Catalan Intransitive Verbs and Argument Realization

Alex Alsina
Universitat Pompeu Fabra

Fengrong Yang
Universitat Pompeu Fabra

Proceedings of the LFG'18 Conference

University of Vienna

Miriam Butt, Tracy Holloway King (Editors)

2018

CSLI Publications

pages 46–66

http://csli-publications.stanford.edu/LFG/2018

Keywords: intransitive verbs, argument realization, mapping theory, verb agreement

Alsina, Alex, & Yang, Fengrong. 2018. Catalan Intransitive Verbs and Argument Realization. In Butt, Miriam, & King, Tracy Holloway (Eds.), *Proceedings of the LFG'18 Conference, University of Vienna*, 46–66. Stanford, CA: CSLI Publications.

Abstract

The goal of this paper is to analyze the behavior of the single direct argument of intransitive verbs in Catalan, including its encoding as a grammatical function, verbal agreement, case assignment, and expression by means of clitics. Our main claim is that the single direct argument of a clause can be a nominative object. We show that the direct argument of intransitive verbs (whether unaccusative or unergative) alternates between subject and object. The proposed analysis diverges from standard versions of LFG, as it allows an external argument to map onto an object and allows a clause to lack a subject, in violation of the Subject Condition. We propose a new mapping theory in which case assignment plays a major role and account for the agreement facts by assuming a set of agreement features of the clause (AGR) that are identified with a grammatical function (GF), not necessarily the subject, by general constraints.

1 Introduction

The topic of this paper is the behavior of the single direct argument[1] of intransitive verbs (the *intransitive argument*, for short). The relevant facts are presented in section 2, showing that the behavior of that argument is split between subject and object. The argument realization theory needed to account for these facts is proposed in section 3, where case assignment plays a crucial role in constraining the mapping of arguments to grammatical functions. The agreement facts are discussed and explained in section 4, adopting the theory of agreement proposed by Alsina and Vigo (2014, 2017). The main conclusions are summarized in section 5.

2 Properties of the sole argument of intransitive verbs

The intransitive argument behaves in some ways like a subject and in some ways like an object. We start by showing its object properties, in 2.1; then, turn to its subject properties, in 2.2, focusing on the agreement facts in 2.3.

2.1 Object properties

En cliticization provides evidence that the intransitive argument can be an object in Catalan. (Other Romance languages, such as Italian and French, show a similar behavior of the cognate clitic *en* or *ne*.) The internal argument of Catalan transitive verbs can be partially or totally expressed by means of the clitic *en*:[2] *en* in (1a) and (1b) replaces *carpetes* 'folders' and *carpetes de plàstic*

† We deeply acknowledge the comments and observations made by the anonymous reviewers and the audience of the 23th LFG conference. Any remaining errors are our own.

[1] By *direct argument* we refer to an argument whose default expression is as a direct grammatical function with an unmarked case feature (nominative or accusative).

[2] Here we are only concerned with one of the two functions of the clitic *en*, which we may call 'partitive' *en*, as it replaces the head noun of an indefinite object and cannot cooccur with it. In the other function—'genitive' *en*—the clitic corresponds to a *de*-complement of the verb or of the verb's object, as shown in (i), where *en* corresponds to the verb's *de*-complement:

noves 'new plastic folders', respectively:

(1) a. Si vols carpetes, en tinc tres de noves.
if want.2p.sg *folder*.f.pl *en*.cl *have*.1p.sg *three of new*.f.pl
'If you need folders, I have three new ones.'

b. Si vols carpetes de plàstic noves, compra'n.
if want.2p.sg *folder*.f.pl *of plastic new*.f.pl *buy*.imp.2p.sg-*en*.cl
'If you need new plastic folders, buy some.' (Alsina 1986:97-98)

The internal argument of Catalan unaccusative verbs patterns with the internal argument of transitive verbs in terms of the *en* cliticization:

(2) Cada dia surten molts trens,
every day leave.pl *many*.m.pl *train*.m.pl
però avui només n'ha sortit un.
but today only en.cl-*have*.sg *leave*.pp *one*
'Everyday many trains leave, but today only one has left.'

Surprisingly, although Catalan transitive verbs do not allow their external arguments to be cliticized by *en*, as in (3), the external argument of unergative verbs nevertheless can be replaced by the *en* clitic, as in (4):[3]

(3) a. * N'aprovaran tres els exàmens.
en.cl-*pass*.fut.pl *three the*.m.pl *exam*.m.pl

b. * N'aprovaran els exàmens tres.
en.cl-*pass*.fut.pl *the*.m.pl *exam*.m.pl *three*
'Three of them will pass the exams.'

(4) a. En ploraran sis quan sàpiguen la veritat.
en.cl *cry*.fut.pl *six when know*.sbjv.pl *the*.f.sg *truth*.f.sg
'Six of them will cry when they find out the truth.'
(Cortés and Gavarró 1997:41)

b. - Com repartirem les conferències?'
how distribute.fut.1p.pl *the*.f.pl *conference*.f.pl
'How should we arrange the conferences?
- Avui en poden parlar dos i demà tres més.
today en.cl *can*.3p.pl *talk*.inf *two and tomorrow three more*
'Today two of them can give a talk and tomorrow three.'
(Gràcia 1989:82)

The possibility of *en* cliticization with unergative verbs in Catalan shows that

---

(i) Podria parlar avui d'aquest problema, però en parlarà demà.
could.3p.sg *speak today of-this problem but en*.cl *speak tomorrow*
'He could speak about this problem today, but he will speak about it tomorrow.'

