

Clitics out of control

On an *Agree* failure and its dramatic LF consequences

Andrés Saab

IIF-SADAF-CONICET

Universidad de Buenos Aires

Al_Saab75@yahoo.com.ar

First draft, 26.09.2020

(comments welcome)

ABSTRACT: Within the framework of a uniform theory of the so-called *se* constructions in Spanish, I propose to explain a control ban that has received almost no attention in the previous bibliography. Specifically, as long as a subject control sentence has an impersonal *se* as controller, the subordinate infinitive clause cannot contain any other instance of the clitic *se*, other than the so-called spurious *se*. The source of this restriction follows, as I will argue, from a legibility problem at LF produced, specifically, by a failed attempt to apply *Agree* between *PRO* and the embedded *se*, which, as we shall see, acts as a probe for A-movement. If the explanation that I offer is correct, it also follows a series of theoretical conclusions that directly affect the way in which we must conceive of the design of *Agree* in the syntax and its effect at the LF interface. In particular, the system tolerates certain *Agree* failures (Preminger 2014) as long as it does not affect legibility in the semantics. Indeed, the theory of *se* constructions that I assume here derives the distinction between paradigmatic and non-paradigmatic *se* as the result of successful or unsuccessful *Agree* applications, respectively. The limit of this tolerance to failed applications of *Agree* must be found in the type of semantic object that can be deduced at LF. This limit is illustrated with the aforementioned restriction in control and impersonal *se* contexts that motivates the present study.

Key-words: *se* constructions, control, *Agree*, thematic theory

RESUMEN: Dentro del marco de una teoría uniforme de las llamadas construcciones con *se* en español, me propongo derivar aquí una restricción que casi no ha recibido atención en la bibliografía previa. En concreto, siempre que una oración de control de sujeto tenga como controlador una instancia de *se* impersonal, la cláusula de infinitivo subordinada no puede contener ninguna otra instancia del clítico *se* con excepción del llamado *se* espurio (e.g., **Se intentó quejarse*, **Se quiso criticarse*, etc.). La fuente de esta restricción se sigue, según argumentaré, de un problema de legibilidad en la Forma Lógica producido, en concreto, por un intento fallido de aplicar *Agree* entre *PRO* y el *se* subordinado, que, como veremos, actúa como sonda para el movimiento-A. Si la explicación que ofrezco es correcta, se sigue también una serie de conclusiones teóricas que afectan directamente el modo en que debemos concebir el diseño de la operación *Agree* en la sintaxis y su efecto en la interfaz de Forma Lógica. En particular, el sistema tolera ciertas fallas de *Agree*

(Preminger 2014) siempre y cuando no afecte cierto tipo de efectos de legibilidad en la semántica. En efecto, la teoría de las construcciones con *se* que aquí asumo deriva la distinción entre *se* paradigmático y no paradigmático como el resultado de aplicaciones exitosas o fallidas de *Agree*, respectivamente. El límite de esta tolerancia a aplicaciones fallidas de *Agree* está en el tipo de objeto semántico que puede deducirse en la Forma Lógica. Dicho límite es ilustrado aquí con la mencionada restricción en contextos de control y *se* impersonal que motiva el presente estudio.

Palabras Clave: Construcciones con *se*, control, *Agree*, teoría temática

1. The “one-or-many” question

The clitic *se* in Spanish, and other Romance languages, occurs in a set of different syntactic-semantic environments (some grammars document 11 or 13 types of *se* depending on the dialect):

Passive *se*:

- (1) a. La policía cerró las puertas para bloquear la salida.
 the police closed the doors for block.INF the exit
 ‘The police closed the doors in order to block the exit.’
- b. Se cerraron las puertas para bloquear la salida.
 SE closed.3PL the doors for block.INF the exit
 ‘The doors were closed in order to block the exit.’

Impersonal *se*:

- (2) a. Juan criticó a Ana.
 Juan criticized DOM Ana
 ‘Juan criticized Ana.’
- b. Se criticó a Ana.
 SE criticized DOM Ana
 ‘One criticized Ana.’

Ergative *se*:

- (3) a. La tormenta hundió al barco.

the storm sank DOM.the ship

‘The storm sank the ship.’

b. Se hundió el barco con la tormenta.

SE sank the ship with the storm

‘The ship sank by the storm.’

Reflexive se

(4) a. Juan criticó a Ana.

Juan criticized DOM María

‘Juan criticized Ana.’

b. Ana se criticó.

Ana SE criticized

‘Ana criticized herself.’

“Aspectual-benefactive” se:

(5) a. Juan comió la manzana.

Juan ate the apple

‘Juan ate the apple.’

b. Juan se comió la manzana.

Juan SE ate the apple

‘Juan ate the apple.’

Inherent se:

(6) a. Juan se quejó.

Juan SE complained

‘Juan complained.’

b. *Juan quejó.

Juan complained

- c. Juan lo quejó.
Juan him complained

Diacritic se:

- (7) a. Juan acordó las condiciones.
Juan agreed the conditions
'Juan agreed the conditions.'
- b. Juan se acordó de Ana.
Juan SE remembered of Ana
'Juan remembered Ana.'

The broad question is this:

The "one-or-many" question:

- (Q) How many clitics *se* does Spanish (and Romance in general) have and which purposes does it /do them serve in the clause?

