

**COPULAR INVERSION AND NON-SUBJECT  
AGREEMENT**

Alex Alsina and Eugenio M. Vigo  
Pompeu Fabra University

Proceedings of the LFG14 Conference

Miriam Butt and Tracy Holloway King (Editors)

2014

CSLI Publications

<http://csli-publications.stanford.edu/>

## Abstract

In this paper we analyze the phenomenon of copular inversion (CI) in Catalan, which consists in the copula not agreeing with its subject but with its complement. We claim this is due to the idea that in CI-languages verbs agree with one of their cosubjects, i.e. the GFs that are coreferential with the subject, including the subject itself. In the case of copular sentences, there are two cosubjects the verb may agree with, namely the subject and the predicate; which one the verb agrees with is determined by a set of OT constraints that implement a Person-Number Hierarchy so that the copula agrees with the most marked cosubject. This theory does not affect the analysis of subject-verb agreement in non-copular sentences, as in such cases there is only one cosubject available, the subject itself. In order to integrate these ideas within an LFG framework, we propose a change in how verbal agreement is formalized: we claim that verbs do not specify the person and number features of the subject, but have a special structure, AGR, specifying their own features. An f-structure constraint requires this AGR to be unified with a GF in the sentence, the choice being licensed through an OT ranking of constraints.

## 1 Introduction

Copular sentences show a puzzling situation with respect to subjecthood properties when the predicative complement is a DP in Catalan and closely related languages such as Spanish and Italian. In some cases, the preverbal DP agrees with the copula, as in (1). All linguistic data in this paper are from Catalan, unless otherwise noted.<sup>1</sup>

- (1) a. Els impostos són el problema  
*the.pl tax.pl be.pres.3.pl the.sg problem.sg*  
'The taxes are the problem.'
- b. \*Els impostos és el problema  
*the.pl tax.pl be.pres.3.sg the.sg problem.sg*

This is an unsurprising situation, given that, as will be argued later, the preverbal DP is assumed to be the subject in the absence of an agreeing clitic or marked intonation. However, we also find situations in which it is the postverbal DP that agrees with the copula, as in (2):

- (2) a. El problema són els impostos  
*the.sg problem.sg be.pres.3.pl the.pl tax.pl*  
'The problem is taxes.'

---

<sup>†</sup>We gratefully acknowledge the observations made by the reviewers and the audience of the 19th LFG Conference. The research presented in this paper is supported by research project FF2011-23046.

<sup>1</sup>Gender information is not glossed in the examples in order to keep the glosses as compact and simple as possible. Only the features relevant for our discussion are annotated.

- b. \*El problema és els impostos  
*the.sg problem.sg be.pres.3.sg the.pl tax.pl*

In examples like (2) not only can the copula agree with the postverbal DP, but it cannot agree with the preverbal DP, as in (2b). We refer to the construction illustrated in (2a) as *copular inversion*, or CI.

In this paper we build on the claim in Alsina (2007) that the agreeing postverbal DP in CI is not the subject, but the complement of the copula. But we depart from Alsina (2007) in claiming that in CI the verb really agrees with the complement and that CI is not an instance of covert subject agreement. This conclusion forces us to abandon the standard LFG idea that, in their lexical entries, finite verb forms specify the grammatical function (GF) that the verb agrees with. We propose that verb forms lexically specify a set of agreement features, without tying them down to any specific GF. It is through the interaction of OT-like constraints that this set of features is required to belong to a particular GF. In languages like Catalan, these constraints allow a non-subject as agreement trigger only in copular sentences. In other languages, such as English, the same constraints, in a different ranking, prohibit non-subject agreement completely, whereas in yet other languages different rankings of the same constraints account for a much more general distribution of non-subject agreement.

In what follows, we first review the evidence from Alsina (2007) for the claim that, in CI, the agreeing postverbal DP is not the subject. We then briefly discuss the analysis of CI proposed in Alsina (2007), which makes the claim that CI is only possible if the construction has a null subject. We provide evidence that this claim is incorrect. We propose an analysis of CI in which the verb agrees with its complement: it is, in fact, a new approach to highest argument agreement that derives subject-verb agreement as a particular case and the most frequent one, but by no means the only one. To conclude, we show that this approach can be straightforwardly applied to other instances of non-subject agreement.

## 2 The properties of the postverbal DP

The fact that the postverbal DP in CI agrees with the verb, as shown in (2), is strong *prima facie* evidence in favor of the subject status of the postverbal DP. Typical examples of copular sentences where the verb agrees with the postverbal DP, when both DPs are 3rd person and the postverbal DP is plural, are found when the preverbal DP denotes an abstract entity, such as ‘the problem’, as in (2), ‘the solution’, ‘the reason’, ‘the best thing’, etc. But it is also found when the preverbal DP denotes a concrete entity, as in (3):

- (3) a. El meu joc preferit són/\*és els escacs  
*the.sg my.sg game.sg favorite.sg be.pres.3.pl/\*sg the.pl chess.pl*  
 ‘My favorite game is chess.’

- b. El sopar d'avui són/\*és verdures a la  
*the.sg dinner.sg of=today be.pres.3.pl/\*sg vegetable.pl on the*  
 planxa  
*grill*  
 'Today's dinner is grilled vegetables.'
- c. Aquesta cadira són/\*és quatre fustes mal clavades  
*this.sg chair.sg be.pres.3.pl/\*sg four wood.pl badly nailed.pl*  
 'This chair is a bunch of pieces of wood poorly nailed together.'

