F-to-c-structure mapping: accounting for inflectional morphology and periphrasis

Alex Alsina

Pompeu Fabra University

Proceedings of the LFG'23 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2023

PubliKon

lfg-proceedings.org

Abstract

The treatment of inflectional periphrasis is problematic in LFG, apparently because of the lexicalist nature of the framework. A close inspection of what is usually understood by lexicalism reveals two distinct, but related, notions: lexicalism and lexical encapsulation. Complex inflectional systems show that one can preserve lexicalism (the idea that words and phrases are different in terms of units and rules of composition), but that it is necessary to reject lexical encapsulation (the idea that words are formed without input from syntax). An adequate theory of inflectional morphology needs a framework that is not constrained by lexical encapsulation. With such a framework, it is then possible to give a correct account of inflectional periphrasis. The paper develops the analysis of two periphrastic constructions, one in Latin and one in Catalan, within a non-encapsulated version of LFG.

1 Introduction

This paper addresses the treatment of inflectional periphrasis in LFG.[1] Inflectional periphrasis, defined as a two-word (or multi-word) expression that alternates with single word forms in an inflectional paradigm, poses a serious problem for the standard LFG conception of the correspondence between the syntax and the morphology (or word formation). LFG standardly assumes an asymmetrical relation between these two components: phrasal syntax uses the information provided by word forms and is therefore constrained by it, but word formation cannot use the information provided by phrasal syntax and therefore cannot be constrained by it. Previous work (Alsina 2020; 2022; Alsina & Vigo 2017) has argued for the need to allow word formation to be sensitive to f-structure information. If the same set of f-structure features can alternatively map onto a single verb form or onto two verb forms, depending on the values of those features, it is hard to see how this can be achieved if the f-structure information has to be constructed from the information provided by the word forms. On the other hand, if word forms are constrained by the features in the f-structure, it becomes feasible to account for an alternative mapping to either one or more than one verb form.

Given the assumption that inflectional morphology takes f-structure information as the input for its rules, as argued in previous work, all we need to assume for periphrasis is that a specific f-structure maps onto two words, the minimal units of c-structure, each one undergoing its own morphology. The approach will be illustrated with two phenomena: the classical problem of the alternation between synthetic and periphrastic forms in the Latin passive/deponent conjugation (Börjars et al. 1997; Sadler & Spencer 2001) and the periphrastic past perfect in Catalan, exemplified in (1) and (2) respectively.

[1]I thank three anonymous reviewers, Ash Asudeh, Miriam Butt, Dan Siddiqi, Nuo Xu, and other members of the audience at the LFG'23 Conference for their questions and comments.

(1) *Fratres* *locuti* *sunt*.
brother.NOM.M.PL speak.PST.PTCP.NOM.M.PL be.PRES.3.PL
'The brothers spoke.'

(2) *Va* *parlar* *l'* *advocat*.
VA.3SG speak.INF the lawyer
'The lawyer spoke.'

Section 2 lays out the main assumptions of the new proposal, which we may call the bidirectional c-f mapping, and compares it with the standard assumptions of LFG. Section 3 presents the relevant data and the analysis of the Latin synthetic/periphrastic alternation within the bidirectional c-f mapping. Section 4 presents the phenomenon and analysis of the Catalan periphrastic past perfect. And section 5 brings out the main conclusions of the paper.

2 The bidirectional c-f mapping approach

2.1 Two hypotheses of the standard LFG framework

One of the distinctive properties of LFG as a framework is that it is a lexicalist framework. This means that it adheres to the lexicalist hypothesis, a design feature of the framework. The lexicalist hypothesis has received different names and slightly different definitions through the years (see O'Neill 2016), but a widely accepted version of this hypothesis within LFG and the one we will use here is Bresnan & Mchombo's (1995, 181), stated as (3):

(3) *Lexical integrity principle (LIP)*: words are built out of different structural elements and by different principles of composition than syntactic phrases. Specifically, the morphological constituents of words are lexical and sublexical categories – stems and affixes – while the syntactic constituents of phrases have words as the minimal, unanalyzable units.

See Mohanan (1995) and Bresnan et al. (2016, 92) for other formulations of this principle. A consequence of (3) is that the rules of phrasal syntax, or c-structure, cannot involve elements below the level of the word, such as affixes and roots. A framework that adheres to the design property in (3) is a lexicalist framework.

Another widely adopted assumption within LFG is the idea that word formation and, therefore, inflectional morphology takes place in a module isolated from syntax called the lexicon. The sense in which the lexicon is isolated from syntax can be expressed as the idea that words are encapsulated, as in the following statement: [2]

(4) *Lexical encapsulation hypothesis (LEH)*: Although words can provide information for constructing syntactic structures, syntax cannot provide information that can be used in forming words.

[2] See Brown & Hippisley (2012, 17) for the notion of encapsulation as used here.

The LEH enforces an asymmetrical relationship between morphology and syntax: words carry much of the syntactic information, particularly, f-structure information for constructing, or licensing, the syntactic structures in which the words appear, but these syntactic structures cannot be used in word formation. If we use a temporal metaphor to express this relationship, we can say that, according to the LEH, morphology takes place before syntax: since words are constructed before syntactic structures, there is no (strictly) syntactic information that word formation rules can access. On the other hand, syntactic structures are built using the information carried by the words in those structures. We can call a framework that incorporates the design feature in (4) lexically encapsulated or, simply, encapsulated.

