org/10.3406/lfr.1983.6413>]
- Guéron, Jacqueline. 2006. Inalienable possession. In *The Blackwell companion to syntax*, ed. by Martin Everaert & Henk van Riemsdijk, 589–638. Oxford: Blackwell.
- Hale, K. (2001). Eccentric agreement. In *On case and agreement*, ed. by P. Albizu & B. Fernández, 15–48. Bilbo: Euskal Herriko Unibertsitatea.
- Harley, Heidi & Elisabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78, pp. 482-526.
- Harris, James & Morris Halle. 2005. Unexpected plural inflections in Spanish: Reduplication and metathesis. *Linguistic Inquiry* 36.2, p. 195-222.
- Hauser, Mark. D.; Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298:1569-1579.
- Laka, Itziar. 1993. The structure of inflection. In *Generative studies in Basque linguistics*, ed. by J. I. Hualde and J. O. de Urbina, 21-70. Amsterdam: John Benjamins.
- Lakarra, Joseba. 2005. Prolegómenos a la reconstrucción de segundo grado y al análisis del cambio tipológico en (proto) vasco. *Palaeohispanica* 5, pp. 407-470.
- Lemus, Jorge. 2014. 'Se venden tortillas': un análisis del uso del clítico "se" en español. *Científica* 13, 7-16.
- López, Luis. 2012. *Indefinite objects: Scrambling, choice functions, and differential marking*. Cambridge, MA: MIT Press.
- López, Luis. 2020. Case, concord, and the emergence of default. *Languages* 5:12, doi:10.3390/languages5020012.
- MacDonald, Jonathan E. 2017. An implicit projected argument in Spanish impersonal and passive *se* constructions. *Syntax* 20:353-383. [<https://doi.org/10.1111/synt.12146>].
- MacDonald, Jonathan & Mathew L. Maddox. 2018. Passive *se* in Romanian and Spanish: A subject cycle. *Journal of Linguistics* 54.2, pp. 389-427.
- MacDonald, Jonathan & Jeriel Melgares. 2018. On (un)grammatical clitic sequences in Spanish impersonal *se* constructions. Ms. U. of Wisconsin.
- MacKenzie, Ian. 2013. *Se azotaron a los delincuentes: a case study in arbitrary exclusion*. *Bulletin of Spanish Studies. Hispanic Studies and Research on Spain, Portugal and Latin*

- America*, 90:6.
- Maddox, Mathew L. 2021. The Development of Se from Latin to Spanish and the Reflexive Object Cycle. In G. Armstrong & J. MacDonald, eds. *Unraveling the complexity of SE*. Dordrecht: Springer, pp. 33-56.
- Mancini, Simona, Frances Postiglione, Alessandro Laudanna & Luigi Rizzi. 2014. On the person-number distinction: Subject-verb agreement processing in Italian. *Lingua* 146:28-38.
- Manzini, Rita & Savoia. 2005. *I dialetti italiani e romanci: L'ausiliare (2a parte); La negazione e l'avverbio; Strutture aspettuali e modali; Il sintagma nominale; Bibliografia* (Vol. 3). Edizioni dell'Orso.
- Martins, Ana Maria. 2005. Passive and impersonal se in the history of Portuguese. In Klaus Push, et al, eds., *Romance corpus linguistics II: Corpora and diachronic linguistics*, Tübingen, pp. 411-430.
- Martins, Ana Maria & Jairo Nunes. 2016. Passives and se constructions. In *The Handbook of Portuguese Linguistics*, ed. by L. Wetzels et al., 318-337. Malden: Wiley-Blackwell.
- McGinnis, Martha. 2005. On Markedness Asymmetries in Person and Number. *Language* 81:3, 699-718.
- Mendikoetxea, Amaya. 1992. On the Nature of Agreement: The Syntax of ARB SE in Spanish. PhD. dissertation, U. Autónoma de Madrid.
- Mendikoetxea, Amaya. 1999. "Construcciones con se: Medias, Pasivas e Impersonales". In *Gramática Descriptiva de la Lengua Española*, ed. by Ignacio Bosque & Violeta Demonte, 1631-1722. Madrid. Espasa-Calpe.
- Mendikoetxea, Amaya. 2008. "Clitic impersonal constructions in Romance: Syntactic features and semantic interpretation". In *Impersonal Constructions in Grammatical Theory*, A. Siewierska (ed.) Special Issue of the Transactions of the Philological Society, 106, 2, 290-336. Oxford: Blackwell.
- Mendikoetxea, Amaya. 2012. Passives and se constructions. In *The Handbook of Hispanic Linguistics*, ed. by J.I. Hualde et al., 477-502. Wiley-Blackwell.
- Mendikoetxea, Amaya, & Adrian Battye. 1990. Arb se/si in transitive contexts: A comparative study. *Rivista di grammatica generativa*, 15:161-195.
- Menuzzi, Sergio. 1999. 1st Person Plural Anaphora in Brazilian Portuguese: Chains and Constraint Interaction in Binding. In *Portuguese syntax: new comparative studies*, ed. by J. Costa, 191-240. Oxford: Oxford University Press.
- Monge, Félix. 1954. Las frases pronominales de sentido impersonal en español. *Archivo de Filología Aragonesa* VII, pp. 7-102.
- Napoli, Donna Jo. 1976. *The two si's of Italian. An analysis of reflexive, inchoative, and indefinite subject sentences*, Indiana University Linguistic Club, Bloomington.
- NGLE = Real Academia de la Lengua Española. 2011.
- Oca, Eduardo. 1914. El pronombre se en nominativo. *Boletín de la Real Academia Española* 1:573-581.
- Odria, Ane. 2017. Differential Object Marking and datives in Basque syntax. PhD. dissertation, University of the Basque Country.
- Ordóñez, Francisco. 2004. "Se Across Romance". Paper presented at the GURT Conference, Georgetown University.
- Ordóñez, Francisco. 2018. On the nature of the impersonal SE: case, interpretation and