Examples (3b) and (3c) show that the postverbal DP does not have to be more specific (or more restrictive in its reference) than the preverbal DP in order for it to agree with the verb. The copula never agrees in the 3rd person singular with the postverbal DP when the preverbal DP is 3rd person plural, as in (1b). When one of the DPs involved in a copular sentence is a first or second person pronoun, whereas the other one is a DP unmarked for person, the former attracts agreement, regardless of its position with respect to the verb:

- (4) a. L'autor sóc/\*és jo  
*the=author.sg be.pres.1.sg/\*3.sg I*  
 'I am the author.'
- b. Jo sóc/\*és l'autor  
*I be.pres.1.sg/\*3.sg the=author.sg*  
 'I am the author.'
- (5) a. Els candidats més ben valorats éreu/\*eren  
*the.pl candidate.pl more well valued.pl be.impf.2.pl/\*3.pl*  
 vosaltres  
*you.pl*  
 'You were the most highly valued candidates.'
- b. Vosaltres éreu/\*eren els candidats més ben  
*you.pl be.impf.2.pl/\*3.pl the.pl candidate.pl more well*  
 valorats  
*valued.pl*  
 'You were the most highly valued candidates.'

Despite what the phenomenon of agreement seems to indicate, all of the other subjecthood properties examined by Alsina (2007) argue for analyzing the postverbal agreeing DP in the CI construction as a complement, not the subject of the verb. The relevant properties are: the evidence from the position of the DP, the impossibility of omitting it, and its behavior with respect to the partitive clitic. (Alsina, 2007 also examines the facts of raising, which point to the same conclusion, although we will not review them here for space reasons.)

In Catalan, as in the other Romance languages that allow null subjects, there is a positional asymmetry between subject and complements: the apparent subject<sup>2</sup> can appear either preverbally or postverbally, whereas complements (by which we mean direct and indirect objects and obliques) are restricted to appearing postverbally. (6) illustrates the alternative position of the apparent subject, while the contrast between (6) and (7) shows the fixed postverbal position of complements. All orders in (7) are ungrammatical with a neutral intonation: a marked intonation of the preverbal elements would render the orders in (7) grammatical, but that would imply a different syntactic structure in which neither of the preverbal DPs occupies this position by virtue of being a complement of the verb, but as a focused phrase.

- (6) a. Els ruixats seguiran la tempesta  
*the.pl shower.pl follow.fut.3.pl the.sg storm.sg*  
 ‘The showers will follow the storm.’  
 b. Seguiran la tempesta els ruixats  
*follow.fut.3.pl the.sg storm.sg the.pl shower.pl*  
 ‘The showers will follow the storm.’
- (7) a. \*La tempesta seguiran els ruixats  
*the.sg storm.sg follow.fut.3.pl the.pl shower.pl*  
 b. \*Els ruixats la tempesta seguiran  
*the.pl shower.pl the.sg storm.sg follow.fut.3.pl*  
 c. \*La tempesta els ruixats seguiran  
*the.sg storm.sg the.pl shower.pl follow.fut.3.pl*

Given this subject/non-subject asymmetry, it follows that, in a sentence like (2a), the preverbal DP can be nothing but the apparent subject and, consequently, the postverbal DP has to be the complement, even though it agrees with the verb.

An additional argument having to do with position is that, in some cases, the agreeing DP cannot be placed in the preverbal position, which is an expected position for a subject. In many cases, the order of the two DPs in a copular sentence can be reversed, without affecting the agreement of the verb, as shown in (1)–(2) and in the pairs of sentences of (4) and (5), or the truth-conditional meaning of the sentences, though the information-structure is different. But, in some cases, the order of the two DPs cannot be reversed. This is what we see if we try to reverse the order of the DPs in (3b)–(3c):

- (8) a. \*Verdures a la planxa són/és el sopar  
*vegetable.pl on the grill be.pres.3.pl/sg the.sg dinner.sg*  
 d’avui  
*of=today*

---

<sup>2</sup>We use the term *apparent subject* to be neutral with respect to whether the preverbal DP in sentences like (1) is a true subject or a topic that anaphorically binds a null pronominal subject. The latter analysis is supported by a lot of evidence and has been argued for in Bonet (1990), Solà (1992), and Vallduví (1992), among others, but it is by no means unanimously accepted, cf. Forcadell (2013).

- b. \* Quatre fustes mal clavades són/és aquesta cadira  
*four wood.pl badly nailed.pl be.pres.3.pl/sg this.sg chair.sg*

If the agreeing DP were the subject, it should have no difficulty in appearing in preverbal position. The fact that, in some cases, it cannot strongly suggests that it is not the subject.

Another property that reveals the non-subject status of the postverbal agreeing DP in CI is the impossibility of omitting it. Catalan, as a null subject language, freely allows the subject to be omitted in any sentence (subject to pragmatic constraints). So, if we take a sentence like (1a), we can leave out the preverbal DP and the sentence is interpreted as if it had a pronominal subject coreferential with a discourse topic, possibly *els impostos* ‘taxes’, as in (9a). In contrast, we cannot leave out the agreeing postverbal DP of a CI sentence: (2a) becomes ungrammatical if we leave out the agreeing DP, as in (9b), although it is fine without the non-agreeing DP, as in (9c).

- (9) a. Són el problema  
*be.pres.3.pl the.sg problem.sg*  
 ‘They are the problem.’  
 b. \* El problema són  
*the.sg problem.sg be.pres.3.pl*  
 c. Són els impostos  
*be.pres.3.pl the.pl tax.pl*  
 ‘It is taxes.’

The same occurs when the agreeing postverbal DP is a first or second person pronoun, as in (4)–(5). Corresponding to (4a), we cannot omit the agreeing DP, although we can omit the non-agreeing DP:

- (10) a. \* L’ autor sóc  
*the=author.sg be.pres.1.sg*  
 b. Sóc jo  
*be.pres.1.sg I*  
 ‘It is me.’

If the agreeing DP in CI were to be analyzed as the subject, the fact that it cannot be elided would be unexplained. If we analyze it as an obligatory complement of the verb, the fact that it cannot be left out follows from this.