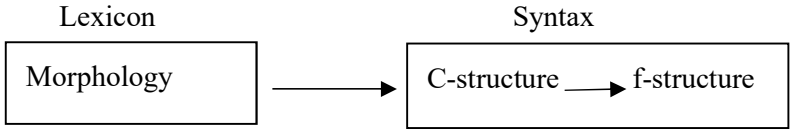
It might appear as though the LEH is just another name for the lexicalist hypothesis. The fact is that, although both the LEH and the LIP are about the relationship between words and larger syntactic structures and, therefore, between morphology and syntax, these two principles are conceptually different and one could adopt the LIP without adopting the LEH. Standard versions of LFG are both lexicalist and lexically encapsulated. They are lexicalist because they do not allow sublexical categories, such as roots and affixes, to appear as syntactic categories in the c-structure, as words are the "minimal, unanalyzable units" of c-structure. And they are lexically encapsulated because morphologically complex words are assumed to exit the lexicon fully formed and be placed as terminal elements in the syntax with all the information provided by the lexicon.[3] Lexical encapsulation, the LEH (4), implies lexical integrity, the LIP (3): it is not possible for a framework to incorporate the former but not the latter. This may give the impression that the implication also goes the other way around. However, one could assume the LIP without assuming the LEH. In other words, one could have a framework that is lexicalist, like current forms of LFG, but is not lexically encapsulated. This is the version of LFG that I will argue in the remainder of this paper that one needs for inflectional morphology and for inflectional periphrasis.

The way the LIP and the LEH are implemented in standard versions of LFG is to assume (a) that morphology, or word formation, is a system of rules located within the lexicon responsible for accounting for the form and content of words and (b) that the lexicon and the syntax are two different modules of the grammar that are related in a unidirectional way such that the lexicon outputs words, which are the terminal nodes of the c-structure, and cannot use information from the syntactic c- and f-structures. See Dalrymple (2015) and Dalrymple et al. (2019) for an implementation of these ideas. These works assume a realizational approach to morphology, following proposals by Sadler and Spencer (2001), Stump (2001; 2006; 2016), Spencer (2013), among others, as opposed to a morpheme-based and incremental approach to morphology, as in previous versions of LFG (e.g. Bresnan

[3] Even if the notion of encapsulation is not always made explicit, it is implicit in many ways. For example, the phrase "When the inflected verb is inserted into the c-structure…" (Bresnan et al. 2016, 53) implies that the inflected verb is formed in a module different from the syntax, the lexicon, where the word formation rules operate without recourse to the information in the syntax.

et al. 2016). Nevertheless, both approaches to morphology are consistent with the LIP and the LEH. The relationship between morphology, contained in the lexicon, and syntax, in a lexicalist encapsulated version of LFG is depicted as in (5):

(5) *Relationship between morphology and syntax in standard LFG*:

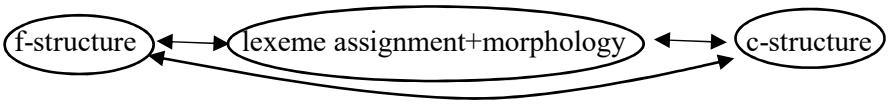


The arrow connecting the lexicon to the syntax in (5) indicates that the lexicon operates without using information in the syntax, whereas the syntax incorporates the output of the lexicon, namely, words (i.e., word forms or lexical entries in Dalrymple et al. (2019)). The arrow connecting the c-structure to the f-structure indicates that the f-structure is projected from the c-structure.

2.2 A non-encapsulated version of LFG

The approach assumed in this paper is one in which the lexicon is not isolated from the syntactic structures but operates alongside these structures and receives input from them. Specifically, inflectional word formation is sensitive to information in the syntactic structures, in particular, f-structure. Within this lexicalist non-encapsulated version of LFG, the relationship between morphology, along with lexeme assignment or lexeme choice, and the syntactic structures is as shown in (6):

(6) *Relationship between morphology and syntax in a non-encapsulated LFG*:



According to this representation, both c-structure and f-structure are in a mutually constraining relation with lexeme assignment and morphology and with each other. Whereas in the encapsulated version of LFG shown in (5) word formation or morphology is in a sequential relation to phrase formation or syntax,[4] in the non-encapsulated version in (4) these two modules are not in a sequential relation, but in a mutually feeding relation.[5]

Evidence for the morphology-syntax interface shown in (6) is given in previous work. Alsina & Vigo (2017) address the phenomenon of direct and inverse agreement in Plains Cree. Choosing the right affix among those that signal either direct or inverse agreement depends on evaluating the relative prominence of subject and object along a person-obviation scale. If the subject is more prominent,

[4] This sequential relation between morphology and syntax is the traditional view in linguistics, according to Julien (2007).

[5] The proposal depicted in (6) does not change the standard LFG idea that the mapping from c- to f-structure is a function, so that the relationship between c- and f-structure is many-to-one, whereas the inverse is merely a relation, allowing for one-to-many relationships. This is not affected by the claim that the c-structure may be constrained by information in the f-structure.

one of the direct affixes is selected; if the object is more prominent, one of the inverse affixes is selected. This means that the f-structure information with the object and the subject and their features needs to be available to the word formation rules.[6] This evidence clearly argues against the encapsulation hypothesis (4).