- variation. Ms. NY University, Stony Brook.
- Ordóñez, Francisco & Esthela Treviño. 2011. Impersonals with Passive Morphology. In *Selected Proceedings of the 13th Hispanic Linguistics Symposium*, ed. by Luis A. Ortiz-López, 314-324. Somerville, MA: Cascadilla Proceedings Project.
- Ordóñez, Francisco & Esthela Treviño. 2016. Agreement and DOM with *se*: A comparative study of Mexican and Peninsular Spanish. In *The Morphosyntax of Portuguese and Spanish in Latin America*, ed. by M. Kato & F. Ordóñez, 236-258. Oxford: Oxford University Press.
- Ormazabal, Javier & Juan Romero. 2007. The Object Agreement Constraint. *Natural Language and Linguistic Theory*, 25:315-347.
- Ormazabal, Javier & Juan Romero. 2013a. Object Clitics, Agreement and Dialectal Variation. *Probus: International Journal of Latin and Romance Linguistics* 25:301-344.
- Ormazabal, Javier, & Juan Romero. 2013b. Differential Object Marking, case and agreement. *Borealis: An International Journal of Hispanic Linguistics* 2, 221-239.
- Ormazabal, Javier & Juan Romero. 2019a. The formal properties of non paradigmatic SE. *Borealis: An International Journal of Hispanic Linguistics*, 8:55-84.
- Ormazabal, Javier & Juan Romero. 2019b. Prolegomena to the study of object relations. In *Linguisticae Investigationes 42. Special issue on Differential objects and datives - a homogenous class?*, ed. by M.A. Irimia & A. Pineda, 102-131.
- Ortiz de Urbina, Jon 1989. *Parameters in the Syntax of Basque*. Dordrecht: Foris.
- Otero, Carlos P. 1972. Acceptable ungrammatical sentences in Spanish. *Linguistic Inquiry* 3:233-242.
- Otero, Carlos P. 1973. Agrammaticality in performance. *Linguistic Inquiry* 4:551-562.
- Planells, Samanta. 2018. Patrones de variación (idiolectal) en torno a las construcciones con *se* no-paradigmático. *Revista de la Sociedad Española de Lingüística* 47:29-47.
- Pujalte, Mercedes. 2013. Argumentos (no) agregados. Indagaciones sobre la morfosintaxis de la introducción de argumentos en español. PhD dissertation, U. de Buenos Aires.
- Pujalte, Mercedes. 2018. Patrones de concordancia en impersonales y pasivas con *se*. Ms. UNGS/UBA/UNAHUR [Talk presented at the *Encuentros de Gramática Generativa*].
- Pujalte, Mercedes & Andrés Saab. 2012. Syncretism as PF-repair: The case of *se*-insertion in Spanish. In *The end of argument structure? Syntax and Semantics* 38, ed. by M. C. Cuervo & Y. Roberge, 229-260. Bingley: Emerald.
- Pujalte, Mercedes & Andrés Saab. 2014. Sobre la interacción entre caso y concordancia en impersonales y pasivas con *se*. *Traslaciones*, 1:30-55.
- Raposo, Eduardo and Juan Uriagereka 1996. Indefinite SE. *Natural Language and Linguistic Theory* 14:749-810. [<https://doi.org/10.1007/BF00133363>]
- Real Academia Española (RAE): Databank (*CORPES XXI*) [online]. *Corpus del Español del Siglo XXI (CORPES)*. <<http://www.rae.es>>
- Real Academia Española (RAE). 2011. *Nueva Gramática de la Lengua Española [NGLE]*, Madrid, Espasa-Calpe.
- Rezac, Milan. 2003. The fine structure of cyclic Agree. *Syntax*, 6:156-182.
- Rezac, Milan. 2008. The syntax of eccentric agreement: The Person Case Constraint and absolutive displacement in Basque. *Natural Language & Linguistic Theory*, 26:61-106.
- Richards, Marc. 2014. Defective agree, case alternations, and the prominence of person. In *Scales and hierarchies*, ed. by I. Bornkessel-Schlesewsky et al., 173-196. Berlin: Mouton