Finally, the partitive clitic *en* may correspond to a nounless DP only when this DP is a complement, never when it is only a subject.<sup>3</sup> The clitic *en* is required in

<sup>3</sup>As is well-known, the single nominative GF of many verbs —unaccusatives— behaves like an object in many respects, and that includes the possibility of having a partitive clitic related to it. Thus, an expression that looks like a subject because it agrees with the verb may have a partitive clitic, as in *N’arriben molts* (*EN arrive.3.pl many*) ‘Many of them are arriving’, but it is reasonable to analyze this expression both as a subject and an object (Alsina, 1995). This is not the situation in copular sentences, where there are two nominative GFs.

(11b) because the direct object lacks the head noun, but it cannot be used in (12), where it corresponds to the missing noun of the subject:<sup>4</sup>

- (11) a. Els estudiants llegeixen molts llibres  
*the.pl student.pl read.pres.3.pl many.pl book.pl*  
 ‘The students read many books.’  
 b. Els estudiants \*(en) llegeixen molts  
*the.pl student.pl \*(EN) read.pres.3.pl many.pl*  
 ‘The students read many (books, magazines, ...).’
- (12) Molts (\*n’)aprovaran  
*many.pl \*(EN)=pass.fut.3.pl*  
 ‘Many will pass.’

In copular constructions, it is the postverbal DP that triggers the presence of the partitive *en* when it lacks a head noun, even if it is the agreeing DP.

- (13) a. El problema són molts estudiants  
*the.sg problem.sg be.pres.3.pl many.pl student.pl*  
 ‘The problem is many students.’  
 b. El problema \*(en) són molts  
*the.sg problem.sg \*(EN) be.pres.3.pl many.pl*  
 ‘The problem is many of them.’

Even though the postverbal DP agrees with the verb in (13b), it triggers the presence of the partitive clitic, which indicates it is a complement.

To summarize, the evidence from position, omissibility, and *en*-cliticization indicates that the postverbal DP in CI is a complement. This leaves the agreement facts as a problem that needs to be solved.

### 3 Previous analyses

Previous analyses of CI have assumed that the general process of subject-verb agreement applies in this construction. Many authors, such as Hernanz and Bruccart (1987), Fernández Leborans (1999), and Ramos (2002), have assumed that the agreeing DP is the subject, without providing an adequate explanation for the facts presented in Section 2 that argue for the complement status of the agreeing DP in CI. Alsina (2007), however, breaks with this tradition and assumes that this DP is a complement and explains the agreement facts by positing a covert (or null) subject in CI which shares its features with the postverbal DP, thus preserving the idea that

<sup>4</sup>Clitic *en* has a second function (in addition to supplying the restrictor of an OBJ): to supply the complement of the OBJ. For this reason, *Molts en llegeixen els llibres* is grammatical with *en* corresponding to the complement of the OBJ *els llibres* (hence, the meaning of the sentence: ‘Many read his/her books’). It does not correspond in any case to the missing noun of the SUBJ.

the verb agrees with its subject even in constructions that appear to contradict this idea. Some aspects of this analysis are shared with Moro (1997), but, for ease of exposition, the rest of this section will discuss the analysis in Alsina (2007).

Alsina (2007) assumes, as in Vallduví (1992) and others, that finite clauses have an optional iterative position at the front of the clause that can be occupied by an XP that is a topic anaphorically connected to a pronoun inside the clause, possibly in a long-distance relation. This pronoun can be a clitic, in which case we get what is known as clitic left dislocation, or a null subject, the claim being that null subjects are pronominal, so that what appears to be a sentence with a clause-initial subject should be analyzed as a sentence with an initial topic DP that binds a null pronominal subject either in the same clause or at a deeper level of embedding. The f-structure conditions and the c-to-f-structure mapping principles in Alsina (2007) are set up so that a subject can only be overtly expressed as a VP-internal phrase (an XP inside the VP). If the subject is not realized in this position, there is no phrase corresponding to the subject, but a PRED ‘PRO’ feature is provided for the subject, allowing it to satisfy the Completeness condition. In this case, there may be a topic DP in clause-initial position that anaphorically binds the pro subject, giving the impression that it is the subject.

This analysis, according to which subjects are either postverbal DPs or null pronominals, applies in all clauses in Catalan, including copular clauses. In this type of clause, a special situation arises. As a lexical property of the copula, the subject and the complement are coindexed.<sup>5</sup> From a syntactic point of view, two coindexed expressions are required to have the same agreement features, i.e. person, number, and gender. Two expressions in an anaphoric binding or control relation are also coindexed, which implies that the pronoun and its antecedent have the same agreement features. Thus, in a copular construction in which there is a null subject bound by a topic, the subject is coindexed with both the topic and the predicative complement and, so, needs to have the same agreement features as both of these expressions.

A null subject has no lexically specified features and therefore is free to “copy” whatever features are present in both coindexed expressions. No conflict arises if both the coindexed topic and the predicative complement happen to have the same agreement features. But a conflict does arise when the two expressions have different features—suppose the topic that binds the null subject is first person singular and the predicative complement is third person singular. Assuming that Consistency, or Uniqueness, cannot be violated, the subject cannot be both first person and third person. In such cases, a set of OT constraints has the effect of requiring the subject to choose, out of the two sets of features of the coindexed expressions, the one that is highest in the following Person-Number Hierarchy:

$$(14) \quad 1/2 \text{ SG} > 3 \text{ PL} > 3 \text{ SG}$$

---

<sup>5</sup>The subject is referentially a subset of the complement in a copular construction, either a proper subset or the same set. We use the term *coindexed* in this paper to refer to both situations.



Given this, it does not matter whether it is the topic anaphorically linked to the subject or the predicative complement that has the features that are higher in the hierarchy in (14): these are the features that the subject will pick and that will have a morphosyntactic effect on the form of the copula. This explains why, when one of the two GFs involved is first or second person and the other one is not, the copula shows agreement with the first or second person GF, regardless of which one is the topic and which one the complement, as shown in (4)–(5), or, when one is third person plural and the other one is third person singular, the copula shows agreement with the former, as shown in (1)–(2) and (3).