It is important to make clear one's reaction to such a question from the beginning, since that any stance one takes with respect to it will affect particular analyses for the particular distribution of any occurrence of the clitic *se* and its agreeing variants. In this sense, my answer can be stated in the following form:

- (A) There is just one *se*, serving always the same purpose: deleting unsatisfied subcategorization features encoded on particular functional heads; *i.e.*, the clitic *se* is a pure syntactic expletive (see AUTHOR).

This leads us now to the question of how to account for attested differences in syntactic distribution and semantic interpretation among the “different” types of *se*. On the view to be presented here, such differences must not be attributed to the clitic *per se*, but to the formal make-up of core functional heads, in particular, *v* and T, and to interactions between θ -theory and the theory of *Agree*. I have defended this project in other places. Here I will focus on a particular constraint involving control sentences whose subject is an instance of impersonal *se*. The ban is this:

- (8) Control Ban (CB): A matrix impersonal *se* subject cannot control an infinitival clause containing *any* other instance of *se* (*modulo* spurious *se*).

Here is the crucial paradigm:

- (9) a. ***Se** intentó criticar-*se*.
 SE_{IMP} tried criticize.INF-SE_{REFL}
 INTENDED READING: ‘One tried to criticize oneself.’
- b. ***Se** quiso comer-*se* una manzana.
 SE_{IMP} wanted eat.INF-SE_{BENEF} an apple
 INTENDED READING: ‘One tried to eat an apple.’
- c. ***Se** intentó quejar-*se* menos.
 SE_{IMP} tried complain.INF-SE_{INH} less
 INTENDED READING: ‘One tried to complain less.’
- d. ***Se** intentó castigarse a los corruptos.
 SE_{IMP} tried punish.INF-SE_{IMP} DOM the corrupt.
 INTENDED READING: ‘One tried to punish the corrupt.’

As far as I know, this ban was first discussed in Pujalte's (2012) dissertation, where a concrete proposal is made. Recently it was also discussed in some detail in MacDonald and Vázquez-Lozares (2020a,b). I do not know of any other work in the Spanish generative tradition in which this paradigm is taken into account. In Romance, similar, but clearly not identical, data are explored in Martins and Nunes (2017) for Portuguese. Space reasons prevent me of reproducing the Portuguese paradigm here but it is worth-mentioning that, as already noticed by MacDonald and Vázquez-Lozares (2020a), the Spanish pattern cannot be resolved as a type of identity avoidance, which is essentially the type of solution proposed by Martins and Nunes (2017) for the Portuguese paradigm. This is so because, as shown by MacDonald and Vázquez-Lozares (2020a), control by impersonal *se* of an infinitival clause containing the so-called spurious *se* is perfectly grammatical.

(10) *Se* intentó mandarselo.

SE_{IMP} tried send.INF-SE_{SPURIOUS}-CL_{ACC}

'They tried to send it to him.'

[MacDonald and Vázquez-Lozares (2020a), ex. (10)]

Importantly, in addition to demonstrate that identity avoidance is not what is behind the ungrammaticality of the Spanish paradigm in (9), a sentence like the one in (10) also shows that impersonal *se* can be a good controller to the extent no other "real" *se* clitic occurs in the infinitival complement. That is, this is not a ban against control by impersonal *se*.

Both Pujalte (2012) and MacDonald and Vázquez-Lozares (2020b) have offered different explanations for the relevant paradigm we are concerned with here. I cannot critically

comment on those proposals in such a short paper. In principle, both are incompatible with the general theory of the clitic *se* I will assume here. For instance, according to MacDonald and Vázquez-Lozares (2020b), the source of the ungrammaticality from (9a) to (9c), in which a form of the so-called paradigmatic *se* occurs (reflexive, benefactive/aspectual and inherent) is due to an *Agree* failure. In a few words, impersonal *se* in the matrix clause licenses a type of defective little *pro*. Such a pronoun lacks number features and, consequently, cannot value the unvalued number features of each instance of paradigmatic *se* in the relevant cases. In turn, the source of the ungrammaticality of (9d) must to be attributed to a more spread hypothesis in the literature on impersonal *se/si* in Romance, according to which the impersonal subject of an impersonal *se* sentence (*se* itself or arbitrary *pro*, depending on different approaches) cannot occur in contexts in which nominative is not available either (see Cinque 1988, Dobrovie-Sorin 1998, AUTHOR, Ormazabal and Romero 2019, 2020, among many others). Infinitival complements of control predicates are among the set of embedded sentences that reject nominative.

Regardless of the internal coherence of this type of approach to the CB and its possible compatibility with empirical data, it is clearly incompatible with my more basic assumptions here, in particular, with my assumption that a mere *Agree* failure do not lead to ungrammaticality. As is well-known, solid arguments in favor of failed *Agree* are given in Preminger (2014). I fully concur with Preminger here. In any case, in addition to this, there are many details of MacDonald and Vázquez-Lozares' (2020b) analysis that, as far as I can tell, remain unclear. For example, there is no explicit comment with respect to the mechanism behind control sentences. The claim is that "*Pro*_{se} in matrix context must share its features with *PRO* in the embedded context." (p. 22). That this is the case is, of course, descriptively correct, as attested by simple cases like the following one in which the matrix subject controls

the inflectional features of *PRO*, which, in turn, determines the same features in the inherent clitic *me*.