- de Gruyter.
- Rigau, Gemma & M. Carme Picallo. 1999. El posesivo y las relaciones posesivas. In *Gramática Descriptiva de la Lengua Española*, ed. by Ignacio Bosque & Violeta Demonte, 973-1024. Madrid: Espasa-Calpe.
- Rivero, M. Luisa. 2001. On impersonal reflexives in Romance and Slavic and semantic variation. In *Romance Syntax, Semantics and L2 Acquisition*, ed. by Joaquim Camps & Caroline R. Wiltshire, 169-195. Amsterdam: John Benjamins.
- Roca, Francesc. 1996. Morfemas objetivos y determinantes: los clíticos del español. *Verba* 23:83-119.
- Rodríguez-Ordoñez, Itxaso. 2016. Differential Object Marking in Basque: Grammaticalization, Attitudes and Ideological Representations. PhD dissertation, Urbana-Champaign: University of Illinois.
- Saab, Andrés. 2014. Syntax or nothing: Some theoretical and empirical remarks on implicit arguments. *Borealis* 3:125-183. [<http://dx.doi.org/10.7557/1.3.2.2952>]
- Saab, Andrés. 2018. *Se*-semantics. An LF motivation for the Case-filter. Ms. CONICET-UBA
- Sánchez, Liliana. 2006. "Clitic Doubling and the Checking of Focus." Ms. Rutgers University.
- Sánchez, Liliana & Pablo Zdrojewski. 2013.
- Sánchez López, Cristina. 2002. Las construcciones con *se*. Estado de la cuestión. In *Las construcciones con se*, ed. by Cristina Sánchez, 13-163. Madrid: Visor Libros.
- Santiago, R. 1975. Impersonal "se le (s), se lo (s), se la (s)". *Boletín de la Real Academia Española*, 55(204), 83-108.
- Shlonsky, Ur. 1989. The Hierarchical Representation of Subject-Verb Agreement, ms. University of Haifa.
- Sigurðsson, Halldór Á. & Anders Holmberg. 2008. Icelandic Dative Intervention: Person and Number are distinct probes. In *Agreement Restrictions*, ed. by R. D'Alessandro et al., 251-279. Berlin: Mouton de Gruyter.
- Silva-Corvalán, Carmen. 1980. La función pragmática de la duplicación de pronombres clíticos." *Boletín de Filología* 31.2, pp. 561-570.
- Torrego, Esther. 1998. *The dependencies of objects*. Cambridge, MA: MIT Press.
- Torrego, Esther. 2008. Revisiting Romance "se". *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*, 785-791.
- Trommer, Jochen. 2008. Third person marking in Menominee. In D. Harbour, D. Adger, and S. Béjar, eds. *Phi theory: Phi-features across modules and interfaces*, Oxford University Press, 221-250.
- Uriagereka, Juan. 1997. Formal and Substantive Elegance in the Minimalist Program. In *The Role of Economy Principles in Linguistic Theory*, ed. by Chris Wilder et al. 170-204. Berlin: Akademie Verlag.
- Villata, Sandra. 2017. *Intervention effects in sentence processing*. Ph.D dissertation, Univ. of Geneva.
- Van Riemsdijk, Henk. 1997. Left dislocation. In *Materials on Left Dislocation*, ed. by E. Anagnostopoulou et al., 1-12. Amsterdam: John Benjamins. 1-12.
- Vergnaud, Jean-Roger & Maria Luisa Zubizarreta. 1992. The definite determiner and the inalienable constructions in French and in English. *Linguistic inquiry* 23.4, pp. 595-652.