[3] It has sometimes been claimed that only unaccusative verbs allow *en* cliticization. Here we are describing the facts of speakers who accept *en* cliticization with unergatives as well as with unaccusatives, like Cortés and Gavarró (1997) for Catalan, or Saccon (1995) for Italian. Independent evidence for the claim that *plorar* 'cry' in (4a) and *parlar* 'talk' in (4b) are unergatives comes from tests such as the participial adjunct test in Cortés and Gavarró (1997). Note that *poden* 'can' in (4b) is a restructuring verb, which inherits the argument structure of the dependent verb.

it is not the 'deep object', i.e., the internal argument, that triggers *en* cliticization. Instead, the fact that both unaccusative and unergative verbs allow their single direct argument to be expressed by means of *en* requires assuming that the argument in question is an object (or the 'surface object', in theories like Burzio 1986, or Cortés and Gavarró 1997, among others).[4]

The second argument for the object status of the intransitive argument is past participle agreement. In Catalan, the past participle optionally agrees in gender and number with a third person object clitic, when cooccurring with the perfective auxiliary *haver* 'have'. But this agreement does not happen with a full NP object:

(5) a. La directora ha defensat/*defensada la proposta.
the director have.sg *defend*.pp.m.sg/*f.sg *the*.f.sg *proposal*.f.sg
'The director has defended the proposal.'

b. La directora **l**'ha defensa**da**.
the.f.sg *director*.f.sg *la*.cl.f.sg-*have.sg defend*.pp.f.sg
'The director has defended it.' (Alsina 1996:95)

Past participle agreement is not only possible with objects of transitive verbs, like the one in (5b), but also with the direct argument of intransitive verbs:[5]

(6) a. Perquè aleshores hi haurà una gran tribulació,
because then hi.cl *have*.fut.sg *one*.f.sg *great distress*.f.sg
com no **n**'hi ha hagu**da** cap des de
like not en.cl-*hi*.cl *have*.3p.sg *have*.pp.f.sg *never from*
la creació del món…
the creation of-the world
'For then there will be great distress, as there has not been one since the creation of the world...' (Bible [Mt 24:21][6])

b. **N**'han arribat**s** molt**s**.
en.cl-*have*.pl *arrive*.pp.m.pl *many*.m.pl
'Many have arrived.' (Fabra 1912:160)

The fact that an intransitive argument expressed as the clitic *en* can trigger past participle agreement further confirms that the argument is an object.

The possibility of expressing the intransitive argument as a bare indefinite NP gives additional evidence for the objecthood of this argument. Bare indefinite NPs, which have a non-specific interpretation, can encode the object of a transitive verb, as shown in (1a). However, they cannot be the subject of the verb, as illustrated in (7) with a transitive verb:

---

[4] Notice that, with respect to the phenomena examined in section 2, there is no difference in behavior among one-argument verbs between unaccusatives and unergatives in Catalan, and we refer to this distinction precisely to make this point.

[5] We have not documented past participle agreement with *en* with unergatives, although it is expected to be possible. This may be due to the fact that this construction is infrequent and formal, and not used by many speakers. We leave it to further research to decide whether unergatives are excluded from this construction.

[6] http://www.biblija.net/biblija.cgi?m=Mt+24%2C1-31&l=ca, visiting time: 18:19, 08/07/2018

(7) a. * Arreglen mecànics el teu cotxe.
fix.pl *mecanics*.m.pl *the*.m.sg *your*.m.sg *car*.m.sg
b. * Arreglen el teu cotxe mecànics.
fix.pl *the*.m.sg *your*.m.sg *car*.m.sg *mecanics*.m.pl
'Mechanics fix your car.' (Alsina 1996:104)

By contrast, the intransitive argument can freely be expressed as a bare NP:

(8) a. Cau aigua de la teulada.
fall.sg *water*.f.sg *from the*.f.sg *roof*.f.sg
'Water is falling from the roof.' (Alsina 1995:13)
b. Treballen nens en aquesta fàbrica.
work.pl *child*.m.pl *in this*.f.sg *factory*.f.sg
'Children work in that factory.' (Cortés 1995:64)

The contrast between examples (7) and (8) indicates that both *aigua* 'water' in (8a) and *nens* 'children' in (8b) are objects and not subjects. The evidence from bare NPs, together with *en* cliticization and optional past participle agreement, indicates that the intransitive argument is an object.

2.2 Subject properties

Catalan is known to be a subject pro-drop language: in Catalan, a subject can be null and be interpreted as having a definite referent, whereas an object cannot be null with a definite reading:

(9) a. Els estudiants solen sortir puntualment,
the.m.pl *student*.m.pl *be-used-to*.pl *leave*.inf *punctually*
però avui Ø surten tard.
but today leave.pl *late*
'Students usually leave on time, but today they are leaving late.'
b. Els estudiants no volen estudiar habitualment,
the.m.pl *student*.m.pl not *want*.pl *study*.inf *usually*
però avui Ø estudien molt.
but today study.pl *a-lot*
'Students usually do not want to study, but today they are studying a lot.'

(10) Joan ha llegit el diari avui,
John have.3p.sg *read*.pp.m.sg *the*.m.sg *newspaper*.m.sg *today*
però no llegirà demà.
but not read.fut.3p.sg *tomorrow*
'John has read the newspaper today, but will not read (*it) tomorrow.'

The contrast between (9) and (10) shows that grammatical functions other than the subject in Catalan cannot be null with a definite reading. Therefore, the fact that the intransitive argument in Catalan can be omitted and have a definite referent, as in (9), requires analyzing it as the subject of the clause.

Another subject property is the possibility of being the controlee in a control

construction, because only the subject of the embedded clause can be controlled by the subject or object of the matrix clause, as shown in (11) for a transitive verb in an embedded clause:

(11) N'he obligat molts a examinar el metge.
en.cl-*have*.1p.sg *obligate*.pp *many*.pl *to examine*.inf *the doctor*
'I have obligated many to examine the doctor.'
* 'I have obligated many to be examined by the doctor.'

In contrast, as the object of the embedded clause, the intransitive argument cannot be controlled by an argument of the embedding clause. Examples (12) and (13) illustrate this contrast.

(12) a. * N'he obligat molts a quedar-se'n.
en.cl-*have*.1p.sg *obligate*.pp *many*.m.pl *to stay*.inf-*se*.cl-*en*.cl
'I have obligated many to stay.'

b. N'he obligat molts a quedar-se.
en.cl-*have*.1p.sg *obligate*.pp *many*.m.pl *to stay*.inf-*se*.cl
'I have obligated many to stay.'

(13) a. * N'he obligat molts a estudiar-ne.
en.cl-*have*.1p.sg *obligate*.pp *many*.m.pl *to study*.inf-*en*.cl
'I have obligated many to study'

b. N'he obligat molts a estudiar.
en.cl-*have*.1p.sg *obligate*.pp *many*.m.pl to *study*.inf
'I have obligated many to study.'