- (11) Yo quiero *PRO* quejarme.
I want *PRO* complain.INF-ME
'I want to complain.'

The problem is how *PRO* and the matrix subject end sharing the same features. The default hypothesis, once *PRO* is assumed as a primitive, is that the underlying mechanism should be *Agree*. This is extensively argued in Landau (2000, 2004). Putting aside many technical details, the minimal assumption is that *PRO* must enter the derivation with a set of unvalued ϕ -features that are valued by the controller in the main clause. Now, this minimal assumption seems to be incompatible with MacDonald and Vázquez-Lozares' (2020b) approach to impersonal *se* and their assumption regarding the fatality of an *Agree* failure. In effect, if this was the case, then a sentence like (10) would be incorrectly ruled out as an *Agree* failure, since *pro_{se}* would not value the number features of *PRO*. Unfortunately, MacDonald and Vázquez-Lozares do not provide any alternative to the default hypothesis, making the proposal hard to evaluate.

The proposal in Pujalte (2012) is also incompatible with my approach to *se* constructions in general because of her commitment with the PF nature of the clitic *se* and its agreeing variants. The theory I will introduce in the following section share many features with Pujalte's approach but differs precisely in the very nature of the clitic *se*. As I have already advanced, on my view, this clitic is a syntactic, not a PF, expletive. This makes both proposals irreconcilable in many aspects that I cannot discuss in full detail here (see

AUTHOR). I will only briefly mention that under the post-syntactic approach to *se* constructions simple control cases in which impersonal *se* is the controller (see, for instance, (10)) cannot be derived in an obvious way. See, however, Pujalte (2012) for an attempt and extensive discussion.

For all these reasons, I will explore an alternative solution to the CB. Given the short nature of this study, I will keep the ongoing discussion in its simpler form. So, in the following section I will resume my general theory of *se* constructions, according to which the apparently irreducible distinction between paradigmatic and non-paradigmatic *se* can be indeed entirely dissolved, if *Agree* failures do not lead to non-convergent derivations *per se*. I think that the particular empirical domain that *se* constructions instantiate in Spanish makes a strong case for Preminger's *Agree* failure model. Then, in section 3, I handle the CB from this perspective and show that in some restricted and well-defined scenarios certain types of *Agree* failures do lead to non-convergent LFs. Put differently, the CB is derived here as a LF legibility crash. A further important consequence of the paradigm emerging from the CB is that it adds another piece of evidence to dissolve well-known taxonomies for the clitic *se* in Spanish.

2. *Se* as a probe for A-movement: A uniform theory of *se* constructions in Spanish

The uniform theory for *se* I favor has as a first crucial ingredient a difference in the formal makeup of a subset of clitic and of regular lexical phrases. Concretely, I assume that certain clitics, and *se* in particular, are probes for A-movement:

(12) *Thesis 1 (syntax)*: *se* is a probe for A-movement.

In order to have a specific implementation of this thesis, I further assume that clitics are structurally defective: they do not project a Kase phrase. By hypothesis, only K heads can be θ -receptors. Whenever a K head is active in the syntactic derivation, it is also a potential receptor of θ -roles. On this theory, as in many others, more than one θ -role can be assigned to an active KP. Therefore, I conceive of the Θ -Criterion just as the prohibition for an argument to lack a θ -role or as the prohibition for having more than one argument with the same θ -role. Indeed, this latter prohibition is at the heart of my explanation of the CB.

Coming back to the basic ingredients of the theory for *se* I am offering, *Thesis 1*, plus this auxiliary assumption about K heads, gives rise to the following formal difference between *se* and regular lexical phrases:

$$(13) D^{min/max}_{[\phi: \text{unvalued}, \text{EPP}]} \text{ vs. } K^{max}_{[i\phi, \theta]}$$

Note that, like in Chomsky (1995), at least a subset of clitics is taken to be phrasally hybrid, having at the same time properties of phrases and of heads. Consider as illustration the impersonal *se* construction in (2) and the *se* reflexive sentence in (4). In both cases, the clitic performs exactly the same function: it merges with Voice and deletes its subcategorization D-feature. Thus the basic underlying argument structures are identical, namely (< > = deleted features):

$$(14) [\text{VoiceP } se_{[\phi: \text{unvalued}, \text{EPP}]} \text{ Voice}_{\langle D \rangle} [\text{VP criticar Ana}]]$$

What is then the essential syntactic difference between reflexives and impersonals? I contend that it is Abstract Case. As shown again in (15), whereas transitive sentences formed with the

clitic *se* have an accusative direct object, in reflexives the same internal argument surfaces as nominative:

Impersonal *se* vs. reflexives

- (15) a. Se criticó **a Ana.**
 SE criticized DOM Ana
 ‘One criticized Ana.’
- b. **Ana** se criticó.
 Ana SE criticized
 ‘Ana criticized herself.’

Such a difference is syntactically quite radical. For the impersonal derivation, this means that *se* as a syntactic probe cannot attract the internal argument, which is inactivated immediately after its Case feature is valued as accusative (Chomsky 2000, 2001). This obviously results in an *Agree* failure:

Scenario #1: *Agree* failure

- (16) [VoiceP *se*_[φ: unvalued, EPP] Voice_[<D>] [VP criticar [KP Ana_[Case: accusative, theme, iφ: valued]]]]

For the reflexive derivation, absence of Case valuation in the lower domain leaves the internal argument *Ana* active for entering into further A-dependencies. Concretely, *Ana* raises to a position in which it can delete the EPP feature *se* has, value *se*’s inflectional features and get an additional agent θ-role from the Voice head.