Alsina (2020) proposes an analysis of the presence or absence of the reflexive "clitic" in the Romance language Catalan.[7] Two conditions have to be met for the appearance of the reflexive "clitic": that it be selected by a reflexive predicate (a predicate that lexically belongs to the reflexive class or that is of the appropriate semantic type), even though it need not be morphologically attached to this predicate but may be attached to another predicate in a "restructuring construction," and that the logical subject of this predicate be encoded as a grammatical subject. If there is a reflexive predicate, but its logical subject maps onto an object, because it is part of a causative complex predicate, there is no reflexive "clitic". Since the formation of these complex predicates takes place in the syntax, not in the lexicon, because the causative predicate and its dependent predicate are syntactically independent words, the mapping of the arguments of this predicate to grammatical functions also has to take place in the syntax. Therefore, the explanation of the variable expression of the reflexive "clitic" crucially depends on information at the f-structure being available to the word formation rules.[8]

Alsina (2022) compares the syntax and morphology of two Romance languages, Catalan and Spanish, of which the former has a partitive "clitic" and the latter does not. In addition to having this morphological difference, they also have an apparently syntactic difference: Catalan does not allow null indefinite objects without a morphological exponent of the object (on the assumption that the partitive "clitic" is a morphological exponent of the object), whereas Spanish does. These two properties do not vary independently of each other: lacking the partitive "clitic" implies allowing null indefinite objects without any morphology to go with it, whereas having the partitive "clitic" implies the opposite. The standard encapsulated versions of LFG cannot account for this co-variation by means of just one parameter of variation. The non-encapsulated version assumes that the two languages have the same syntax and only differ on the morphological side. The co-variation is explained only if we allow the syntax to feed the morphology.

---

[6] Whether these facts can be "captured by an equation (or set of equations) introduced by each affix stating the relevant person-obviation relation(s) between SUBJ and OBJ," as suggested by a reviewer, is hard to determine, because an analysis along these lines has not been developed. Also, such an analysis entails a morpheme-based incremental approach to inflectional morphology, which suffers from the objections that have been raised to morpheme-based incremental approaches in general by Anderson (1992), Stump (2001), Spencer (2013), among others.

[7] The particles traditionally called clitics in Romance linguistics have been shown to be affixes (see references in Alsina (2020; in press)) and, since they are a special type of affixes, they are referred to here as "clitics", in quotes, to avoid confusion.

[8] An anonymous reviewer suggests an alternative analysis according to which reflexive verbs would have a disjunction in their lexical entry by which either the presence of a reflexive clitic is required, by means of a constraining equation, or its logical subject is distinct from its grammatical subject. I cannot adequately evaluate this proposal, because it is underdeveloped.

2.3 Problems with the Lexical Encapsulation Hypothesis

Proponents of the encapsulated LFG – any version of LFG that adheres to the LEH (4) – who do not assume the more traditional morpheme-based incremental approach to morphology, but have adopted a realizational approach, such as Luís and Sadler (2003), Dalrymple (2015) and Dalrymple et al. (2019), try to circumvent the kinds of objections just presented by incorporating all the syntactic information that is needed for word formation into the lexicon.[9] This is done by copying the necessary f-structure information in the form of m-features (for *morphological features*). The name m-features is a source of confusion because it includes truly morphological features, such as declension class or conjugation class, and the copies of f-structure features. Alsina (2022) suggests calling the latter ms-features (for morphosyntactic features). By including the ms-features in the lexicon, we make it possible for word formation rules to access the necessary syntactic information while still apparently complying with the LEH.

If the type and amount of syntactic information that needs to be copied into the lexicon for the purpose of word formation rules were very limited and restricted in clear ways, it might provide an argument for maintaining the LEH: the access to syntactic information in word formation would be limited and constrained in very specific ways. Spelling out exactly which f-structure information has to be copied as m-features and which f-structure information cannot be copied as such would constitute a testable hypothesis about what syntactic information can be used in word formation and what syntactic information cannot. But the fact is that no limits have been argued to exist. If we look at the arguments against the LEH mentioned in the preceding paragraphs, we see that a considerable amount of f-structure information would have to be copied into the lexicon: the distinction between subject and object, with their respective person and obviation features, for Plains Cree; the distinction between subject, object, and oblique, with features such as case, person, number, definiteness, the distinction between old and new information, for the Romance languages. The conclusion one can draw from this is that, by copying into the lexicon all the f-structure information that is needed for word formation, the LEH is preserved in appearance, but this is done at the very high cost of having the same information represented in two different places in the grammar: a massive violation of the simplicity criterion, or Occam's Razon, which governs theory design.

Furthermore, as the f-structure information has to be copied into the representation of each word that occurs in a sentence, the same f-structure information may have to be copied several times—once for each word that needs that information in its formation. To illustrate this point, take a sentence in a language like Catalan,

[9] Morpheme-based incremental approaches to morphology, which have been adopted in earlier versions of LFG (e.g., Bresnan and Mchombo 1987; Bresnan et al. 2016), cannot be criticized on the same grounds. While they also assume the LEH, they do not copy f-structure information as m-features. The idea of such approaches is that morphemes such as affixes contribute part of the f-structure information associated with the fully inflected word form.

or any other language where the head noun in an NP agrees in gender and number with its determiner and with modifying adjectives and with an AP as a predicative complement of the verb (traditionally called an XCOMP) and agrees in person and number with the verb, as in (7):

(7) *Aquestes formes estranyes semblen imaginàries*.
this.F.PL form.F.PL strange.F.PL seem.3PL imaginary.F.PL
'These strange forms seem imaginary.'