The fact that the control relation in (12b) and (13b) is grammatical indicates that the intransitive argument of the embedded clause is the subject. This is further confirmed by the ungrammaticality of (12a) and (13a), in which the clitic *en* appears in the embedded clause. If we assume that *en* cliticization is an object property, the ungrammaticality of (12a) and (13a) follows naturally: as an object, the argument of the embedded clause cannot be controlled.

2.3 Verbal agreement

It is commonly assumed that the agreement trigger of the verb is the subject (Chomsky 1981, 1995, among others). In a simple example with a transitive verb like (14), the auxiliary *haver* is in the third person plural form, agreeing with the subject *els estudiants* 'the students':

(14) Els estudiants han/*ha llegit aquest llibre.
the.pl *student*.pl *have*.3p.pl/*sg *read*.pp.m.sg *this*.m.sg *book*.m.sg
'The students have read this book.'

Intransitive verbs regularly agree with their single direct argument. But we would have a problem if we should assume that the agreement trigger is necessarily the subject: *molts* in (15) would have to be both a subject (as the agreement trigger) and an object (as it is expressed by means of the *en* clitic):

(15) Avui en surten/*surt molts.
today en.cl *leave*.pl/*sg *many*.pl
'Today many are leaving.'

The verbal agreement facts of languages like Icelandic or Hindi indicate that, in such languages, the verb can agree with a grammatical function other than the subject, provided that it is in *nominative* case:

(16) a. Henni líkuðu hestarnir.
she.dat.3p.sg *like*.past.3p.pl *horse*.nom.3p.pl
'She liked the horses' (Icelandic, Sigurðsson 2004:139)

b. Ravii-ne / niinaa-ne kelaa k^{h}aayaa
Ravi-erg.m.sg/ *Nina*-erg.f.sg *banana*-nom.m.sg *eat*.perf.m.sg
'Ravi/Nina ate a banana.' (Hindi, Mohanan 1994:104)

The same assumption will allow us to solve the paradox of (15): the verb agrees with a nominative argument, whether it is a subject or an object, and in (15) the verb in fact agrees with the object, which is nominative.

Independent evidence for the claim that the argument with which the verb agrees is nominative comes from the contrast between nominative and accusative with respect to the use of the preposition *a* 'to'. An indefinite pronoun allows *a*-marking optionally only if it is animate and accusative:

(17) a. (*A) molts llegeixen el llibre.
to *many*.m.pl *read*.3p.pl *the*.m.sg *book*.m.sg
'Many read the book.'

b. En veiem (a) molts.
en.cl *see*.1p.pl *to many*.m.pl
'We see many.'

Nominatives never allow *a*-marking, whether SUBJ (as in (17a)) or OBJ (as in (18)):

(18) En surten (*a) molts.
en.cl *leave*.pl *to many*.m.pl
'Many are leaving.'

From the facts listed above, we conclude that the intransitive argument in Catalan alternates between subject and object, and is always nominative. The intransitive verb agrees with this argument, regardless of the function it takes.[7]

3 Argument realization

In this section we propose the theory of argument realization needed to account for the facts reported in the previous section concerning the expression of the intransitive argument in Catalan. In 3.1, we briefly point out the difficulties that existing theories of argument realization within LFG would face in

[7] The properties discussed in subsection 2.2 cannot be attributed to nominative case, rather than to subjecthood. For example, the controlee has to be the subject and not just a nominative argument, as shown in (12) and (13).

accounting for these facts. In 3.2, an alternative argument realization theory is proposed, in which case assignment is a central element. In 3.3, we show how some of the main facts are derived from this theory, and, in 3.4, we show some constraints on the subject-object alternation.

3.1 Current LFG mapping theories

Current LFG theories of argument realization face two problems with respect to the facts considered in this paper: the treatment of multiple objects and the difficulty in accounting for the subject-object alternation of external arguments, which we will address in turn.

Since its inception, LFG has assumed as a general property of all languages that clauses have at most one unrestricted object and possibly one or more restricted objects. These two kinds of GFs have been designated by different names, including OBJ and OBJθ, to refer to unrestricted and restricted object, respectively, which we shall use for brevity. While the distinction between these two types of object finds strong motivation in asymmetrical languages such as Chicheŵa (see Alsina and Mchombo 1990, 1993, and Bresnan and Moshi 1990, among others), it is unmotivated in many other languages, particularly in languages that make use of grammatical case such as Catalan and the other Romance languages.[8] Therefore, assuming the OBJ/OBJθ distinction for all languages constitutes an unnecessary complication of the analysis of multiple objects in the latter type of language.

As noted already in Alsina (1996), the relevant distinction among objects in Catalan (as well as other Romance languages) is in terms of grammatical case: dative vs. non-dative objects. Stipulating that one of the two objects is an OBJ and the other one an OBJθ plays no role in accounting for the facts in this language and does not allow us to maintain that this distinction has a cross-linguistically valid empirical reflex. The behavior of objects in Catalan is entirely predictable from the presence or absence of dative case. Stipulating that the dative object is the OBJθ is redundant, as it would be to stipulate that the dative object is the OBJ and the non-dative object is the OBJθ. Both dative and non-dative objects can be expressed by means of pronominal clitics (and in some cases dative objects are preferentially expressed in this way), which can be taken to be the equivalent of object marking in the Bantu languages, a property not available to OBJθ. Both dative and non-dative objects can be reflexivized (and reciprocalized), which is the equivalent of reciprocalization in Bantu, another property in which OBJθ does not take part. The failure of dative objects (in contrast with non-dative objects) to alternate with the SUBJ function (i.e., to passivize) is best analyzed by means of a language-particular constraint disallowing dative subjects (see Nominative Subject Constraint (23) below). As is well known, other case-marking languages lack this constraint

[8] Bresnan and Moshi (1990:167) already note that "many languages (including Romance) lack restricted objects altogether". We thank an anonymous reviewer for bringing up this point.

and allow dative subjects, or other subjects with other marked cases (e.g., Icelandic, Hindi-Urdu, etc.).