- Zawiszewski, Adam, Mikel Santesteban & Itziar Laka. 2016. Phi-features reloaded: An ERP study on person and number agreement processing. *Applied Psycholinguistics* 37:601-626.
- Zdrojewski, Pablo. 2008. *Por quién doblan los clíticos. Restricciones morfosintácticas sobre la duplicación pronominal en el español rioplatense*. Ph.D. dissertation, Universidad Nacional del Comahue.
- Zdrojewski, Pablo. 2018. La marcación diferencial de objetos inanimados: veinte visitas a la gramática del español inspiradas por Ángela Di Tullio. Ms. Universidad Nacional de General Sarmiento / Universidad de Buenos Aires.

Javier Ormazabal

University of the Basque Country (UPV/EHU)
Basque Group of Theoretical Linguistics (HiTT)
Department of Linguistics & Basque Studies (HEIS)
Unibertsitateen Ibilbidea 5, 01006 Vitoria-Gasteiz, Spain
javier.ormazabal@ehu.eus

Juan Romero

University of Extremadura (UNEX)
Basque Group of Theoretical Linguistics (HiTT)
Department of Hispanic Philology
Avenida de la Universidad s/n, 10071 Cáceres, Spain
juantxoromero@gmail.com

Appendix: full examples and sources:

Short notes on our searches in CORPES XXI:

i. It is not our goal to present a precise dialectal analysis of agreement patterns in SECs in Spanish but to introduce a broader picture of their (a)systematicity and the range of variation within each agreement pattern, independently of whether speakers allow various strategies or just one.

Moreover, the searches have not been designed to have any statistical relevance. In the vast majority of the cases, the size of the samples is irrelevant, since the overall numbers depend on the frequency of other, more common, structures that coincide superficially with the sequence we targeted, and do not say much about the real frequency of the structure we are analyzing. For instance, a great number of “se V+pl + a (prep.)” cases are inchoative/ unaccusative examples of the form *se acercan a* (‘they approach’) or “inherently reflexive” verbs like *se dedican a* (‘they occupy themselves with’). The search itself was a tool to find instances of a particular frame, and we generally had to delve into the sets of examples to find a relatively small number of relevant ones. But we think that the mere existence/lack of a pertinent number of cases having a direct bearing on the issue under discussion is enough to support our goals.

ii. Yet, when relevant, we tried to obtain as much information as possible about the geographical distribution of the different constructions. In all cases we performed a pre-search where data was organized following the zonal parameter to see whether it revealed some possible dialectal cut. In those cases where either the bibliography on the topic or our pre-search indicated some possible dialectal effect, we included areal and country parameters, to have the information more organized for that purpose, and selected the same number of hits for each country we analyzed.

iii. In the cases in which no worth mentioning results emerged in the pre-search, we decided to run our search with no organizing criterion. By default, *Corpes XXI* organizes the sample chronologically beginning from the earliest instances (2002-2016). That guarantees that the sub-corpus analyzed is as random and hypothesis-neutral as possible concerning all parameters, also including a variety of geographic areas.

iv. Given the size of the corpus, we restricted our search to the “non-fiction” subcorpus, where non-paradigmatic *se*-constructions are much more common than in fiction, and in some cases to present tense, where *se*-constructions are also more present.