Intuitively, there is only one set of person, number, and gender features, which we can think of as being part of the subject. These features, which, in (7), are third person feminine plural, are reflected morphologically in the form of the determiner, the noun, and the adjectival modifier in the subject NP, of the verb, and of the adjective in the predicative complement. If we allow word formation rules to refer to f-structure information, as advocated here, one set of f-structure features is sufficient, as the rules will license several different exponents of these features. If, on the other hand, we are determined to preserve the LEH, albeit only in appearance, we are forced not only to have these features in the f-structure, but also to repeat them in the form of m-features in the lexical information set of each of the five words in (7). These features need to be in the lexical entry of each word so that the morphological rules, which operate in the lexicon, can refer to them. In short, an approach that formally adheres to the LEH entails a radical dismissal of the simplicity criterion.

It is clear that the LEH needs to be abandoned and the framework has to be rethought without this design principle. The consequence of abandoning the LEH is that word formation (at least, inflectional word formation) takes place simultaneously with syntax and, crucially, word formation rules can refer to information in the syntactic structures, such as f-structure. As proposed in Alsina (2020; 2022), the f-structure should play the role of the morphosyntactic representation (MR) or set of morphosyntactic properties of Anderson (1992), Stump (2001), and others. It is not necessary to introduce a new level of representation as the f-structure has all the information that is assumed to be required in the MR. The rules of the morphology take into account the f-structure information in order to determine the phonological form of the word (its fully inflected morphological form). If we assume that the phonological form of a word is part of the c-structure, we can say that the morphology involves a mapping of f-structure information onto c-structure, together with a given lexeme choice, as depicted in (6).[10]

The model represented in (6) is not only the adequate one for inflectional morphology, but is also the right model for accounting for inflectional periphrasis. We

[10] An anonymous reviewer observes that, from a computational perspective, it is surprising to claim that, in standard versions of LFG, syntactic structures cannot take part in word formation, because syntactic rules influence word formation – during generation – and word formation influences syntactic structure – during parsing –. It may well be that computational linguists are not constrained by the same principles as theoretical linguists and that work within computational linguistics has reached similar conclusions to those reached here.

can assume that there are two kinds of rules that map f-structure information onto c-structure: (a) word formation rules or rules of the morphology, which affect the phonological shape of words, and (b) lexeme selection rules, which require the presence of a particular lexeme. A lexeme selected by one of these rules, just like any other lexeme, undergoes word formation rules in order to obtain the right form of the lexeme in a given syntactic context. The remainder of this paper illustrates how the model represented in (6) is adequate for accounting for inflectional periphrasis by means of two periphrastic constructions: the Latin perfect forms of the passive/deponent conjugation and the periphrastic past perfect in Catalan.

3 The Latin periphrastic perfect

3.1 Latin synthetic and periphrastic forms of verbs

The Latin conjugation shows an alternation between synthetic and periphrastic forms that has led many works, from pedagogical grammars to studies such as Börjars et al. (1997) and Sadler & Spencer (2001), to assume that the same system of rules is responsible for deriving both classes of forms. The morphology of verb forms depends on a set of morphosyntactic features such as tense, aspect, mood, person and number of the subject, and a strictly morphological classification of verb lexemes into different conjugation classes. In addition, many verbs have different forms depending on whether they are used in the active voice or in the passive voice. We can say there is an "active" conjugation and a "passive" conjugation, which we can refer to as the A conjugation and the P conjugation respectively. For verbs that can be used either in an active structure or in a passive structure, the A conjugation signals the former and the P conjugation signals the latter. As an illustration of the two conjugations for the same verb lexeme, (8) shows the different forms of the present indicative of the lexeme AMO 'love' (specific lexemes are designated by the citation form of the lexeme in small caps):

(8)

Subj. features	A conjugation	P conjugation
1.SG	*amo*	*amor*
2.SG	*amas*	*amaris*
3.SG	*amat*	*amatur*
1.PL	*amamus*	*amamur*
2.PL	*amatis*	*amamini*
3.PL	*amant*	*amantur*

There is not a perfect correspondence between these two conjugations and a passive vs. active structure, as some verbs can only be inflected according to the P conjugation and yet have an active structure—so-called deponent verbs—and another class of verbs are inflected according to the A conjugation in the imperfective forms and according to the P conjugation in the perfective forms, within an active structure—so-called semideponent verbs. This suggests that the A/P conjugation

can be modeled as strictly morphological features, or m-features, as opposed to morphosyntactic features, which in the present proposal are features of the syntactic structures. The conjugation class of a verb (whether the verb belongs to one of the four regular conjugation classes or to an irregular paradigm) is also an m-feature, but, unlike the A/P conjugation, is a property that is listed as part of the information of each lexeme. In contrast, the assignment of the A or P conjugation to a given verb form may depend on the lexeme, on the syntactic information of the verb form, or on a combination of the two.

We can formalize this idea by proposing that all verb forms have a morphological representation that includes the feature [CONJUG A/P], a feature that takes one of the two values A or P, corresponding to the A and P conjugation respectively, in addition to other features. All features that appear in the morphological representation are strictly morphological features, since morphosyntactic features are part of the syntactic structures. The assignment of either of the two values of the feature [CONJUG A/P] proceeds as follows:

(9) Assignment of the features [CONJUG A] and [CONJUG P]:

The feature [CONJUG P] is selected
a) by verb forms whose lexeme information includes this feature;
b) by verb forms whose lexeme information includes this feature only in cooccurrence with the f-structure feature [PERFECTIVE +];
c) by passive verb forms (verb forms with an a-structure whose highest argument is suppressed).