This analysis correctly explains the agreement facts in copular sentences with a preverbal and a postverbal DP, while preserving the observation made in Section 2 of this paper that the postverbal DP in CI is a complement even though it agrees with the verb. But the analysis makes the claim that CI is only possible in languages and in constructions with null subjects. As we shall see now, this claim is not correct.

#### 4 Problems with Alsina (2007)

The main problem with the approach defended by Alsina (2007), as well as Moro (1997), is that it claims that CI is possible only in *pro*-drop languages and, within languages of this kind, in constructions that have a null subject. However, we will show this is not true, as there is CI in Catalan copular constructions in which no null subject is possible and there is data that shows the existence of CI in a non-*pro*-drop language like German.

Consider the following examples in Catalan, where a VP-internal subject is found:<sup>6</sup>

- (15) a. Són/\*és            els    impostos un    problema  
           *be.pres.3.pl/\*sg the.pl tax.pl    a.sg problem.sg*  
           ‘Taxes are the problem.’
- b. Són/\*és            un    problema els    impostos  
           *be.pres.3.pl/\*sg the.sg problem.sg a.pl tax.pl*  
           ‘The problem is taxes.’

These sentences do not have a null subject. According to the sentence structure defended by Alsina (2007) and Vallduví (1992), a null pronominal subject is required to comply with the Subject Condition if there is no postverbal DP corresponding to the subject and there is no displaced constituent (such as an interrogative DP) that fills the subject through structure-sharing. In such cases, a null pronominal subject

<sup>6</sup>In these examples, we have decided to give data that use the indefinite article for *problema* instead of the definite article which we have been using throughout this paper, as we have noticed that the latter produces some unwanted effects that seem to be due to the information structure of the sentence, as in example ??*Són el problema els impostos* ‘The problem is taxes.’ This seems to be constant regardless of which nouns are involved.

is assumed to be present at f-structure: this null pronoun can be anaphorically dependent on a discourse topic not present in the sentence, as in (9a) and (9c), or on a discourse topic expressed as a preverbal DP, as in examples (1) through (5). In the latter class of examples, the preverbal DP is assumed not to be the subject, but a topic that is the antecedent of the null subject, hence the name “apparent subject” used for this DP.

It seems to be irrelevant to know which of the two DPs is the subject. The data in (15) show that regardless of word order, the copula must agree in the plural, thus yielding exactly the same paradigm we have shown for CI: the plural wins over the singular.

In fact, Alsina (2007) predicts that in cases like (15) the verb should agree with its subject. Leaving aside the problem that it is quite difficult to determine which DP is actually the subject, the data presented above show that this prediction is wrong. If the verb agreed with its subject, there are two possible situations: if we assume a flat structure where either DP may be assigned the SUBJ function, the verb should be able to agree with either DP, or if we assume a strict interpretation of Vallduví’s (1992) theory, which proposes a VOS canonical word order for Catalan, then the final DP should be regarded as the subject, thus leaving (15a) unexplained.

As has been already mentioned above, there is further data showing that approaches that explain CI by means of null subjects are incorrect. Consider these data from German, for instance:

- (16) a. Die Steuern sind/\*ist das Problem  
*the.pl tax.pl be.pres.3.pl/\*sg the.sg problem.sg*  
 ‘Taxes are the problem.’
- b. Das Problem sind/\*ist die Steuern  
*the.sg problem.sg be.pres.3.pl/\*sg the.pl tax.pl*  
 ‘The problem is taxes.’

In (16) we see exactly the same paradigm that is shown by Catalan: the copula agrees with the element that is higher in the Person-Number Hierarchy, regardless of its position.

If we follow Berman (2003), one could argue that German is a language that allows for a certain freedom of word order, such that any DP may be placed in Spec-CP (the so-called *Vorfeld*) in order to be focalized. This could lead to the false impression that in (16b) *das Problem* is a focalized predicate, such that the subject, *die Steuern*, remains internal to the VP. The main issue with this hypothesis is that it does not explain why singular agreement is barred in (16a) which should be the consequence if we consider *die Steuern* a focalized predicate, or why singular agreement is ungrammatical in (16b), which should be the consequence if *das Problem* was a focalized subject.

We conclude, therefore, that German does indeed have CI phenomena. Since German is not a *pro*-drop language, it is clear that the analysis of CI cannot depend on the presence of a null subject. The paradigm shown by (15) and (16) is explained

in a unified way if we assume a CI analysis, where the Person-Number Hierarchy is the key factor determining the agreement features, such that only plural agreement is possible because it ranks higher than singular agreement.

In summary, the data presented in this section imply that any approach using null subjects as some kind of “proxy” that “copies” agreement features from a non-subject to the subject, like Moro (1997) and Alsina (2007), is incorrect.<sup>7</sup> CI, as we will show below, is about two GFs competing for agreement with a verb whose arguments are in a very particular relationship. Whereas in transitive sentences the object does not share its reference with the subject, in copular sentences the predicate (the “object”) is indeed coreferential with the subject of the sentence. The proposal we defend here is that the copula is not different from any other verb with respect to agreement; this special relationship between the subject and complement of the copula unravels the nature of verbal agreement, namely, that all verbs seek to agree with a GF that is coreferential with the subject. Of course, copular sentences are the only case where two GFs share their reference, thus allowing the copula to agree with the predicate.<sup>8</sup>

## 5 Proposed analysis

Our claim is that in some languages, the verb does not actually agree with its subject, but with some GF that is coreferential with the subject, which may, of course, be the subject itself. We call the GFs that comply with these conditions cosubjects (COSUBJS), more formally:

- (17) A GF  $f$  is a COSUBJ iff  $ref(f) = ref(g)$ , where  $g$  is SUBJ and  $ref(x)$  is a function that returns the reference of some GF  $x$ .

If  $f$  is the subject, then  $f = g$ , thus  $ref(f) = ref(g)$  by definition. Therefore, the subject is always considered a cosubject and, in fact, in non-copular sentences it is the only one available.