Scenario #2: Agree by A-movement:

- (17) [VoiceP [KP Ana_[agent, theme, iφ: valued, Case: unvalued]] *se*_[φ: valued, <EPP>] Voice_[<D>] [VP criticar < [KP Ana_[theme, iφ: valued, Case: unvalued]] >]]

After T (or C, depending on some assumptions) is introduced, *Ana* values its Case feature as nominative. Thus, the theory just sketched reduces the differences between impersonals and reflexives of the relevant type to a simple difference in the Case-*Agree* system in each case. As advanced in the introduction, the theory is committed to tolerate *Agree* failures in the system, as essentially proposed by Preminger (2014). In particular, my analysis of reflexives and impersonals exploits *Agree* successes and *Agree* failures to account for their differences in form and interpretation. On the one hand, successful or failed applications of *Agree* automatically give us paradigmatic and non-paradigmatic instances of *se*. This is self-evident: whenever *Agree* succeeds the form of the clitic will depend on the inflectional features of the lexical subject:

Paradigmatic *se*:

- (18) [VoiceP Ana/yo/vos *se/me/te* Voice_[<D>] [VP criticar <Ana/yo/vos >]]

In contradistinction, as shown in (16), if *Agree* fails then the clitic itself surfaces as third person singular by default.

On the other hand, and this is crucial for the ongoing discussion, successful or failed applications of *Agree* results in two different LF realizations, as well. This is stated as follows:

(19) Thesis 2 (semantics): The LF realization of *se* depends on the syntactic output. Either *Agree* applies in the syntax between *se* and its goal and LF receives the instruction for predicate abstraction or *Agree* fails and, as a consequence, there is no abstraction. If the latter is the case, *se* satisfies the individual argument Voice requires and is realized as an indefinite in Heim’s (1982) sense (probably, under existential closure).

The idea is very simple. The LF correlate of a successful application of the *Agree* operation between *se* and its goals results in an LF in which the clitic itself is a mere index that triggers predicate abstraction as defined, for instance, in Heim and Kratzer (1988):

(20) Predicate Abstraction Rule:

Let α be a branching node with daughters β and γ , where γ dominates only a numeric index i . Then, for any variable assignment g , $\llbracket \alpha \rrbracket^g = \lambda x. x \in D. \llbracket \gamma \rrbracket^{g[i \rightarrow x]}$.

[Heim and Kratzer 1998: 186]

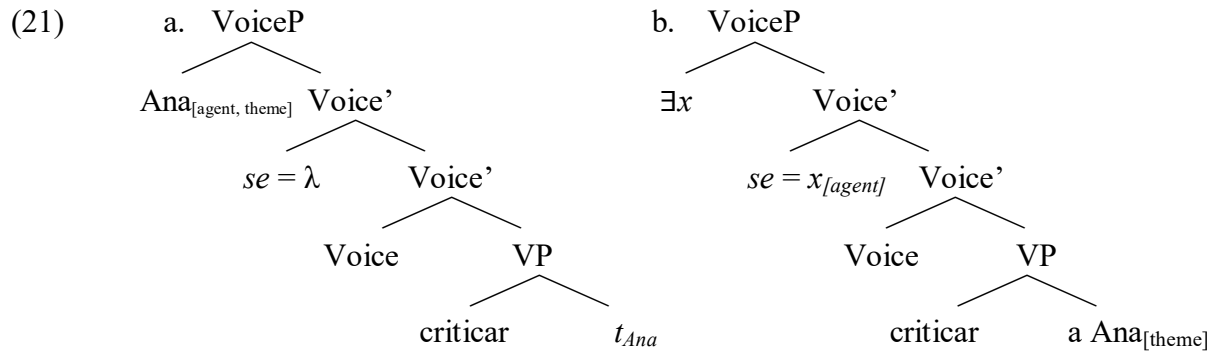
If *Agree* fails, like in the impersonal *se* scenario, then the clitic cannot receive a referential index and, consequently, abstraction is not triggered. In this situation, LF reads the clitic as an indefinite variable in Heim’s (1982) sense (see also Chierchia 2004 and Mendikoetxea 2008).¹ The two LF just commented can be roughly represented with the following two trees:

¹ Existential closure (or whatever other mechanism for getting the impersonal reading is relevant) must, of course, be sensitive to the intervention of other possible operators present in the Syntax-LF. Thus, if generic operators are active, existential closure does not apply and the variable *se* instantiates in that particular case is bound by the relevant operator. As is very well-known at least since Cinque (1988), this particular scenario in which a generic operator intervenes licenses what Cinque called [- argument] *se*, which only occurs in such generic environments. I will come to this distinction in section 3.