In addition, importing the OBJ/OBJθ distinction into Catalan would render this distinction devoid of any cross-linguistically valid empirical effect. On the basis of asymmetrical languages such as Chicheŵa, in which the OBJ/OBJθ distinction does play an important role, we can observe that certain properties are only available to OBJ, such as expression by means of an object marker, possibility of passivization, or accessibility to reciprocalization. In Catalan, the two types of objects are available for expression by means of a verbal clitic and for reflexivization. If dative objects were assumed to be OBJθ and non-dative objects were assumed to be OBJ, it would no longer be possible to maintain that certain properties (such as expression by means of object markers or clitics and accessibility to reflexization or reciprocalization) are cross-linguistically properties of OBJ (that is, unavailable to OBJθ).

Therefore, we do not assume that objects in Catalan are represented as either OBJ or OBJθ. Instead, we assume that, cross-linguistically, there can be multiple instances of the GF OBJ and that, in some languages, objects are distinguished by means of grammatical case. Catalan is one of these languages, in which objects can be either dative or non-dative. In languages such as Chicheŵa, where there are no grammatical case distinctions, objects are distinguished between restricted and unrestricted at the level of argument structure. As proposed in Alsina (2001), internal arguments may be marked as R at the level of argument structure, so that there may be at most one internal argument not marked with this feature. This feature makes the argument so marked unavailable to the morphosyntactic properties noted above (object marking, reciprocalization, possibility of passivization).

The proposal that objects are not distinguished in terms of grammatical function, since they all bear the GF OBJ, but may be distinguished either in terms of grammatical case (as in Catalan) or in terms of the presence or absence of the feature R at the level of argument structure (as in Chicheŵa) entails rejecting the four-way classification of grammatical functions found in current versions of the Lexical Mapping Theory (LMT), as in, for example, Levin (1986), Bresnan and Kanerva (1989), Bresnan and Moshi (1990), Kibort (2001, 2009, among others), and Findlay (2016). These versions of LMT assume that there are four basic GFs: SUBJ, OBJ, OBJθ, and OBLθ. These theories also assume a decomposition of these GFs by means of the features [±r] and [±o] and that arguments are classified by means of these features. Since these features combine to yield the four GFs just mentioned, they also need to be discarded in the theory to be advanced in subsection 3.2.

The second problem with current LFG mapping theories can be seen as a consequence of the featural decomposition of GFs just discussed. The classification of an argument by means of one of these features implies the possibility of an alternation between two GFs. If an argument is classified at a-structure as [–r], as is assumed for internal arguments, it can map onto either

SUBJ or OBJ; if it is classified as [+o], it can map onto either OBJ or OBJθ, and so on. This restricts the possible GF alternations. External arguments, such as agents, are assumed to have the [–o] classification, which limits the possible realizations to SUBJ and OBLθ. What is not assumed in current versions of LMT is for external arguments to show a SUBJ/OBJ alternation, but what we find in Catalan is that the intransitive argument, whether internal or external, shows the SUBJ/OBJ alternation.[9] In contrast, the external argument of transitive verbs is constrained to map onto the SUBJ function. This shows that a [–o] argument classification is inadequate for external arguments and that the mapping of external arguments depends in part on the other arguments in the argument structure.

3.2 Argument-to-function mapping theory

The present mapping theory assumes a level of argument structure, or a-structure, and three sets of principles of argument realization, which relate a-structure to f-structure: case assignment principles, argument-to-GF linking rules, and constraints on case features.

A-structure

A-structure consists of the list of arguments of a predicate, without any thematic information, ordered according to the thematic hierarchy, such as the commonly assumed hierarchy based on Givón (1984), Kiparsky (1987), and Bresnan and Kanerva (1989), among others:

(19) Thematic Hierarchy:
ag > ben > recip/exp > inst > th/pt > loc

Arguments are classified into *core* arguments (C) and *non-core* arguments (NC). As we shall see, core arguments are the ones that map onto direct grammatical functions (i.e., SUBJ and OBJ). Core arguments are further divided into external argument (E) and internal argument (I) and represented as such in the a-structure. The external argument E, if there is one, is the most prominent argument in the argument structure. Non-core arguments are those that map onto the indirect function OBL.

Case assignment principles

In this theory, case assignment is crucial for argument realization. For Catalan, we assume that there are three case values—dative, accusative, and nominative—for the core arguments, and that all core arguments must be assigned a case value, according to the following case assignment principles, ordered by priority:

[9] The subject-object alternation of the intransitive argument (both internal and external argument) is also found in Norwegian, Swedish, and Danish, according to Lødrup (1999).

(20) Case Assignment Principles:

i. Assign *dative* case to the more prominent of two internal arguments, or to a goal;[10]

ii. Assign *accusative* case to the less prominent of two *core* arguments that lack case;

iii. Elsewhere, assign *nominative* case to a *core* argument.

Argument-to-GF linking rules

We propose two rules to license the correspondence between arguments and GFs—the Core Argument Rule and the Elsewhere Mapping Rule—and Passivization, as an instance of a morphosyntactic operation that affects the argument-to-GF linking.

The Core Argument Rule requires a core argument (C) to map onto a direct grammatical function (DGF), the class of GFs that consists of SUBJ and OBJ:

(21) Core Argument Rule: C
|
DGF

This rule allows the external argument, as well as an internal argument, to be either SUBJ or OBJ, which is not possible in previous mapping theories like Bresnan and Kanerva (1989), Kibort (2001), or Findlay (2016), for, as noted earlier, the proposal that the external argument is associated with [–o] prevents linking this argument to an OBJ.

The operation of passivization blocks the linkage of the highest argument to a DGF:[11]

(22) Passivization: $\hat{\theta}$
|
DGF

Finally, the Elsewhere Mapping Rule optionally links an argument to OBL:

(23) Elsewhere Mapping Rule: A
|
(OBL)

This rule is ordered after the other linking rules and therefore it applies to arguments to which the Core Argument Rule (21) cannot apply: non-core arguments as well as arguments that have their linkage to DGF cut off by morphosyntactic operations like passive or antipassive. The optionality of this rule captures the idea that in general OBLs are not obligatory. Moreover, this optionality may be overridden by having a lexical entry specifying that an

[10] See Alsina (1996:175) for a detailed discussion.