2. Agreement with plural [+DOM] objects

CORPES SEARCH:

Lema: SE: pronombre personal, 3ª persona, Caso inespecífico + **proximidad** + **subcorpus:**

Proximidad:	1. verbo	3ª persona plural	[distancia=1derecha]
	2. lema: <i>a</i>	preposición	[distancia=2derecha]
	3. artículo	plural	[distancia=3derecha]
Subcorpus:	“no ficción”		

■ Examples in the text:

(16a): Instit. de Estudios de Promoción de la Autonomía: *Cultura política en el Atlántico Norte*.

Managua, 2001

En 1996 se eligen a las primeras autoridades municipales de las regiones autónoma

‘In 1996 the first municipal authorities of the autonomous regions were elected’

(16b): Guzmán, Humberto: *Los extraños*. México D. F.: Tusquets Editores, 2001

De esta manera mantenían controladas las principales instalaciones civiles y militares, habían ametrallado la fachada del Museo Nacional y se rastreaban a los miembros del Partido Comunista Checoslovaco.

‘This way civil and military installations were under control, and they machine-gun the Museum National facade, the members of the Czechoslovakian Party Communist were tracked down’

(16c): *Arqueoweb*. Madrid [España]: ucm.es/arqueoweb, 2001-12-03

Al iniciarse la menstruación se aislaban a las jóvenes en un refugio donde cumplían con varios tabúes de comida y aprendían normas de conducta apropiada.

‘When the menstruation begins, young women were isolated in a refuge where they keep to several food tabus and learn the rules of proper behavior’

3. “Long distance” plural agreement

We designed a narrow search, where we analyzed the entire corpus alphabetically organized by main verb. Out of the 7160 items, more than three quarters were modals or causatives [*deber, poder, querer, soler; hacer, mandar*] and perception verbs [*oir, escuchar, sentir, ver*]. See results below.

Search:

Lema: SE pronombre personal, 3ª Caso inespecífico + **proximidad** + **subcorpus:**

Proximidad:	1. verbo	3ª persona plural	[distancia=1derecha]
	2. verbo	infinitivo	[distancia=2derecha]
	3. artículo	plural	[intervalo=4derecha]
Subcorpus:	“no ficción”		

Organizado por: lema primero derecha

Analyzed: **entire corpus** (7160 items)

Results: the vast majority of the cases are bridge verbs or verbs in the edge of becoming so [*intentar*, *buscar*(7), *pretender*, *desear*(9), *necesitar* (10), *pensar*, *lograr*, *conseguir* (14)], but there are clear cases with other verbs [e.g. *osar*; *requerir*, *valorar*].

■ Examples in the text:

(18a): <http://universidadydiscapacidadeniberoamerica.fundacionuniversia.net/profesiones-nunca-podra-sustituir-robot/1>

En esta profesión **se requieren hacer** *evaluaciones* tanto objetivas como subjetivas de los casos, [...].

‘In this job, evaluations, both objective and subjective for each situation, are required’

(18b) Amanda Mars: “Grecia da por cerrados los recortes y confía en lograr el rescate”, *El País* 2012-02-18. [Spain]

El Gobierno [...] acabó de pactar una rebaja de las pensiones [...] que, según el Ejecutivo, está aún por concretar, aunque **se valoran reducir** *las superiores a 1300 euros*.

‘The Government concluded an agreement on a reduction of pensions [...] which, according to the executive, is yet to be decided, although reducing those higher than 1300 euros will be considered.’

(18c): <https://www.ultimahora.com/esperan-que-cargo-ipta-sea-meritocracia-n333587.html>

Se requieren nombrar a los responsables de la Dirección Nacional

‘It is necessary to appoint National Directorate leaders’

4. Plural agreement with non-arguments

There are intrinsic difficulties to design a search that would give us a general result: a) the cases of agreement with non-arguments is low, and they do not follow a specific pattern, at least not one that might be identified as easily as in the previous cases. For both pseudo-arguments and adjunct-cases we based our results on Google searches, and targeted the ones in Gallego’s and our original example on adjuncts as well as some additional ones, so the search is lexically restricted in this case. In response to an anonymous reviewer’s inquiry, in March 2020 we extended our search to other temporal NP-modifiers.

(20a)

(20b)

(23): <http://www.ambito.com/886018-macri-se-reunio-con-periodistas-de-espectaculos-y-deslizo-que-podria-ir-por-la-reeleccion>

Brey [...] aclaró que tampoco "se hablaron de temas de la farándula

‘Brey explained that there was no talk about showbusiness issues’

(24a): *ABC Digital*. Asunción [Paraguay], 2004-03-06

Indicó que se deben recurrir a otras figuras para penalizar el hecho punible

(26a): Attributed to Juanxto López (LAB) in an interview; *El Correo*, Bilbao 2009/01/18.