- (i) a. Cuando *se* desaparece de esa manera, *se* causan problemas.
when SE disappears of that way SE cause problems
‘When one disappears in that way, troubles are caused.’
b. * Ayer *se* desapareció de repente.
yesterday SE disappeared suddenly
[generic: OK vs. episodic: *]

LF #1: Successful Agree

LF #2: Agree Failure



Summing up the main points made so far, the theory I favor dissolves any particular taxonomy of *se* constructions in Spanish. There is only one *se* in the grammar. Differences among “types” of *se* must not be looked for in the clitic *se per se* but in the formal properties of the clauses in which *se* occurs. If the theory is correct, any occurrence of *se* in the clause (*modulo* the so-called “spurious *se*”, e.g., *Se lo dijo* ‘SE CL.ACC said’) univocally indicates the presence of a syntactic expletive merged with the Voice head. As I have shown in AUTHOR, the theory extends successfully to other cases of paradigmatic and non-paradigmatic *se*. Let me just briefly consider the case of benefactive-aspectual and inherent *se*, whose analyses would be important for the discussion in the next section.²

- (22) a. Juan **se** comió la manzana.
 Juan SE ate the apple
 ‘Juan ate the apple.’
- b. Juan **se** quejó.
 Juan SE complained
 ‘Juan complained.’

² As for passive *se* (see (1)), the other instance of non-paradigmatic *se*, I assume, following AUTHOR and Ormazabal and Romero (2020), that its syntax is the same as impersonal *se*, with agreement differences between both “types” arising at PF.

As for benefactive-aspectual *se*, I assume the simplified structure in (23a), according to which the subject is base-generated as an argument of a high ApplP (see Pylkkänen 2008), which assigns a benefactive θ -role to its argument. The clitic *se* is merged with Voice as already indicated and attracts the benefactive to an extra Spec, VoiceP position in which the benefactive gets an additional agent θ -role from Voice (see (23a)). As for inherent *se*, it instantiates a case in which the verbal root selects a DP, but it does not θ -mark it (*contra* a very well-known assumption in Chomsky 1981; see also Postal and Pullum 1988). Then, the clitic *se* is merged in the already usual way and attracts the internal complement of the verbal root. Again, in its landing position, this argument receives its unique agent θ -role (see (23b)).

- (23) a. [VoiceP Juan_[agent, benefactive, Case: unv., i ϕ : val.] se_[\phi: valued, <EPP>] Voice_[<D>] [ApplP <Juan_[benefactive, Case: unv., i ϕ : val.]> Appl [VP comió la manzana]]]
- b. [VoiceP Juan_[agent, Case: unv., i ϕ : val.] se_[\phi: valued, <EPP>] Voice_[<D>] [VP quejar_[<D>] <Juan_[Case: unv., i ϕ : val.]>]]

I refer the reader to AUTHOR for a detailed justification and further discussion on these and other *se* “constructions”. For the purposes of the next section, these analyses will be enough.

3. A dramatic *Agree* failure

Let’s see now how the present theory accounts for the CB, repeated below:

- (24) Control Ban (CB): A matrix impersonal *se* subject cannot control an infinitival clause containing *any* other instance of *se* (*modulo* spurious *se*).

Recall the basic paradigm:

(25) a. ***Se** intentó criticar-*se*.

SE_{IMP} tried criticize.INF-SE_{REFL}

INTENDED READING: ‘One tried to criticize oneself.’

b. ***Se** quiso comer-*se* una manzana.

SE_{IMP} wanted eat.INF-SE_{BENEF} an apple

INTENDED READING: ‘One tried to eat an apple.’

c. ***Se** intentó quejar-*se* menos.

SE_{IMP} tried complain.INF-SE_{INH} less

INTENDED READING: ‘One tried to complain less.’

d. ***Se** intentó castigarse a los corruptos.

SE_{IMP} tried punish.INF-SE_{IMP} DOM the corrupt.

INTENDED READING: ‘One tried to punish the corrupt.’

I will adopt a simplified *Agree*-based theory of control sentences, like the one proposed by Landau (2000, 2004). As far as I can tell, the simplifications I will make in what follows do not affect the spirit of such a theory. Consider an obligatory subject control sentence as a starting point:

(26) Ana quiere trabajar.

A. wants work.INF


‘Ana wants to work.’

The basic assumption is this: *PRO* enters the derivation with a set of Case and ϕ -features unvalued. With Pesetsky and Torrego (2001), I assume that unvalued features can be interpretable. This is what occurs with *PRO*, whose ϕ -set is interpretable but unvalued. Finally, I assume that *PRO* also enters the derivation with an unvalued referential index. I think that this latter assumption can be seen as a way of interpreting Landau's [- R] feature, *i.e.*, a referential index that depends on the referential properties of the controller in order to get its semantic value. Nothing hinges on any of these concrete implementations, though. The important point, mostly uncontroversial, is that *PRO* does not have inherent, valued ϕ -features. The infinitival clause can be then represented as follows:

(27) [CP *PRO*_[Case: unvalued, $i\phi$: unvalued, Referential Index: unvalued] trabajar]

Somewhat simplifying the set of *Agree* relations that take place after the controller and other functional heads are added to the derivation, I will assume that *PRO* gets all its features valued after an *Agree* relation with the controller:

(28) [Ana ... [CP *PRO*_[Case: Nominative, $i\phi$: valued, Referential Index: 2] trabajar]]



The index 2 is just a convenient way to state that after an application of *Agree* for the *Ana-PRO* pair is done, *PRO*'s referential index must be read as the following assignment function:

(29) $\llbracket g(2) \rrbracket = \text{Ana}$

Again, the reader should take this as a convenient simplification. Using a [-R] feature as in Landau would not affect my main point here. In both cases, we obtain the desired result that the controller of *PRO* will be *Ana* after the said *Agree* relation.