Elsewhere, the feature [CONJUG A] is selected.

Case a) in which the feature [CONJUG P] is selected corresponds to deponent verbs, such as LOQUOR 'speak' or SEQUOR 'follow'. Any property that is part of the information of a lexeme is also part of each of the lexeme's forms. As for case b), we can assume that semideponent verbs carry a conditional requirement stating that, if a verb form is associated with the feature [PERFECTIVE +], it is marked with the feature [CONJUG P]. According to case c), a verb form that corresponds to an a-structure whose highest argument is suppressed also includes this feature. In all other cases, the alternative feature [CONJUG A] is required, which captures the idea that the "active" conjugation is the default. Both cases b) and c) of (9) show that a morphological feature (i.e. [CONJUG A/P]) may depend on a syntactic feature: in case b), the morphological feature [CONJUG P] is only triggered in the presence of the syntactic feature [PERFECTIVE +]; and, in case c), it is triggered by a suppressed highest argument.

What is relevant about the A and P conjugation from the point of view of inflectional morphology and periphrasis is that all forms of the A conjugation (both perfective and imperfective) and all imperfective forms of the P conjugation are synthetic, i.e. single word forms, whereas the perfective forms of the P conjugation are periphrastic, specifically consisting of a form of the perfect participle and a form of the verb SUM 'be' in an imperfective tense. To illustrate this point, (10)

shows an imperfective form, present, and a perfective form, past, in the A conjugation and the corresponding forms in the P conjugation, for the lexeme AMO, all of these forms showing third person singular agreement:

(10)		*Present, imperfective*	*Past, perfective*
	A conjugation:	amat 's/he loves'	amavit 's/he loved'
	P conjugation:	amatur 's/he is loved'	amatus est 'he was loved'

The periphrastic form *amatus est* is functionally equivalent to the synthetic *amatur*, except that the former is perfective and the latter is imperfective.[11] And, just as in the active form the notion of past perfective is conveyed by the suffixal chunk *–avit* in *amavit*, in the passive form the same notion is conveyed by the combination of the past participle *amatus* and the present form *est*.

If we tried to account for the periphrastic forms in the Latin P conjugation within a standard lexically encapsulated version of LFG, we would be faced by many problems. The claim of a framework that adheres to the LEH is that words are inserted in the syntax with a set of syntactic features that are contributed to the syntax. However, the features associated with each of the words that make up the periphrasis are not the features associated with the periphrasis. E.g., *est* (elsewhere) is present and imperfective, whereas the periphrasis *amatus est* is past and perfective. What happens to the features of the lexical item *est*? Where do the features of the periphrasis come from? What prevents *fuit* (past perfective counterpart of *est*) from being used with the same value in the periphrasis?

3.2 The Latin periphrasis rule

If, on the other hand, we adopt a non-encapsulated version of LFG, as advocated here, these problems disappear. In this approach, the phonological (or fully inflected) form of words is conditioned by information in the f-structure. Given that words are the terminal nodes of the c-structure, it follows that rules that map the f-structure onto the c-structure may affect the form of words. We can assume that they may also license the presence of words. This is what happens in a periphrasis: the presence of specific features in the f-structure requires the presence of specific words in the c-structure. We can call rules that require the presence of a form of a particular lexeme on the basis of f-structure information *lexeme selection rules*.

An example of a lexeme selection rule is the one we need for the periphrastic forms of the P conjugation in Latin. The f-to-c mapping principles may simply involve the application of morphological rules to the words in the structure. If the lexeme AMO is used in its active form, the morphological rules produce a word

[11] I am setting aside the fact that the periphrastic form, as it includes a participle, which, being categorially an adjective, is inflected for gender, number, and case, signals the gender of the subject, in addition to its person and number, whereas a finite verb form such as *amatur* only signals the latter two features of the subject. So, the P conjugation counterpart of the past perfective *amavit* is *amatus est*, *amata est*, and *amatum est*, which indicate that the subject is of masculine, feminine and neuter gender, respectively.

form with the morphology of the first conjugation and the A conjugation. If it is used in its passive form, the passive a-structure triggers the assignment of the m-feature [CONJUG P], as seen in (9). If, in addition to having this m-feature, the word is associated with an f-structure containing the feature [PERFECTIVE +] (for perfective aspect), a specific lexeme selection rule is activated which makes the f-structure in question map onto two words: the past participle of the verb associated with the feature [CONJUG P] and a form of SUM in the imperfective aspect. A formalization of this rule is given in (11):

(11) Periphrastic SUM licensing rule:

$$\left[\begin{array}{ll}\text{PERF} & +\\ \text{FINITE} & +\end{array}\right]_{1,f} \wedge \left[\begin{array}{ll}\text{LEXEME} & \text{X}\\ \text{CONJUG} & \text{P}\end{array}\right]_{1,m} \longrightarrow \left\{\begin{array}{l}\left[\begin{array}{ll}\text{LEXEME} & \text{X}\\ \text{V-FORM} & \text{P-PART}\\ \text{FINITE} & -\end{array}\right]_{1,m}\\ \left[\begin{array}{ll}\text{LEXEME} & \text{SUM}\\ \text{PERF} & -\end{array}\right]_{1,m}\end{array}\right\}$$

The left-hand subscript on the feature structures in (11) distinguishes f-structures, with subscript *f*, from morphological structures, with subscript *m*. Coindexation of structures, with a right-hand subscript, signals correspondence between levels: the f-structure on the left of the arrow in (11) is in a mapping or correspondence relation with the coindexed morphological structures (or words). Thus, the rule in (11) says that, if there is an f-structure with the features [PERF +, FINITE +] which maps onto a word with the m-feature [CONJUG P], this word, whatever its lexeme is, has to be a non-finite past participle form and there also has to be a [PERF −] form of the lexeme SUM in a mapping relation with the same f-structure. The two morphological structures on the right of the arrow in (11) undergo the rules of the morphology that yield the appropriate inflected forms of the two lexemes involved.