In a copular sentence, of course, there are at least two cosubjects, the subject and the predicate of the sentence. This means that the verb must somehow choose which GF it agrees with, a choice that is determined by the Person-Number Hierarchy proposed earlier. The problem is that the aforementioned hierarchy is not a principle by itself, but needs to be integrated in a theory of syntax that predicts in which situation the verb will agree with a non-subject and in which it will agree with a subject. The statement that the verb must agree with the cosubject ranking higher in the hierarchy is just an informal description. In order for this statement

---

<sup>7</sup>This kind of analysis resembles the approach defended by Perlmutter (1983) for unaccusative constructions, in which the verb agrees with a non-subject postverbal DP triggered by the presence of a dummy null subject.

<sup>8</sup>There is one exception to this claim, namely reflexive objects. Such cases, though, pose no problem for our theory, as will be shown later.

to be incorporated in a formal LFG theory of agreement, we need to change some aspects of the LFG framework in a substantive way, as we shall see next.

In the first place, we assume that verb forms specify agreement features in a feature structure called AGR as shown in Figure 1 (see next page).

$$\left[ \begin{array}{ll} \text{PRED} & \text{'be} \langle \text{arg(ument)} \rangle \\ \text{TENSE} & \text{PRES} \\ \text{MOOD} & \text{INDIC} \\ \text{AGR} & \left[ \begin{array}{ll} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \end{array} \right] \end{array} \right]$$

Figure 1: Lexical entry for Catalan *són* ‘(they) are’

The AGR of the verb must be unified with the AGR of some other GF according to the well-formedness condition GF-VERB AGREEMENT shown below:<sup>9</sup>

$$(18) \text{ GF-VERB AGREEMENT: } \left[ \begin{array}{ll} \text{AGR} & \boxed{\text{I}} \\ \text{GF}_j & \boxed{\text{I}} \end{array} \right]$$

The condition in (18) must be fulfilled by the f-structure of the sentence. In a language with CI,  $\text{GF}_j$  is a cosubject, whereas in languages without this phenomenon, it is just the subject.

It is easy to understand how this is a radical departure of what may be called “standard” LFG (Bresnan, 2001, Dalrymple, 2001, among others). In previous theories, verbal agreement is formalized by means of a functional annotation of the verbal lexical entry that specifies the agreement features of the SUBJ. These features have to unify with those of the DP in subject position. However, the data we have presented are not compatible with this approach, as the GF the verb agrees with cannot be determined at a lexical level. For instance, a form like *són* ‘they are’ agrees with the subject in (1a) and with the complement in (2a).

For a sentence like (2a), the c-structure and f-structure are those shown in Figures 2 and 3 (next page), respectively.<sup>10</sup> These assume that the preverbal DP is effectively a topic, following Alsina (2007) and Vallduví (1992), but it is perfectly possible to represent both structures by assuming that the preverbal DP is mapped to a SUBJ, for example, if one follows theories that reject VOS as the “neutral”

<sup>9</sup>In languages where a single verb form simultaneously agrees with SUBJ as well as with other arguments (e.g. Basque, Swahili, etc.), the verb specifies information about OBJ (and possibly other non-subject arguments) as well as AGR.

<sup>10</sup>The correspondence between arguments at a-structure and GFs at f-structure is not governed by the standard Completeness and Coherence conditions, but by principles such as those proposed by Alsina (2007, pp. 32–33), including the Subject Condition, which licenses a subject in the f-structure, even without a corresponding argument at a-structure.

word order in Catalan, e.g. Forcadell (2013). The only difference, besides the obvious ones at the level of c-structure (the preverbal DP would be Spec-IP, not an adjunct of IP), would be that there is no instance of a null subject at the f-structure.

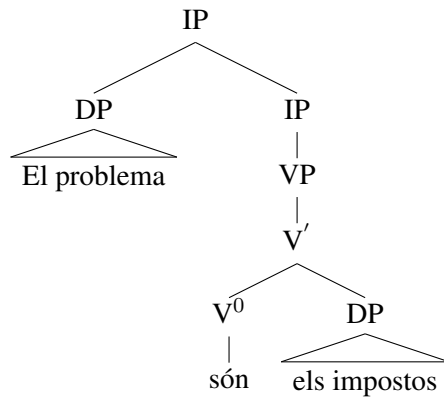


Figure 2: c-structure of (2a)

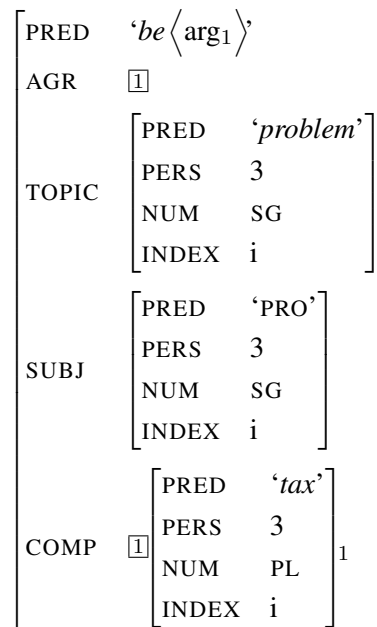


Figure 3: f-structure of (2a)

As stated so far, we still do not have any mechanism that determines which GF the copula must agree with. We have the Person-Number Hierarchy, but, as noted before, in its current fashion it is nothing but a useful descriptive device, yet devoid of any explanatory power as it is not integrated into a theoretical framework that answers the questions of why and how CI arises.

To do so, we propose an implementation of the Person-Number Hierarchy in an OT-LFG framework (Bresnan, 2000; Kuhn, 2003). OT-LFG operates under the standard assumptions of LFG that syntax is the product of interactions between different levels of grammatical information that are independent of each other, but it adds the OT hypothesis that grammaticality is actually the product of a competition between different possible candidates, such that the grammatical one is the one that violates the least important constraint the least amount of times (Prince and Smolensky, 2004). From a philosophical stand, using OT-LFG embodies a vision of language that states that the forms that are deemed part of the language (i.e. grammatical) are actually those that are the “less defective” ones with respect to different criteria represented in the structures of the LFG architecture as a result of the complex interaction of potentially clashing tendencies that are found within a language, across languages and in human language in general.