Let's move on and see how our basic pattern is derived under the present theory. I will focus on the impossibility for the impersonal *se* to control an infinitival complement with reflexive *se* in it:

(30) **Se* intentó criticar-*se*.

SE_{IMP} tried criticize.INF-SE_{REFL}

INTENDED READING: 'One tried to criticize oneself.'

Consider first the following derivation step inside the complement clause (RI = Referential Index):

(31) [VoiceP *se*_{[ϕ : unvalued, EPP]] Voice[<D>] [VP criticar *PRO*_[Case: unv., $i\phi$: unv., theme, RI: unv.]]]]}

Here, *se* is a probe and *PRO* is a defective pronominal in the sense already commented above. Now, note that although *PRO* does not possess valued ϕ -features, such features are interpretable. This fact, together with the fact that it is active (*i.e.*, its Case feature is unvalued), renders *PRO* a goal for the probe that *se* instantiates; so, *PRO* moves to a position in which *c*-commands *se*.

(32) [VoiceP *PRO*_{[Case: unv, $i\phi$: unv., theme, agent]] *se*_{[ϕ : unvalued, <EPP>]] Voice[<D>] [VP criticar *t*_{*PRO*}]]}}

As we already know, within the infinitival complement there is another *Agree* failure between *PRO* and *se*, but this time such a failure results in a dramatic legibility problem at LF. Recall that whenever *se* does not get a referential index as a result of *Agree*, it must be read as an indefinite variable (cf. *Thesis 2* in (19)). But if this happens, we end up in a scenario where both *PRO* and *se* are read as the agent of the event. This is a flagrant violation of the Θ -criterion. Crucially, *PRO* and *se* cannot be referentially linked because of the abovementioned *Agree* failure. Absence of *se* in the infinitival complement is grammatical, although depending on the predicate involved is felt as a bit marginal for some. At any rate, the following sentence is perfectly grammatical:

- (35) En este país, nunca *se* quiso castigar
in this country never SE_{IMP} wanted punish.INF
a los corruptos.
DOM the corrupts

Crucially, the syntactic derivation of a sentence like (35) also contains multiple *Agree* failures, but none leads to the same legibility problem at LF observed with cases like (30). As shown in the rough representation in (36), matrix *se* fails to attract a goal and, as a consequence, an *Agree* fails obtains. This is the kind of *Agree* failure we assume derives impersonal/passive *se* in general, so there is nothing new here: a default mechanism repairs the inflectional set *se* encodes. Now, embedded *PRO* also fails to get its features valued, unless it gets its inflectional set valued with *se* after default valuation for *se*. At LF, *PRO*, which bears the agent θ -role, is read as an indefinite variable. Existential closure in the matrix

clause would give us the desired result that both indefinite variables are bound by the same existential operator:

(36) ... *se*_[ϕ : unv., EPP] quiso [*PRO*_[agent, $i\phi$: unv., Case: unv., RI: unv] castigar a los corruptos]

Beyond the implementation details one favors, it is clear that, unlike the CB pattern, no offense to the Θ -criterion arises here.

In order to get a more complete picture of the approach to the CB I defend, let me briefly show now how the same explanation generalizes to the other two cases in (25) involving a paradigmatic *se* in the subordinate clause: aspectual-benefactive *se* (25b) and inherent *se* (25c).

As for benefactive *se*, recall the analysis proposed in (23a) and repeated below:

(37) [_{VoiceP} Juan_[agent, benefactive, Case: unv.. $i\phi$: val.] *se*_[ϕ : valued, <EPP>] Voice_[<D>] [_{AppIP} <Juan_{[benefactive, Case: unv., $i\phi$: val.]>} Appl [_{VP} comió la manzana]]]

Now, like in the reflexive case, we find exactly the same legibility problem at LF in (25b): *PRO* moves to a θ -position, Spec,VoiceP, but crucially fails to agree with *se*, and, consequently, we end up with an illegitimate LF configuration in which *PRO* and *se*, which do not form a referential chain, should be *both* the agent of the event.

- (38) [CP ... *se*_[ϕ : unv., EPP] *quiso* ... [CP [TP *PRO*_[agent, benefactive, $i\phi$: unv., RI: unv, Case: unv.] [VoiceP <*PRO*_{[agent, benefactive, $i\phi$: unv., RI: unv, Case: unv.]]> *se*_[ϕ : unv., <EPP>] Voice_{[AppIP <*PRO*_{[benefactive, $i\phi$: unv., RI: unv, Case: unv.]]>} *comer una manzana*]]]]]}}

And the same illegible LF arises whenever inherent *se* occurs in the infinitival complement.