The morphological structures of words only specify strictly morphological information, not present at f-structure. The rules of the morphology have access to the information both in the morphological structures and in the f-structure. In Latin a finite verb form shows agreement in person and number with its subject. The subject information is in the f-structure that corresponds to that verb: the morphological rules access this information and assign the appropriate inflections to the verb. In the case of an f-structure that maps onto two words – a periphrasis –, such as the structure that results from the application of rule (11), both the copula and the non-finite past participle form of the main verb reflect the features of the subject. The copula agrees in person and number and the past participle, as an adjective, reflects the gender, number, and case features of the subject. This accounts for contrasts such as the following:

(12) a. *Discipulus* *amatus* *est*.
student.NOM.M.SG love.PST.PTCP.NOM.M.SG be.PRES.3.SG
'The (male) student was loved.'

b. **Discipulus* *amatae* *sunt*.
student.NOM.M.SG love.PST.PTCP.NOM.F.PL be.PRES.3.PL

c. *Discipulae* *amatae* *sunt*.
student.NOM.F.PL love.PST.PTCP.NOM.F.PL be.PRES.3.PL
'The (female) students were loved.'

d. **Discipulae* *amatus* *est*.
student.NOM.F.PL love.PST.PTCP.NOM.M.SG be.PRES.3.SG

(13) a. *Fratres* *locuti* *sunt*.
brother.NOM.M.PL speak.PST.PTCP.NOM.M.PL be.PRES.3.PL
'The brothers spoke.'

b. **Fratres* *locuta* *est*.
brother.NOM.M.PL speak.PST.PTCP.NOM.F.SG be.PRES.3.SG

c. *Mater* *locuta* *est*.
mother.NOM.F.SG speak.PST.PTCP.NOM.F.SG be.PRES.3.SG
'The mother spoke.'

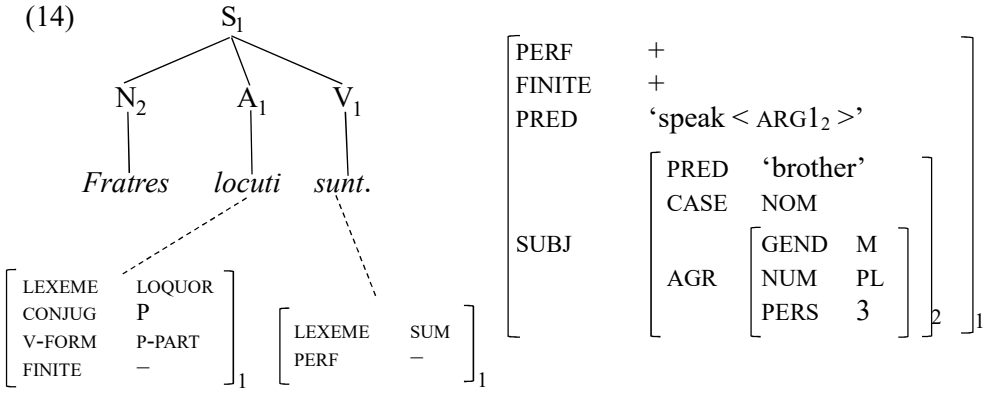
d. **Mater* *locuti* *sunt*.
mother.NOM.F.SG speak.PST.PTCP.NOM.M.PL be.PRES.3.PL

The two sets of examples (12) and (13) illustrate the use of a perfective form of the P conjugation. In (12) the P conjugation is required by the passive a-structure and in (13) it is required by the lexical stipulation of the deponent lexeme LOQUOR. Regardless of the condition that triggers this conjugation, the syntactic and morphological properties of the construction are the same in both (12) and (13): the past participle has to agree in gender and number, as well as case, with the subject and the finite form of SUM has to agree with it in person and number.

Crucially, although the morphology has to have access to the f-structure features, the access to a particular f-structure feature is blocked in case that same feature is specified with a different value in the morphological structure. Thus, for example, the form of the lexeme SUM in the output of rule (11) is assigned the feature [PERF −]: this feature is incompatible with the feature [PERF +] in the corresponding f-structure and consequently the morphology ignores the latter. *Est* and *sunt* in (12) and (13) are morphologically imperfective, but the corresponding f-structure, and semantics, is perfective. Rules that provide a phonological spell-out of morphosyntactic features give preference to the features in the morphological structure when they are inconsistent with equivalent f-structure features.

The c- and f-structures corresponding to sentence (13a) are shown in (14), together with the morphological representation of the two words that make up the periphrasis. The correspondence between nodes in the c-structure and feature structures in the f-structure is shown by coindexation.

(14)



There are aspects of these representations that I am assuming without argument, as they are not relevant for the issues under discussion in this paper. For example, whether the noun that maps onto the subject is part of an NP or not has no bearing on these issues. Latin is well-known to allow “discontinuous grammatical functions”: constituents that map onto the same GF, such as subject or object, may be separated by constituents belonging to other GFs. We can capture this idea by allowing non-phrasal categories to map onto GFs on the basis of their case features. The agreement features of the subject are grouped as the feature structure AGR, following proposals such as Alsina & Vigo (2014) among others, although, again, this is not a crucial aspect.