The OT constraints we propose are the following ones, ranked as in (22) and in (23) for CI and non-CI languages, respectively:

- (19) COSUBJAGR:  $GF_j = \text{COSUBJ}$
- (20) SUBJAGR:  $GF_j = \text{SUBJ}$
- (21) MARKEDAGR:
- a. AGRPERS:  $\left[ \text{AGR} \left[ \text{PERS} \ 1 \vee 2 \right] \right]$
  - b. AGRNUM:  $\left[ \text{AGR} \left[ \text{NUM} \ \text{PL} \right] \right]$
  - c. AGRPERS  $\gg$  AGRNUM
- (22) For CI languages:  
 COSUBJAGR  $\gg$  MARKEDAGR  $\gg$  SUBJAGR
- (23) For non-CI languages:  
 SUBJAGR  $\gg$  MARKEDAGR  
 SUBJAGR  $\gg$  COSUBJAGR

The way to implement the Person-Number Hierarchy is the constraint “bundle” MARKEDAGR with its two internal constraints: agreement in the 1st or 2nd person plural means no violation of MARKEDAGR; agreement in the 1st or 2nd person singular, one violation against AGRNUM; agreement in the 3rd person plural, one against AGRPERS; and agreement in the 3rd person singular, two against both constraints that make up MARKEDAGR.<sup>11</sup>

Given that MARKEDAGR favors verb forms in the 1st and 2nd person, why don’t we always get verb forms with these features? The reason is that the verb’s AGR must unify with a GF in the clause, as required by (18) and it must do so preserving Consistency, which we consider to be an inviolable principle. In addition, the sentence must satisfy either COSUBJAGR or SUBJAGR, in CI or non-CI languages, respectively.

The ranking in (23) for non-CI languages may require further explanation. This device, formalized by Anttila and Cho (1998) and used in OT-LFG by Belyaev (2013), expresses an underspecified ranking where the relative order between the MARKEDAGR and COSUBJAGR constraints is irrelevant and is, therefore, left undetermined, as it does not affect the predictions about verb agreement in non-CI languages. The only relevant ranking in this case is that SUBJAGR must be higher than MARKEDAGR and COSUBJAGR.

Let us now explain the basic facts of CI making use of this theory. For cases like those in (2), repeated here for convenience, the optimization is the one shown in Table 1:

---

<sup>11</sup>Instead of MARKEDAGR one could have a disjunction of four constraints as suggested by a reviewer, but not only would it be extremely inelegant; it would not be able to explain the facts in (26) and (27).

- (24) a. El problema són els impostos  
*the.sg problem.sg be.pres.3.pl the.pl tax.pl*  
 ‘The problem is taxes.’
- b. \*El problema és els impostos  
*the.sg problem.sg be.pres.3.pl the.pl tax.pl*

	COSUBJAGR	MARKEDAGR		SUBJAGR
		AGRPERS	AGRNUM	
☞ (24a)		*		*
(24b)		*	*!	

Table 1: Optimization for (24)

The explanation is quite straightforward in this case. Neither candidate violates COSUBJAGR, as in both cases the verb agrees with a cosubject: the complement in (24a) and the subject in (24b). Both violate AGRPERS once, as in both cases agreement is in the 3rd person, not in the 1st or the 2nd. The constraint that rules out (24b) is AGRNUM: this constraint is violated in (24b) because the verb is in the singular (i.e. its AGR would have the NUM feature singular), whereas it is not violated in (24a), where the verb is in the plural form.

The COSUBJAGR constraint is the key for rejecting any possible instance of a transitive verb agreeing with its object in Catalan just because the object ranks higher than the subject in the Person-Number Hierarchy. As the object is not a cosubject, a verb agreeing with its object violates COSUBJAGR, which is the highest ranking constraint, thus it follows that the verb must agree with the subject (the only cosubject available in a transitive sentence) in order to comply with the COSUBJAGR constraint. An interesting situation is presented by transitive sentences with a reflexive object that is coreferential with the subject: here the reflexive object is a cosubject. Per definition, the reflexive shares its agreement features with its antecedent, in this case, the subject. Therefore, the reflexive will always have the same ranking as the subject in the Person-Number Hierarchy, so MARKEDAGR does not choose which candidate is grammatical as it is violated exactly the same amount of times in both cases. SUBJAGR is the constraint which rejects the candidate where  $GF_j$  is the object, thus licensing the candidate where the verb is stated to agree with the subject as the only grammatical form.

In a non-CI language, the optimization forces the verb to agree with the subject. The ranking of constraints shown below is licensed by (23), but it is just one of the two possible rankings due to the underspecification of the relative order between MARKEDAGR and COSUBJAGR in non-CI languages. Therefore, for a minimal pair in English like the one presented below, the OT tableau is the one shown in Table 2:

- (25) a. The problem is taxes

- b. \* The problem are taxes

		MARKEDAGR		
	SUBJAGR	AGRPERS	AGRNUM	COSUBJAGR
☞ (25a)		*	*	
(25b)	*!	*		

Table 2: Optimization for (25)

Let us now consider the case in which one cosubject is in the 1st person and the other one in the 2nd person. As the reader will recall, the Person-Number Hierarchy states that the 1st and 2nd person both rank the highest. Consider the following set of data, where the four mathematically possible combinations of 1st and 2nd person singular in the subject and complement function of the copula are shown:

- |      |    |                       |            |      |    |                         |          |
|------|----|-----------------------|------------|------|----|-------------------------|----------|
| (26) | a. | Jo sóc                | tu         | (27) | a. | Tu ets                  | jo       |
|      |    | <i>I be.pres.1.sg</i> | <i>you</i> |      |    | <i>you be.pres.2.sg</i> | <i>I</i> |
|      |    | ‘I am you.’           |            |      |    | ‘You are me.’           |          |
|      | b. | * Jo ets              | tu         |      | b. | * Tu sóc                | jo       |
|      |    | <i>I be.pres.2.sg</i> | <i>you</i> |      |    | <i>you be.pres.1.sg</i> | <i>I</i> |

The description of the data above can be captured by the simple statement that, if both cosubjects rank equally, then the verb must agree with its subject. This is predicted by our theory as is, as shown by the tableau below for (26):<sup>12</sup>

		MARKEDAGR		
	COSUBJAGR	AGRPERS	AGRNUM	SUBJAGR
☞ (26a)			*	
(26b)			*	*!