[11] This allows for cross-linguistic variation. In Spanish and Catalan, passivization prevents the linkage of the highest argument to a direct grammatical function, thus accounting for the *se* passivization/impersonalization with both unergative and unaccusative verbs, but in languages like German or Dutch, we need to rewrite the operation of passivization as 'blocking the linkage of the external argument to a direct grammatical function', since there is no passivization/impersonalization with unaccusative verbs in these languages.

argument is obligatorily mapped onto an oblique.

Constraints on case features

There are some constraints on the association of particular case features with particular GFs. Catalan, along with other Romance languages, but unlike languages such as Icelandic and Hindi-Urdu, requires subjects to be in the nominative case (or, conversely, rules out subjects in a case other than nominative). For example, in Catalan there are no dative subjects (see Alsina 1996) or accusative subjects. To account for this fact, we posit the Nominative Subject Constraint:

(24) Nominative Subject Constraint (specific to Catalan):
*SUBJ [CASE ¬NOM]

The effect of this constraint is to rule out structures with a non-nominative subject. Notice that the implication is unidirectional: subjects must be nominative, but it is not required for a nominative expression to be a subject.

A second case constraint that we need to consider is what we may call the 1 Non-Dative Object Constraint (or 1NDO): a structure allows at most one object that is not dative:

(25) 1 Non-Dative Object Constraint (1NDO):
*[CASE ¬DAT] [CASE ¬DAT]
| |
OBJ OBJ

This constraint rules out a structure with two accusative objects, or with two nominative objects, or with a nominative object and an accusative object. Together with constraint (24), it has the effect of requiring a nominative argument to be the subject if it co-occurs with an accusative object. Notice that the principles and constraints stated so far do not require the presence of a subject in the clause and so it is the 1NDO constraint that forces a nominative to be the subject if there is an accusative in the structure.

3.3 Illustration of the theory

We now provide some examples of how the proposed argument realization theory works in Catalan.

A ditransitive verb like *donar* 'give' is lexically specified with one external and two internal arguments, as represented in (26). The goal argument is the more prominent internal argument, thus, by case assignment principle (20i), it will get dative case. The theme argument, as the less prominent of the two arguments—agent and theme—lacking case, is assigned accusative case, according to principle (20ii). Finally, the external argument receives nominative case by principle (20iii). As for the argument-to-GF mappings, the three arguments, being core arguments, are required to map onto a direct GF by the Core Argument Rule (21). However, the goal and theme arguments can only be realized as OBJ according to the Nominative Subject Constraint (24) and the nominative agent argument must be realized as SUBJ in order to avoid

violating the 1NDO constraint (25). The representation in (26) and subsequent ones show the thematic roles of the arguments involved merely for convenience, as they are not part of the a-structure or of the f-structure; the a-structure is shown in angled brackets; the case features assigned to each argument are shown on the line below it and, on the next line, are the corresponding GFs; the relevant principles are given in parentheses.

(26) *Donar* 'give'

(ag)	(go)	(th)	
< E	I	I	>
nom	dat	acc	(Case assignment principles (20))
SUBJ	OBJ	OBJ	(Rules (21), (24), (25))

'Like' type verbs in Catalan (as well as other Romance languages like Spanish or Italian) have two internal arguments and no external argument. The experiencer argument gets dative case by principle (20i) and maps onto OBJ because of the Nominative Subject Constraint (24). The theme argument is assigned nominative case by principle (20iii), thus being compatible with both SUBJ and OBJ:[12]

(27) *Agradar* 'like'

(exp)	(th)	
< I	I	>
dat	nom	(Case assignment principles (20i, iii))
OBJ	SUBJ/OBJ	(Rules (21), (24))

Intransitive verbs, whether unergative or unaccusative, only have one *core* argument (an external and an internal argument, respectively), as exemplified in (28) for the unergative *treballar* 'work'. Case assignment principle (20iii) applies assigning nominative case. This core argument, as we have seen in section 2, alternates between SUBJ and OBJ.

(28) *Treballar* 'work'

(ag)		
< E	>	
nom		(Case assignment principle (20iii))
SUBJ/OBJ		(Rule (21))

Transitive verbs like *llegir* 'read' have an external and an internal argument. Since there is only one internal argument, dative case is not assigned;

[12] Although space does not permit us to give detailed examples of this class of verbs, only the theme argument can be expressed by the clitic *en*; the experiencer cannot, due to the case restriction in constraint (32).

accusative case is assigned to the less prominent argument (i.e., the internal argument); by principle (20iii), nominative case is assigned to the external argument. In accordance to the Nominative Subject Constraint and the 1NDO constraint, the external argument maps to the SUBJ and the internal argument to the OBJ:

(29) *Llegir* 'read'

(ag)	(th)	
< E	I >	
nom	acc	(Case assignment principles (20ii, iii))
SUBJ	OBJ	(Rule (21), (24), (25))

When the transitive verb is passivized, the linkage of the external argument to a direct grammatical function is blocked. Since there is only one internal argument, case assignment principles (20i, ii) will not be used. Then, by principle (20iii), the internal argument gets nominative case. This internal argument can map onto either SUBJ or OBJ:[13]

(30) *Llegit* 'read-PASS'

(ag)	(th)	
< E	I >	
	nom	(Rule (22), case assignment principle (20iii))
(OBL) DGF	SUBJ/OBJ	(Rules (21), (23))

From the representations in (26)-(30), which illustrate different patterns of argument realization, we can see that a clause in Catalan: **i)** *may* contain at most one SUBJ; **ii)** *need not* contain a SUBJ, and **iii)** may contain more than one OBJ. The uniqueness of the subject and the multiplicity of objects can be handled in a variety of ways (see e.g. Alsina 1996 and Patejuk and Przepiórkowski 2016). This proposal can be implemented within the standard LFG formalism by assuming that the SUBJ is single-valued and OBJ is set-valued. But we will not go into further details of this topic in this paper.

3.4 Constraints on the subject-object alternation

In section 2 we saw that the intransitive argument can alternate between SUBJ and OBJ. However, if this SUBJ/OBJ alternation were completely free, nothing would require the presence of the clitic *en* in (31), as shown by the contrast between the grammatical (15), repeated as (31a), with the clitic, and the absence of the clitic in the ungrammatical (31b):

[13] The nominative argument in a passivized clause can be either SUBJ or OBJ, showing the same behavior as the direct argument of intransitive verbs, described in section 2. We will discuss the conditions under which the subject-object alternation happens in subsection 3.4. We leave detailed issues about passivization and impersonalization for further study.