What is significant is that both the past participle (*locuti* in (14)) and the form of the copula (*sunt* in (14)) show agreement with the subject. In the present proposal, these agreement phenomena are not syntactic phenomena, but strictly morphological or pertaining to the morphology-syntax interface. Forms such as *locuti* or *sunt* do not have lexical entries that specify the person, gender, and number features of their subject; in other words, they do not carry f-structure specifications signalling that their subject is masculine plural or third person plural. Once a lexeme is selected, in this case, LOQUOR and SUM, the appropriate form of the lexeme to be used will depend on the morphological and syntactic features associated with the form in question and on the rules of the morphology, or word formation rules, which are sensitive to those features. For example, the feature [V-FORM P-PART] in the morphological representation of the form of the lexeme LOQUOR in (14) triggers a rule or set of rules that associates the stem /lokut/ with this form. In addition, the presence of the f-structure features of nominative case, masculine, and plural of the subject of the clause that maps onto the form in question triggers a rule that adds the suffix /i/ to that stem, so that the phonological sequence /lokuti/ is selected as the right form to use. If the subject were masculine singular, instead of plural, the rules of the morphology would select the suffix /us/ to be added to the stem /lokut/, giving as the final form /lokutus/.

This view allows us to drastically simplify the information that is part of lexical entries. We do not have separate lexical entries for each of the forms of a lexeme, each one with an array of m-features that copy information present in the f-

structure in order for the rules of the morphology, or word formation rules, to have access to the syntactic information, which would be unavailable as such because of the LEH. We just have a lexical entry for the lexeme, which has to include strictly morphological information such as its morphological class, the various stems it has, with their syntactic distribution, etc., and a set of rules of the morphology needed for selecting the right phonological form in each syntactic context.

What is special about a construction like that represented in (14) is that a periphrasis-licensing rule like (11) maps an f-structure onto two verb forms. Each of these forms undergoes the set of morphological rules to select the appropriate phonological form of the lexemes specified in the morphological representation of the words. In order for the structures in (14) not to incur a violation of LFG's Uniqueness Condition, we can assume that the auxiliary verb, in this case, the lexeme SUM, has an optional PRED feature. When this verb is used as the only PRED-bearing element in the construction, it is selected with its PRED feature. When it co-occurs with a PRED-bearing element (such as the past participle adjective in the perfective periphrasis), the option without the PRED feature is chosen.

We can assume that a semantically empty verb, such as the auxiliary SUM, without a PRED feature, can only be used in order to satisfy a principle of the grammar, appealing to Bresnan et al.'s (2016: 90) Economy of Expression. In this way, there is a single lexeme SUM, whether used as auxiliary or as main verb: it is used as an auxiliary, without PRED, if required by the periphrasis-licensing rule (11). Otherwise, it is used as a main verb.

4 The Catalan periphrastic past perfect

4.1 The Catalan past perfect periphrasis rule

Catalan has two ways of expressing the past perfect tense of a verb: a single word form and a periphrastic construction. The former option is illustrated in (15a) and the latter in (15b).

(15) a. *Parlà* *l'* *advocat*.
speak.PST.PRF.3SG the lawyer
'The lawyer spoke.'

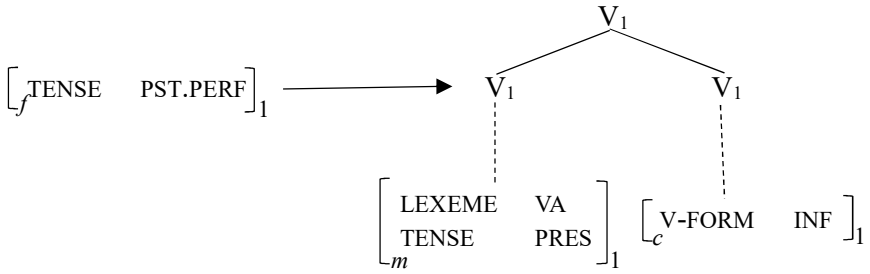
b. *Va* *parlar* *l'* *advocat*.
VA.3SG speak.INF the lawyer
'The lawyer spoke.'

The two forms are semantically equivalent. However, in many geographic dialects, including the Barcelona area, the synthetic form is highly restricted in its register: it is only used in very formal styles. In contrast, the periphrastic form is used in all registers. In some other dialects, such as central Valencian, the two forms are equivalent also with respect to register. In what follows, only the periphrastic construction will be discussed.

The periphrastic past perfect consists of a form of the auxiliary *va* and an infinitive. The auxiliary is partially homophonous with the present indicative of *anar* 'go', as it is historically descended from this form (Cruschina & Kocher 2022), but the lack of complete homophony precludes assuming that in contemporary Catalan the past perfect auxiliary is a form of *anar* 'go'. The first and second person plural forms of the present indicative of this verb are *anem* and *aneu*, whereas the corresponding forms of the auxiliary are *vam* and *vau* respectively. In addition, some dialects have more differences between the two paradigms, with forms such as *vares* and *varen* for the second person singular and third person plural of the auxiliary, contrasting with the corresponding forms *vas* and *van* of *anar*. For this reason, the auxiliary is glossed as VA.