Table 3: Optimization for (26)

As shown above, in such cases, SUBJAGR is the deciding factor that rules out the ungrammatical structure (26b) in favor of (26a). This shows that, even in CI languages, where agreement with a cosubject is the strongest of the constraints being considered here, SUBJAGR becomes relevant when the higher ranking constraints do not provide a way to distinguish between the available candidates.

In summary, the claim we defend here is that all verbs are subject to cosubject-verb agreement in CI languages, but non-subject agreement only arises in copular

<sup>12</sup>We only provide the tableau for (26) as the optimization is identical to the one for (27).



sentences because the copula has the unique function of requiring its complement to be coreferential with its subject and therefore only copular sentences have more than one cosubject that the verb can agree with. When two cosubjects are available, we claim that the grammatical structure is the one in which the copula violates MARKEDAGR the least amount of times, i.e. by agreeing with the DP that is higher in the Person-Number Hierarchy.

## 6 The Norman Bates Problem: when there is no coreference

The analysis provided above presents an apparent problem that we have called “The Norman Bates Problem,” in which CI does not arise:<sup>13</sup>

- (28) a. Norman Bates és moltes persones a *Psicosi*  
*Norman Bates be.pres.3.sg many.pl person.pl in Psycho*  
 ‘Norman Bates is many people in *Psycho*.’  
 b. \* Norman Bates són moltes persones a *Psicosi*  
*Norman Bates be.pres.3.pl many.pl person.pl in Psycho*

If we blindly applied our theory, we would predict exactly the opposite of what the data show: we would predict that the copula should be in the 3rd person plural, as such a candidate violates just AGRPERS, whereas (28a) violates both the higher ranking AGRNUM constraint as well as AGRPERS.

This is not restricted to Catalan and closely related languages. German, which we already have claimed to be a CI-language, also shows the exact same issue as in the sentences above:

- (29) a. Norman Bates ist mehrere Personen in *Psycho*  
*Norman Bates be.pres.3.sg/\*pl many.pl person.pl in Psycho*  
 ‘Norman Bates is many people in *Psycho*.’  
 b. \* Norman Bates sind mehrere Personen in *Psycho*  
*Norman Bates be.pres.3.sg/\*pl many.pl person.pl in Psycho*

The key to this is the lack of coreference between the subject and the predicate in a case like (28a). The meaning of the sentence above is not that Norman Bates really is many people, but that he plays *the roles* of many people. This introduces a layer of fiction whereby the copula is linking an entity with the representation of another entity, i.e. the role of that entity, which is in no case that entity. This implies that the subject and the predicate are not coreferential and thus, that the predicate is not a cosubject. Therefore, the only cosubject available in (28a) is the subject, which explains why the singular is the only possible form in this case; this

<sup>13</sup>We thank a reviewer for bringing up this potential problem.

leaves a situation that is exactly the one in transitive sentences, where the object is not a cosubject. The optimization for (28) follows in Table 4:


	COSUBJAGR	MARKEDAGR		SUBJAGR
		AGRPERS	AGRNUM	
 (28a)		*	*	
(28b)	*!	*		*

Table 4: Optimization for (28)

In the case of (28) the CI version of the structure is impossible, as there cannot be coreference between the subject and the complement because it is contradictory to believe that one person is many people.

However, a distinction must be made between people acting as, or playing the role of, other people as in (28) and inanimate objects that represent people such as statues and pictures. In the former case, coreference between the two expressions denoting the people involved is disallowed, whereas in the latter case coreference between the image and the person represented by it is required. For example, in the context in which a person is pointing at a picture of herself, she can utter (30a) but not (30b):

- (30) a. *Aquesta noia sóc jo*  
*this girl be.pres.1.sg I*  
 ‘I am this girl.’  
 b. \* *Aquesta noia és jo*  
*this girl be.pres.3.sg I*

In some cases, a given object can be taken to be either the representation of a specific person or not. In such cases, the DP corresponding to the object, if used as the subject of the copula, may either be coreferential with the complement or not, so that the two agreement forms shown below are possible:

- (31) a. *Aquest és jo*  
*this be.pres.3.sg I*  
 ‘This is me.’  
 b. *Aquest sóc jo*  
*this be.pres.1.sg I*  
 ‘This is me.’

The case in (31) is data taken from a TV program in the Catalan public broadcasting service. At a certain point, the speaker is holding a puppet that looks like him and utters (31a) while holding, and pointing at, the puppet. The context makes it quite clear that the intended meaning is ‘This (puppet) represents me.’ In the same context, example (31b) can be used meaning ‘I am this (puppet).’ In our

interpretation, there is no coreference between subject and complement in example (31a), whereas there is in (31b). Without coreference the verb has to agree with its subject, as in (28a); with coreference, it agrees with the complement, since the complement is a cosubject that is higher than the subject in the Person-Number Hierarchy.

The conclusion we draw from these data is that they actually support our claim that it is coreference that drives the existence of CI in copular sentences. In some copular sentences, coreference between the subject and the complement is not possible for semantic reasons and, therefore, CI does not arise.

## 7 Conclusions and possible extensions

As has been shown in the previous pages, our proposal radically modifies our way of thinking about verb agreement. We have claimed that the best approach towards explaining CI is to abandon the hypothesis that verbs specify the person and number features of the subject; our claim implies that verbs and GFs have their features specified in a special f-structure named AGR so that the verb's AGR must unify with one of the GFs in the verb's clause, according to the GF-VERB AGREEMENT well-formedness constraint.