(31) a. Avui en surten molts.
today en.cl *leave*.pl *many*.pl
'Today many are leaving.'

b. * Avui surten molts.
today leave.pl *many*.pl
'Today many are leaving.'

Consider the information the *en* clitic provides: the *en* clitic corresponds to an OBJ that is pronominal and indefinite, which can either be nominative or accusative, but not dative, as illustrated in the f-structure in (32):

(32) *En*: $\begin{bmatrix} \text{OBJ} \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{DEF} & - \\ \text{CASE} & \neg\text{DAT} \end{bmatrix} \end{bmatrix}$

The presence of this clitic indicates that it corresponds to an object, which may be expressed by an NP lacking a head N, as is the case of *molts* 'many' in (31a). However, if the core argument of a verb like *sortir* 'leave' were free to also be expressed as a subject, we would expect (31b), without the clitic *en*, to be grammatical, as this clitic cannot correspond to a subject. In order to explain the ungrammaticality of (31b), we assume that the subject-object alternation of the intransitive argument is constrained by definiteness and posit a constraint that penalizes an indefinite subject:[14]

(33) Indefinite Subject Ban:
*SUBJ [DEF -]

For an intransitive verb whose single direct argument is indefinite, constraint (33) penalizes the subject realization and favors the object realization. This explains the obligatoriness of *en* in (31). But notice that this constraint has no effect on transitive verbs, within an Optimality Theory (OT) conception (see Kuhn 2003), provided 1NDO (25) ranks higher than (33): the subject realization of the external argument of a transitive verb is the optimal candidate, even if it is indefinite and violates (33).

By contrast, when the sole argument of the intransitive verb is definite, it is the subject of the clause, like the NP *els estudiants* in (34):

(34) Avui surten els estudiants tard.
today leave.pl *the*.m.pl *student*.m.pl *late*
'Today the students are leaving late.'

The reasoning is that we also assume the Subject Condition (SC) (see Bresnan and Moshi 1990, among others), which requires every clause to have a subject, as an OT constraint: SC is a low-ranking constraint and, in particular, lower than the Indefinite Subject Ban (33) in Catalan.[15] When the intransitive

[14] According to Bartra (2009:3), Spanish and Catalan allow plural indefinites as objects of the verb but not as external subjects. The claim refers to bare NPs, a subset of indefinites, and it is also made by Espinal (2010) and Espinal and McNally (2010).

[15] Notice that, because of this ranking of constraints and because, in languages like Catalan, SC ranks below the faithfulness constraint requiring every GF to correspond to an argument, there

argument is definite and is not constrained by (33), the SC will penalize the candidate that lacks a subject and select the one in which the argument maps onto the subject.[16]

An additional fact that needs to be considered is that the *en* clitic cannot be licensed by a preverbal NP, even if this NP is indefinite:

(35) a. Ja n'han sortit quatre de l'ou.
already en.cl-*have*.pl *leave*.pp *four from the-egg*

b. Quatre ja (*n') han sortit de l'ou.
four already en.cl *have*.pl *leave*.pp *of the-egg*
'Four of them have already come out of the egg.'

(based on GLC 2016:699)

We adopt the assumption in Vallduví (2002) that preverbal NPs in Catalan (such as *quatre* in (35b)) are topics (not subjects) anaphorically related to an in-clause GF. Since the topic is the antecedent of an anaphoric pronoun (possibly null, as with null subjects) and anaphoric pronouns must be definite, it follows that topics cannot be related to the clitic *en*, because the lexical information of the *en* clitic specifies that it corresponds to an indefinite object. This makes it incompatible with its being an anaphoric pronoun dependent on the preverbal topic, thus explaining the ungrammaticality of the *en* clitic in (35b).

At this point, one may ask if it is possible to use a definite object clitic in place of the indefinite *en*, as it would qualify as a topic-anaphoric pronoun; the fact is that the definite object clitics *el/la/els/les* are incompatible with intransitive verbs:

(36) * Avui els surt/surten tard.
today them.obj.m.pl *leave*.sg/*leave*.pl *late*
'Today they are leaving late.'

Whichever agreement form of the verb is chosen, the core argument of the intransitive verb in (36) cannot be expressed by means of *els*. According to our analysis of (34), a definite argument of an intransitive verb is the subject. Since clitics like *el, la, els,* and *les* are (non-dative) *object* pronouns, they cannot be used as subjects, which explains the ungrammaticality of (36).[17]

---

are no expletive subjects in Catalan and there is no subject in a sentence like (31a). Languages with expletive subjects, such as French and English, have the opposite ranking of SC and this faithfulness constraint.

[16] Although we do not have space to give a detailed OT analysis of the phenomena considered here, the following ranking of constraints is assumed for Catalan: (25) » (33) » SC.

[17] The argument realization theory presented in this section has been developed in order to account for the facts of Catalan. It is beyond the scope of this paper to include this theory in a general theory of argument realization, defining the parameters of variation needed to account for cross-linguistic variation in this area. The approach may have points in common with Kiparsky's (1987) linking theory, but space limitations prevent us from making a detailed comparison.