Recall the proposed analysis in (23b):

- (39) [VoiceP *PRO*_[agent, Case: unv., $i\phi$: unval.] *se*_[ϕ : unvalued, <EPP>] Voice_[<D>] [VP *quejar*_[<D>] <*PRO*_[Case: unv., $i\phi$: unval.] >]]

Now, when we try to embed this type of structure into a subject control configuration in which impersonal *se* occupies the external argument position, *PRO* and embedded *se* are both interpreted as the agent of the subordinate event without forming a referential chain:

- (40) [CP ... *se*_[ϕ : unv., EPP] *intentó* ... [CP [TP *PRO*_[agent, $i\phi$: unv., RI: unv. Case: unv.] > [VoiceP <*PRO*_[agent, $i\phi$: unv., RI: unv. Case: unv.] > *se*_[ϕ : unvalued, <EPP>] Voice_[<D>] [VP *quejar*_[<D>] <*PRO*_[$i\phi$: unv., RI: unv. Case: unv.] >]]]]]

Put it in event semantic terms, the problem for any attempt of controlling a paradigmatic *se* by impersonal *se* can be schematized in the following way:

- (41) *Illegible LF*: $\exists e.[\text{Agent}(\textit{PRO}, e), \exists x \text{Agent}(\textit{se}, e), P(e), \dots]$

So far, I have explained those situations in which we try to control an infinitival complement containing an instance of some paradigmatic *se*. Yet, as we already now, impersonal *se* in the embedded clause is also ruled out:

- (42) **Se* intentó castigarse a los corruptos. .
 SE_{IMP} tried punish.INF-SE_{IMP} DOM the corrupt.
 INTENDED READING: ‘One tried to punish the corrupt.’

For those that believe that *se* requires nominative case, this is of course ruled out for reasons not related to the type of *Agree* failures we are exploring here, but because the embedded *se* is in a clause in which nominative is not available. However, this correlation has at least two important gaps. First, there are non-finite contexts in which nominative is not available and, yet, impersonal *se* is allowed. The case at point is infinitival complements of perception verbs. For many speakers, impersonal *se* is licensed both in Spanish and Italian:

- (43) a. Non ho mai visto spendersi così tanti soldi come quest’ anno.
 ‘I’ve never seen *si* spend so much money as this year.’
 b. Non ho mai visto acquistarsi così tante merci come quest’ anno.
 ‘I’ve never seen *si* purchase so many goods as this year.’

[Cinque 1988: 561, footnote 48]

- (44) a. Nunca he visto bailarse un tango de esa manera.
 ‘I’ve never seen dance a tango in that way.’
 b. Nunca escuché criticarse tanto a alguien.
 ‘I’ve never heard criticize someone so much.’

[AUTHOR]

Cinque's (1988) suggestion is that these are cases of middle *se*, not of impersonal-passive *se*. Yet, this cannot be correct because accusative case is assigned inside the infinitival clause (note the differential object marker in (44b)). One could argue then that at least for some speakers impersonal *se* has the distribution of overt subjects in general, covering thus accusative subjects of ECM-constructions. But again this cannot be on the right track as *causee* subjects in *hacer* causatives cannot be replaced by *se*, as shown with (45b).

- (45) a. Juan hizo a Pedro criticarse, comerse una manzana, quejarse...
 'Juan made Pedro criticize himself, eat an apple, complain...'
 b. Juan hizo castigar(*se) a los culpables. (ok as reflexive/reciprocal)
 'Juan made someone/one punish the culprits.'

For space reasons, I will not discuss here the contrast between causatives and ECM sentences. For detailed discussion on this, see AUTHOR. The second gap is due to an observation also made by Cinque and can be stated in the following way: only impersonal *se* associated to transitive and unergative predicates are licensed in those nonfinite configurations in which nominative case is available, like in the following infinitival absolute clauses.³

Transitives:

- (46) a. Al castigar **el gobierno/ellos** a los culpables, se consiguió la paz.

³ The distribution of impersonal *se* associated to unergatives in Spanish doesn't behave exactly as predicted by Cinque's split. They are degraded when compared with transitives but not so degraded as the rest of [-argument] *se*. E.g., (?)*De trabajarse así... 'if ones works in that way...'. Maybe, impersonal *se* with unergatives responds to a different syntactic configuration (impersonal passives like in Romanian?, see Dobrovie-Sorin 1998). I don't have an answer to this problem at the moment. As far as I know, the issue was not discussed in the literature.

‘Once the government/they punished the culprits, peace was obtained.’

b. Al castigarse a los culpables, se consiguió la paz.

‘Once someone/one punished the culprits, peace was obtained.’

Unaccusatives:

(47) a. Cuando se desaparece de esa manera, se causan problemas.

b. *Al desaparecerse de esa manera, se causan problemas.

‘When one disappears in that way, troubles are caused.’

Passives:

(48) a. Cuando se es condenado sin razón, uno se rebela.

b. *Al serse condenado sin razón, uno se rebela.

‘When one is condemned, one rebels.’

Copulatives:

(49) a. Cuando se es amable, se es aceptado socialmente.

b. *Al serse amable, se es aceptado socialmente.

‘When one is kind, one is socially accepted.’

So, for Cinque, a second split in the *se* construction realm is needed: [+argument] *se* vs. [-argument] *se* (see also footnote 1). In effect, according to Cinque, the distribution of the impersonal *se* sentences formed with unaccusative and passives forces to make a further division inside the non-paradigmatic slot. Crucially, the division is sensitive to the type of predicate involved in the relevant impersonal *se-si* sentence. The impersonal *se* in unaccusative-passive sentences is akin to an expletive that requires syntactic licensing by agreement.