The GF which the verb agrees with is determined in a language-specific way. In the case of non-CI languages, it will always be SUBJ, due to SUBJAGR being the highest-ranking constraint. On the other hand, in CI languages the GF is a cosubject and, in case there is more than one, the set of constraints MARKEDAGR determines which one the verb agrees with, namely the one that ranks higher in the Person-Number Hierarchy.

There is, of course, room for extending this research further. There are non-subject agreement phenomena that we propose may be analyzed by means of the theory that we have defended in this paper.

One of the phenomena that may be of greatest interest to explore is the case of non-subject agreement in Dargwa, as described by Belyaev (2013). Belyaev (2013) claims that in certain Dargwa dialects transitive verbs agree with the GF that is more prominent in terms of person,<sup>14</sup> regardless of which GF this is. This situation is exactly the one that is predicted if we ranked our constraints like this:

(32) MARKEDAGR  $\gg$  SUBJAGR, COSUBJAGR

The only obstacle that must be faced if our analysis is to be extended to the data presented by Belyaev (2013) is the fact that in these dialects the 2nd person takes precedence over the 1st, which would require a more fine-grained AGRPERS constraint than the one that we have proposed so far.

Finally, there are phenomena of “oblique SUBJS” that also allow for an analysis like ours, namely Icelandic quirky case and English locative inversion. Both consist

---

<sup>14</sup>Number seems to be an underspecified feature in this language, so Belyaev (2013) does not take it into account.

in having a non-nominative element (an oblique case DP in Icelandic, a locative PP in English) taking the SUBJ function, such that the verb must agree with a non-SUBJ nominative DP. This situation can be perfectly analyzed within our framework by stating an AGRCASE principle that requires the case of the GF which the verb unifies its AGR with to be nominative. This constraint must be the highest-ranking constraint in these languages in order to predict the desired results:

(33) AGRCASE:  $\left[ \text{AGR} \left[ \text{CASE NOM} \right] \right]$

(34) AGRCASE  $\gg$  SUBJAGR  $\gg$  MARKEDAGR, COSUBJAGR

As has been shown in the previous pages, even though the theory presented implies a significant departure from common assumptions on agreement in LFG, it allows for an explanation of CI that integrates seamlessly with cases of subject-verb agreement. The only difference with respect to standard treatments of subject-verb agreement is that our theory treats this phenomenon as a particular instance of a more general verb agreement mechanism available in all languages with agreement, the differences between CI and non-CI languages being explained as a result of different rankings of the constraints involved in agreement.

## References

- Alsina, A. (1995). The fall of function-argument biuniqueness. In G. V. Morill and R. T. Oehrle (Eds.), *Formal grammar: Proceedings of the Conference of the European Summer School in Logic, Language and information*, pp. 1–16. Barcelona: Universitat Politècnica de Catalunya.
- Alsina, A. (2007). Apparent copular inversion and the theory of null subjects. In A. Zaenen, J. Simpson, T. Holloway King, J. Grimshaw, J. Maling, and C. Manning (Eds.), *Architectures, rules, and preferences*, pp. 21–47. Chicago: The University of Chicago.
- Anttila, A. and Y.-m. Y. Cho (1998). Variation and change in Optimality Theory. *Lingua 104*, 31–56.
- Belyaev, O. (2013). Optimal agreement at m-structure: person in Dargwa. In M. Butt and T. H. King (Eds.), *Proceedings of the LFG13 Conference*, pp. 90–110. Stanford: CSLI Publications.
- Berman, J. (2003). *Clausal syntax of German*. Stanford: CSLI Publications.
- Bonet, E. (1990). Subjects in Catalan. In L. L. S. Cheng and H. Demirdash (Eds.), *Papers on wh-movement*, pp. 1–26. Cambridge, Mass.: Massachusetts Institute for Technology. Department of Linguistics and Philosophy.

- Bresnan, J. (2000). Optimal syntax. In J. Dekkers, F. van der Leeuw, and J. van der Weijer (Eds.), *Optimality theory: phonology, syntax and acquisition*, pp. 334–385. Oxford: Oxford University Press.
- Bresnan, J. (2001). *Lexical-functional syntax*. Malden, Mass.: Blackwell Publishing.
- Dalrymple, M. (2001). *Lexical functional grammar*. San Diego: Academic Press.
- Fernández Leborans, M. J. (1999). La predicación: las oraciones copulativas. In I. Bosque and V. Demonte (Eds.), *Gramática descriptiva de la lengua española*, Volume 2, pp. 2357–2460. Madrid: Espasa Calpe.
- Forcadell, M. (2013). Subject informational status and word order: Catalan as an SVO language. *Journal of Pragmatics* 53, 39–63.
- Hernanz, M. L. and J. M. Brucart (1987). *La sintaxis*. Barcelona: Crítica.
- Kuhn, J. (2003). *Optimality-Theoretic syntax: a declarative approach*. Stanford: CSLI Publications.
- Moro, A. (1997). *The raising of predicates*. Cambridge: Cambridge University Press.
- Perlmutter, D. M. (1983). Personal vs. impersonal constructions. *Natural language and linguistic theory* 1(1), 141–200.
- Prince, A. and P. Smolensky (2004). *Optimality Theory: constraint interaction in generative grammar*. Malden, Mass.: Blackwell Publishing.
- Ramos, J.-R. (2002). El SV, II: la predicació verbal no obligatòria. In J. Solà, M. R. Lloret, J. Mascaró, and M. Pérez Saldanya (Eds.), *Gramàtica del català contemporani*, Volume 2, pp. 1951–2044. Barcelona: Editorial Empúries.
- Solà, J. (1992). *Agreement and subjects*. Ph. D. thesis, Universitat Autònoma de Barcelona.
- Vallduví, E. (1992). A preverbal landing site for quantificational operators. *Catalan working papers in Linguistics* 2, 319–343.