¿hay alguna diferencia de ventas entre el País Vasco, donde no se abren los domingos, y el Estado español, donde sí se trabaja?

‘Is there any difference in sales between the Basque Country, where shops are closed on Sunday, and the Spanish State, where they do open?’

5. Clitic strategy

Search:

Lema: SE pronombre personal, 3ª Caso inespecífico + **proximidad** + **subcorpus:**

Proximidad: 1. pronombre personal 3ª persona plural, acusativo [distancia=1 derecha]

2. verbo 3ª persona singular [distancia=2 derecha]

Subcorpus: “no ficción”

Organizado por: area lingüística > país

Results:

i) In Argentina, Chile, Uruguay, Bolivia and, very especially, Paraguay and Ecuador there are plenty of examples both with inanimate objects and with animate ones.

ii) In Venezuela there are a few ones, but not many (the entire corpus revised), and in Colombia numbers are very low.

iii) In Central America there are very few cases (e.g. in the entire corpus for El Salvador we found just one example)

iv) All the examples found in Peninsular Spanish involve the feminine clitic *la(s)*, which suggests that they appear in *laísta* speakers, but more work is needed.

■ The corpus is full of coordinated sentences where a new expression is first introduced and agrees with the verb in a postverbal fashion (*se sacan tres cuadros; se compraban huevos*; see section 2.1) and then are referred anaphorically by means of a clitic (*se los coloca; se los vendía*) or agreement (*se guardaban*). In fact, example (B) combines both possibilities in the same sentence.

(A): «Apicultura» abc.com.py, Asunción (Paraguay), 2003-02-12

Seguidamente **se sacan tres cuadros con crías operculadas**, con todas las abejas adheridas y **se los coloca** en cada uno de los cajones que sirven de núcleo.

‘Next, three frames with opercular breeding are extracted with all the bees stucked on, and they are set in each of the boxes serving as nucleae’

(B): «El industrial que la tiene clara» Montevideo (Uruguay), diarioelpais.com, 2001-11-08
[...] hace unos 14 o 15 años **se compraban huevos** en época de producción alta, **se guardaban** en cámaras frigoríficas y **se los vendía** cuando faltaba en plaza,
‘14 or 15 years ago, eggs were bought in high production seasons, they were stored in walk-in freezers and sold when there was need’

■ In a Google search restricted to Peninsular Spain, the sequence *se los mire* is found in different types of journals and publications:

(C) www.revistaceramica.com/detalle2.aspx?id=1486
... sabemos de la propiedad de dos colores que tienen *algunos cuerpos* de presentar dos coloraciones diferentes según la dirección en que **se los mire**
‘We know the two-color property by which some bodies present two different colorings according to the direction they are looked at’

(D): abcblogs.abc.es/.../demuestran-la-existencia-de-los-cristales-del-tiempo-17128.asp
... *los cristales* rompen espontáneamente esas simetrías espaciales, ya que cambian según el ángulo desde el que **se los mire**
‘The crystals break spontaneously those spatial simmetries because they change according to the angle they are looked at’

(E): www.tindas.es/decoracion-de-balcones/
Son *balcones* cargados de belleza que, **se los mire** como se los mire, son únicos.
‘Those are balcon full of beauty, unique no matter how you look at them’

■ Examples in the text:

(36a): Direcc. Gral. de Planif y Ordenam. Territ, *Guía Metodológica para la formulación de los Planes Departamentales* [La Paz (Bolivia), Unidad de Ordenam. Territorial, 2001]
Para efectos operativos *estas ponderaciones se las realiza* por sectores
‘For operational reasons these weightings are made by areas’

fn. 17.(ia): H. ARIDJIS, *La zona del silencio*. México D. F., Punto de Lectura, 2005
Cuando **se la** desencaja, *la luz de la lámpara* alumbra unas facciones ocres, [...]
When it is disarranged, the light of the lamp illuminates his ochre features