The preceding discussion aims to show that besides initial appearances the distribution of impersonal *se* does not constitute an argument in favor of the nominative *vs.* non-nominative *se* division, at least not straightforwardly.⁴ The way in which the important observations made by Cinque were taken in the subsequent literature followed his main insights. But as is clear from the two gaps briefly commented here, one could take another route, according to which, strictly speaking, *se* is always [- argument], *i.e.*, in the terms of the proposed theory, just a syntactic expletive. As I have shown in AUTHOR, this way the two gaps discussed here can be explained in a rather straightforward manner. At any rate, if I am correct, the source of the ungrammaticality in (42), cannot be attributed to absence of nominative case. One alternative is ruling out this case, exactly in the same way I have ruled out those examples in which the infinitival complement contains some type of paradigmatic *se*. This would amount to forcing the introduction of *PRO* in control clauses. I think this is a plausible alternative in view of the type of algorithm behind the control calculus (see in particular Landau 2004). Yet, this alternative is not forced. Suppose, for instance, that only *se*, not *PRO*, is introduced in the embedded infinitival clause. Roughly, this would suppose two instances of *Agree* failures:

(50) $se_{[\phi: \text{unv.}, \text{EPP}]}$ intentó [_{CP} ... [_{VoiceP} $se_{[\phi: \text{unv.}, \text{EPP}]}$ cerrar ...

Now, recall that according to *Thesis 2* in (19), this situation implies existential closure in the relevant domain. In this case, however, existential closure in the embedded *and* in the matrix clause would amount to blocking subject control and trigger an illegitimate disjoint reference

⁴ It is important to insist in the weakness of Cinque's argument because even nowadays the division is taken as irreducible even for researchers who favor dissolution of particular taxonomies of *se* construction in Spanish and Romance. This is the case of Ormazabal and Romero (2020), who propose dissolving the passive *vs.* impersonal division for non-paradigmatic *se* (see also AUTHOR), but who keep with the nominative *vs.* non-nominative distinction. Yet, the two gaps in Cinque's original division are not accounted for in their work.

reading regarding the matrix and the embedded subject. Thus, the present theory derives the entire Control Ban paradigm as concrete *Agree* failures that create non-convergent LFs.

4. Conclusion

I have conceived of the Control Ban (repeated below) in Spanish as a case in favor of a particular theory of *se* constructions in Spanish and, more generally, as an argument in favor of a particular model for *Agree*, according to which *Agree* failures do not lead to non-convergent outputs.

- (51) Control Ban (CB): A matrix impersonal *se* subject cannot control an infinitival clause containing *any* other instance of *se* (*modulo* spurious *se*).

Agree failures can, however, lead to non-convergent failures, in particular, to legibility problems at LF, whenever other aspects of the clause conspire for such a result. This is precisely what the CB shows in Spanish. In the relevant infinitival complements, *se* attracts *PRO*, but *PRO* itself can never value its own ϕ -features and, consequently, the ϕ -features of *se*. As I have tried to show here, this particular scenario results in an LF in which both *PRO* and *se* receive the agent interpretation, a non-convergent semantic output under any plausible conception of θ -theory or, more generally, of event and argument structure interpretation.

References

Boeckx, Cedric, Norbert Hornstein and Jairo Nunes. 2010. *Control as movement*. Cambridge, Cambridge University Press.

- Chierchia, Gennaro. 2004. A semantics for unaccusatives and its syntactic consequences. In A. Alexiadou, E. Anagnostopoulou and M. Everaert (Eds.), *The unaccusativity puzzle: Explorations of the syntax-lexicon interface* (pp. 22-59). Oxford: Oxford University Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT press.
- Cinque, Guglielmo. 1988. On *si* constructions and the theory of *arb*. *Linguistic Inquiry* 19: 521-581.
- Dobrovie-Sorin, Carmen. 1998. Impersonal *se* constructions in Romance and the passivization of unergatives. *Linguistic Inquiry* 29 (3): 399-437.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. University of Massachusetts dissertation.
- Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Maiden, Mass.: Blackwell.
- Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30: 69-96.
- Landau, Idan. 2000. *Elements of control: Structure and meaning in infinitival constructions*. Dordrecht: Kluwer.
- Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22:811–877.
- MacDonald, Jonathan E. and Almike Vázquez-Lozares. 2020a. On (un)grammatical sequences of *ses* in Spanish. Ms.
- MacDonald, Jonathan E. and Almike Vázquez-Lozares. 2020b. Spanish impersonal *se* in control infinitivals and the ungrammaticality of *se se* sequences. Ms..
- Martins, Ana María and Jairo Nunes. 2017. Identity avoidance with reflexive clitics in European Portuguese and the theory of control. *Linguistic Inquiry* 48:4. 627-649.

- 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.
- Ormazabal, Javier and Juan Romero. 2019. The formal properties of non-paradigmatic *se*. *Borealis: An International Journal of Hispanic Linguistics* 8(1): 55-84.
- Ormazabal, Javier and Juan Romero. 2020. Deconstructing *se* constructions: number agreement and post-syntactic variation. Ms..
- Pesetsky, David and Esther Torrego. 2001. T-to-C Movement: Causes and Consequences. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 355–426 Cambridge, Mass.: MIT press.
- Postal, Paul and Geoffrey Pullum. 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* 19(4). 635–670.
- Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, Mass.: MIT press.
- Pujalte, Mercedes. 2012. *Argumentos (no) agregados. Indagaciones sobre la morfosintaxis de la introducción de argumentos en español. Phd diss.*, University of Buenos Aires.
- Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, Mass.: MIT press.