4 Verbal agreement

In order to account for the idea that a verb can agree with either a subject or an object, if nominative, we follow Haug and Nikitina (2012, 2016), and Alsina and Vigo (2014, 2017), among others, in assuming that verbal agreement is mediated by the feature bundle AGR, which contains the agreement features encoded by the verb. Two general constraints, adopted from Alsina and Vigo (2014, 2017), are relevant to account for the agreement of the verb with one of its dependent GFs: the requirement that the clausal AGR feature be shared with that of a dependent GF (AGRSHARE (37a)), and the requirement that the agreeing GF be nominative (*AGRCASE (37b)):

(37) a. AGRSHARE: $\left[\begin{array}{l}\text{AGR } \boxed{1} \\ \text{DGF}\left[\text{AGR } \boxed{1}\right]\end{array}\right]_f$

For f-structure f that maps to a constituent of category V

b. * AGRCASE: $*\left[\begin{array}{ll}\text{AGR} & \boxed{1} \\ \text{GF} & \left[\begin{array}{ll}\text{AGR} & \boxed{1} \\ \text{CASE} & \neg\text{NOM}\end{array}\right]\end{array}\right]_f$

For f-structure f that maps to a constituent of category V

Thus, verbal agreement with a subject and with an object is represented as in (38a) and (38b), respectively:

(38) a. Avui surten els estudiants tard.
today leave.pl *the*.pl *student*.pl *late*
'Today the students are leaving late.'

$$\left[\begin{array}{ll}\text{PRED} & \text{`}\textit{leave}\text{ <Arg}_1\text{>'} \\ \text{AGR } \boxed{1} & \left[\begin{array}{ll}\text{PERS} & 3 \\ \text{NUM} & \text{PL}\end{array}\right] \\ \text{SUBJ} & \left[\begin{array}{ll}\text{PRED} & \text{`student'} \\ \text{DEF} & + \\ \text{AGR} & \boxed{1} \\ \text{CASE} & \text{NOM}\end{array}\right]_1\end{array}\right]$$

b. Avui en surten molts.
today en.cl *leave*.pl *many*.pl
'Today many are leaving.'

$$\left[\begin{array}{ll}\text{PRED} & \text{`}\textit{leave}\text{ <Arg}_1\text{>'} \\ \text{AGR } \boxed{1} & \left[\begin{array}{ll}\text{PERS} & 3 \\ \text{NUM} & \text{PL}\end{array}\right] \\ \text{OBJ} & \left[\begin{array}{ll}\text{PRED} & \text{`PRO'} \\ \text{DEF} & - \\ \text{AGR} & \boxed{1} \\ \text{QUANT} & \text{`many'} \\ \text{CASE} & \text{NOM}\end{array}\right]_1\end{array}\right]$$

In Catalan, a raising verb like *semblar* 'seem' can agree with the nominative object of the embedded clause:

(39) Semblen [arribar-ne molts.]
seem.3p.pl *arrive*.inf-*en*.cl *many*.pl
'Many seem to arrive.'

This is an instance of (apparent) long-distance agreement, as the inflected verb form *semblen* 'seem' in (39) doesn't seem to agree with any of its dependent GFs, but with the object *molts* 'many' in the infinitival complement clause. The only GF in the f-structure of *semblen* 'seem' that this verb could agree with is its complement clause, but, if the verb were to agree with it, it would

have to be in the third person singular form on the assumption that clauses agree in the third person singular. To solve this problem, we assume that long-distance agreement like the one in (39) is a combination of two local agreement relations, as in Alsina and Vigo (2017): **i)** the sharing of the AGR of the raising clause with the AGR of its infinitival complement, and **ii)** the sharing of this AGR with that of the object of the infinitive.

But not all verbs allow AGR sharing with the AGR of their embedded clause: only raising verbs do. To be formal, we assume a constraint, i.e., Clausal Opacity, which blocks the sharing of either AGR or GF in a given clause with either the AGR or a GF of its embedded clause. Raising verbs include a lexical specification overriding Clausal Opacity.

(40) Clausal Opacity:

$$*\begin{bmatrix} G & \boxed{1} \\ GF & [F\ \boxed{1}]_f \end{bmatrix}_g$$

For f-structures *f, g* that map to constituents of category V, and F, G = {DGF, AGR}

The cross-clausal agreement in (39) is possible because *semblar* 'seem' is a raising verb; thus, Clausal Opacity does not apply to f-structures whose PRED belongs to this verb, allowing both the structure-sharing of its subject with the subject of its infinitival complement (raising, as standardly understood) and the structure-sharing of its AGR with that of its infinitival complement ("raising" of the agreement features). So, the f-structure of (39) can be represented as:

(41)
$$\begin{bmatrix} \text{PRED} & \text{'}\textit{seem}\text{ <Arg1>'} \\ \text{AGR} & \boxed{1} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'}\textit{arrive}\text{ <Arg2>'} \\ \text{AGR} & \boxed{1}\begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO'} \\ \text{DEF} & - \\ \text{AGR} & \boxed{1} \\ \text{QUANT} & \text{'many'} \\ \text{CASE} & \text{NOM} \end{bmatrix}_2 \end{bmatrix}_1 \end{bmatrix}$$

Just like the raising of a subject is unbounded and can cross as many clauses as contain a raising verb, the raising of the agreement features is likewise potentially unbounded. All that is required is for there to be a chain of raising verbs overriding Clausal Opacity, as can be seen in the following example, where both *semblen* 'seem' and *tendir* 'tend' are raising verbs:

(42) Semblen [tendir [a arribar-ne molts]].
seem.3p.pl *tend*.inf *to arrive*.inf-*en*.cl *many*
'Many seem to tend to arrive.'

Once we have assumed that an intransitive argument can be a nominative object, we can explain the agreement facts, namely, the observation that the

verb agrees with its object and can be involved in long-distance agreement, adopting the agreement theory of Alsina and Vigo (2014, 2017) without additional assumptions.

5 Conclusion

This paper has argued for the claim that the single direct argument of an intransitive verb in Catalan can be a nominative object. This argument shows a subject-object alternation, but is invariably in the *nominative* case. The alternation is constrained by definiteness, so that the argument is a subject if it is definite and is an object if it is indefinite. As a subject, it displays the expected properties of a subject, including the possibility of pro-drop; as an object, it displays the expected properties of an object, including expression by means of the object clitic *en*. The claim that it is a nominative expression explains the observation that it agrees with the verb, even when it is an object, applying a theory of verbal agreement proposed independently of the facts of Catalan.

The theory of argument realization proposed in this paper is a simple one, as it assumes only two argument-to-GF mapping rules, three case assignment principles, and a small set of constraints restricting the GF assignment on the basis of the case features and definiteness of the arguments, in addition to morphosyntactic operations such as passivization.

References

Alsina, Alex. 1986. Assaig de definició de les funcions del pronom "en". In *Estudis de Llengua i Literatura Catalanes XII, Miscel·lània Antoni M. Badia i Margarit*, vol. 4, pp. 95-121. Barcelona: Publicacions de l'Abadia de Montserrat.