(ib): P. RAMIS, *Esencia prejurídica del derecho*. Mérida, Ed. Venezolana C.A., 2002
En cuanto a *las incorpóreas*, no **se las** veía como cosas, sino como derechos puros.
With respect to the incorporeal ones, they were not seen as things, but as pure rights’

(ic): F. Henriquez Gratereaux 2002, *Disparatario*. Sto. Domingo (Rep. Domin.), Alfa y Omega
Al tratar de actualizar y hacer comprensibles *las palabras del bautizo*, **se las** degrada.
‘When trying to update and make understandable baptism words, they are degraded’

(id): «Francisco Rodríguez Cascante», itcr.ac.cr/revistacomunicacion, Cartago (C. Rica), 2002-01-03

Por el contrario, si **se los** considera variables, [...]e

‘On the contrary, if they are considered variables’fn19.(ia): «Esta noche disfrute la lluvia del cometa Halley» [elheraldo.hn, Tegucigalpa (Honduras), 2011-05-06

[... estrellas fugaces], y hoy es el día que **se las** puede ver con mayor facilidad.

“and today is the day when they (the shouting stars) may be seen the easiest’

fn18.(ib): G. Morales-Luna. 2009. Las Matemáticas y su aplicación en las Comunicaciones Digitales.

Revista.unam.mx., México D. F., 2009-01

Sus funciones primitivas entrañan una complejidad exponencial cuando se las quiere simular por esquemas clásicos

‘Their primitive functions result in an exponential complexity when someone wants to simulate them by classical schemes’

5. Clitic climbing with clitics

Search:

Lema: **SE** pronombre personal, 3ª Caso inespecífico + **proximidad** + **subcorpus:**

Proximidad: 1. pronombre personal 3ª persona plural, acusativo [distancia=1derecha]

2. verbo 3ª persona singular [distancia=2derecha]

3. verbo infinitivo [distancia=3derecha]

Subcorpus: “no ficción”

Organizado por: area lingüística > país

Results: plenty of examples from Argentina, Paraguay, Uruguay, Chile, Ecuador, Bolivia. Examples are attested in all linguistic areas, although much less prominently.

I Google search: “hacen dos”

Vivir en una carpa o una casucha sin calefacción cuando hacen dos grados bajo cero es inaceptable

[https://www.mendozapost.com/nota/141825-una-villa-de-emergencia-poblada-de-](https://www.mendozapost.com/nota/141825-una-villa-de-emergencia-poblada-de-refugiados-y-adictos-se-instalo-en-paris/)

[refugiados-y-adictos-se-instalo-en-paris/](https://www.mendozapost.com/nota/141825-una-villa-de-emergencia-poblada-de-refugiados-y-adictos-se-instalo-en-paris/)

6. Number harmony vs. clitic mutation: relative clauses

Search:

LEMA: **SE** pronombre personal, 3ª Caso inespecífico + **proximidad** + **subcorpus:**

Proximidad: 1. **que** [distancia=1izquierda]

2. verbo 3ª persona singular [distancia=1derecha]

Subcorpus: “no ficción”

■ Examples in the text:

(87a): «EL NORTE VALLADOLID» (El Norte de Castilla) [España]

se valora muy positivamente [...] así como la participación en programas europeos, **los servicios que se proporciona** a los alumnos y la preparación de idiomas.

‘It is highly valued [...] as well as involvement in European programs, the services provided to the students and language learning’

(87b): M. Tello & E. Basqueto. 2006. *II Seminario de Formación Profesional*. Córdoba (Argentina), Universidad Nacional de Río Cuarto

Selección de **las personas a las que se entrevistaron** para obtener más datos

‘Selection of the people interviewed to obtain more data’

■ Google search: “se vende” vs. “se venden” (see text for details)

<https://www.silicon.es/submit-la-herramienta-de-comixology-que-permite-a-los-autores-publicar-online-sus-propios-comics-111091>

Algo similar a lo que sucede [...] con los libros que **se vende** a través de las plataformas en Internet.

‘Something similar to what happens with the books sold through the Internet’

fn 40. (ii) **Essi li si** può trovare in commercio in qualunque store apposito della vostra città

[<https://www.giornalenotizie.online/stucco-veneziano-come-si-fa/>]

(iii) In pratica **esse le si** deve scegliere sulla base di due parametri principali

<https://www.borsainside.com/fisco/70659-confronto-fondi-pensione/>