Alsina, Alex. 1995. The fall of function-argument biuniqueness. In Glyn Morrill and Richard 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, Alex. 1996. *The Role of Argument Structure in Grammar: Evidence from Romance*. Stanford: CSLI Publications.

Alsina, Alex. 2001. On the nonsemantic nature of argument structure. *Language Sciences*, 23, pp. 355-389.

Alsina, Alex and Sam Mchombo. 1990. The syntax of applicative in Chichewa: problems for a theta theoretic asymmetry. *Natural Language and Linguistic Theory*, 8(4), pp. 493-506.

Alsina, Alex and Sam Mchombo. 1993. Object asymmetries and the Chichewa applicative construction. In Sam Mchombo (Ed.), *Theoretical Aspects of*

Bantu Grammar, pp. 17-45. Stanford: CSLI publications.

Alsina, Alex and Eugenio M. Vigo. 2014. Copular inversion and non-subject agreement. In Miriam Butt and Tracy Holloway King (Eds.), *Proceedings of the LFG14 Conference*, pp. 5-25. Stanford: CSLI Publications.

Alsina, Alex and Eugenio M. Vigo. 2017. Agreement: interactions with case and raising. In Miriam Butt and Tracy Holloway King (Eds.), *Proceedings of the LFG17 Conference*, pp. 3-23. Stanford: CSLI Publications.

Bartra, Anna. 2009. Some remarks about the grammaticalization process of the DP functional domain in Old Romance. In Teresa Espinal, Manuel Leonetti, and Louise McNally (Eds.), *Proceedings of the IV Nereus International Workshop: Definiteness and DP Structure in Romance Languages*, pp. 1-26. Universität Konstanz.

Bresnan, Joan and Jonni Kanerva. 1989. Locative inversion in Chichewa: a case study of factorization in grammar. *Linguistic Inquiry* 20 (1), pp. 1-50.

Bresnan, Joan and Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21 (2), pp. 147-186.

Burzio, Luigi. 1986. *Italian Syntax: a Government-Binding Approach*. Dordrecht: Reidel.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge MA: The MIT Press.

Cortés, Corinne. 1995. Subject-object asymmetries and verb classes. *Linguistica Atlantica* 17, pp. 63-78.

Cortés, Corinne and Anna Gavarró. 1997. Subject-object asymmetries and the clitic *en*. In James. R. Black and Virginia Motapanyane (Eds.), *Clitics, Pronouns and Movement*, pp. 39-62. Amsterdam: John Benjamins.

Espinal, Teresa. 2010. Bare nominals in Catalan and Spanish. Their structure and meaning. *Lingua* 120(4), pp. 984-1009.

Espinal, Teresa and Louise McNally. 2010. Bare nominals and incorporating verbs in Spanish and Catalan. *Journal of Linguistics* 46, pp. 1-42.

Fabra, Pompeu. 1912. *Gramática de la Lengua Catalana*. Barcelona: L'Avenç.

Findlay, Jamie. 2016. Mapping theory without argument structure. *Journal of Language Modelling*, 4(2), pp. 293–338.

Givón, Talmy. 1984. *Syntax: a Functional-Typological Introduction*. Amsterdam: John Benjamins.

Gràcia, Lluïsa. 1989. *Els Verbs Ergatius en Català*. Maó, Menorca: Institut Menorquí d'Estudis.

Haug, Dag and Tanya Nikitina. 2012. The many cases of non-finite subjects: the challenge of "dominant" participles. In Miriam Butt and Tracy Holloway King (Eds.), *Proceedings of the LFG12 Conference*, pp. 292-331. Stanford:

CSLI Publications.

Haug, Dag and Tanya Nikitina. 2016. Feature sharing in agreement. *Natural Language and Linguistic Theory*, 34(3), pp. 865-910.

Institut d'Estudis Catalans. 2016. *Gramàtica de la Llengua Catalana* (GLC), pp. 697-699. Barcelona: Institut d'Estudis Catalans.

Kibort, Anna. 2001. The Polish passive and impersonal in Lexical Mapping Theory. In Miriam Butt and Tracy Holloway King (Eds.), *Proceedings of the LFG01 Conference*, pp. 163-183. Stanford: CSLI Publications.

Kibort, Anna. 2009. Intermediary agents and unexpressed pronouns. In Miriam Butt and Tracy Holloway King (Eds.), *Proceedings of the LFG09 Conference*, pp. 378-398. Stanford: CSLI Publications.

Kiparsky, Paul. 1987. *Morphology and Grammatical Relations*, ms. Stanford CA: Stanford University.

Kuhn, Jonas. 2003. *Optimality-Theoretic Syntax: a Declarative Approach*. Stanford: CSLI Publications.

Levin, Lorraine. 1986. *Operations on Lexical Forms: Unaccusative Rules in Germanic Languages*. PhD dissertation. MIT.

Lødrup, Helge. 1999. Linking and optimality in the Norwegian presentational focus construction. *Nordic Journal of Linguistics*, 22, pp. 205-230.

Mohanan, Tara. 1994. *Argument Structure in Hindi*. Stanford: CSLI Publications.

Patejuk, Agnieszka and Adam Przepiórkowski. 2016. Reducing grammatical functions in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King and Stefan Müller (Eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, pp. 541-559. Stanford: CSLI Publications.

Saccon, Graziella. 1995. *Ne*-cliticization does not support the unaccusative/intransitive split. In Glyn Morrill and Richard Oehrle (Eds.) *Formal Grammar: Proceedings of the Conference of the European Summer School in Logic, Language, and Information*, pp. 227-238. Barcelona: Universitat Politècnica de Catalunya.

Sigurðsson, Halldór Á. 2004. Icelandic non-nominative subjects: facts and implications. In Peri Bhaskararao and Karumuri Venkata Subbarao (Eds.), *Non-nominative Subjects*, vol. 2, pp. 137-159. Amsterdam: John Benjamins.

Vallduví, Enric. 2002. L'oració com a unitat informativa. In Joan Solà, Maria Rosa Lloret, Joan Mascaró, and Manuel Pérez Saldanya (Eds.), *Gramàtica del Català Contemporani*, vol. 2, pp. 1221-1279. Barcelona: Editorial Empúries.