The f-structure feature [TENSE PAST.PERF] triggers a rule (a lexeme assignment rule) that maps an f-structure containing that feature to two c-structure categories (word forms). One is a form of the lexeme VA, which lacks a PRED feature, in the present tense, and the other one is an infinitive of any lexeme. Both are verbs, indicated by the presence of the grammatical category V for each of the words involved in the rule in (16):

(16)



Each c-structure category referred to by this rule undergoes the rules of the morphology, which produce the appropriate word form. The morphology has access to the c- and f-structure features, except for those that are also specified in the morphological structure. This is the case of the tense feature in (16): the relevant tense feature for the form of the lexeme VA is PRES, as specified in the morphological structure, even though the f-structure contains the tense feature PAST.PERF. The other features relevant for the form of the auxiliary are present in the f-structure, specifically those of the subject. The first person singular has an irregular ending: *vaig*. The remaining forms have the expected endings: *vas* (2nd.sg.), *va* (3rd.sg.), *vam* (1st.pl), *vau* (2nd.pl.), *van* (3rd.pl.). The main verb has the categorial V-FORM feature INF, which ensures that it has the infinitival morphology.[12]

[12] I am assuming that the V-FORM feature INF is a c-structure feature, as opposed to a morphological feature. The facts illustrated in (20) show that the infinitive in the periphrastic past construction can be coordinated, suggesting that the feature "infinitive" is a syntactic feature, but cannot be an f-structure feature.

4.2 The periphrastic past: two words in a co-head relation

One might be tempted to assume that the sequence of the past perfect auxiliary and the infinitive is a single word (some kind of compound). Following is evidence that the sequence of the *va*-form and the infinitive consists of two separate words.

"Clitic" attachment. If the f-structure contains information that maps onto affixal elements of the kind known as "clitics" in the literature on Romance languages, these affixes can attach to either of the two verb forms in the construction. "Clitics" attach as suffixes to non-finite forms and imperatives and as prefixes to all other forms: this accounts for the alternative placement of "clitics" in the periphrastic past perfect construction, as in (17):

(17) a. *El va llegir*.
PRO.ACC.M.3SG VA.3SG read.INF
'S/he read it.'

b. *Va llegir-lo*.
VA.3SG read.INF-PRO.ACC.M.3SG
'S/he read it.'

(18) a. *Us vam esperar*.
PRO.2PL VA.1PL wait.INF
'We waited for you.'

b. *Vam esperar-vos*.
VA.1PL wait.INF-PRO.2PL
'We waited for you.'

The phonological alternation that we observe in the "clitic"—*el* vs. *lo* in (17) and *us* vs. *vos* in (18)—is an allomorphy conditioned by the prefixal vs. suffixal status of the "clitic": forms such as *el* and *us* are used in prefixal position and forms such as *lo* and *vos* are used when the "clitic" is suffixed following a consonant.

The fact that "clitics" can be either prefixed or suffixed to the periphrastic past perfect would be completely unexpected if this construction were a single word (a compound). But this is exactly what we expect if the two components of the past perfect are separate words. Being verbs, each one can host an affix of the "clitic" kind.

Separability of the two components. The two components of the past perfect, as independent words, can be separated by certain syntactic elements, such as the emphatic negative particle *pas*, the focus expressions *ni* and *ni tan sols* 'not even' and *fins i tot* 'even':

(19) a. *No li va pas dir que no*.
not PRO.3SG.DAT VA.3SG EMPH-NEG say.INF that no
'S/he certainly did not say no to him/her.'

b. *Vaig fins i tot recórrer a les amenaces*.
VA.1.SG even resort.INF to the threats
'I even resorted to threats.'

If a word cannot appear inside another word, the two components of the past perfect cannot be a single word. On the other hand, we expect a word to appear between two other words. Elements such as *pas*, *ni*, or *fins i tot* can also appear outside the periphrasis with no discernible difference in meaning.

Conjoinability. The infinitive in the past perfect periphrasis can be conjoined with another infinitive:

(20) a. *Van [entrar i sortir] diverses vegades*.
VA.3PL [enter.INF and exit.INF] several times
'They entered and exited several times.'

b. *Li vaig [dir i repetir] que portés el carnet.*
PRO.3SG.DAT VA.1.SG [say.INF and repeat.INF] that bring.PAST.SUBJV.3SG the ID card
'I said and repeated to him/her to bring his/her ID card.'

Coordination is a syntactic (not morphological) process. If the aux-infinitive sequence were a single word, we would not expect coordination to involve parts of a word. The auxiliary and the infinitive of the periphrasis, as separate words, can be involved in syntactic phenomena such as coordination.

The preceding evidence shows that the two components of the periphrasis are separate words. We might then assume that they constitute a verb-complement construction, in which the auxiliary would be a complement-taking verb like *prometre* 'promise', *permetre* 'allow' or *provar* 'try', and the infinitive would be the head of an infinitival complement. However, the periphrasis behaves unlike verb-complement sequences in the following respects:

a) phrases cannot appear between the auxiliary and the infinitive;
b) coordination of the infinitive cannot involve phrases;
c) the infinitive cannot be left out or pronominalized.

A verb-complement construction allows a full phrase, such as a subject, to appear separating the verb from its infinitival complement, (21a), but this is not possible with the past perfect periphrasis, (21b).

(21) a. *Primer prova el mestre d' aclarir el concepte*.
first try.PRES.3SG the teacher DE clarify.INF the concept
'First, the teacher tries to clarify the concept.'

b. **Primer va el mestre aclarir el concepte*.
first VA.3SG the teacher clarify.INF the concept
'First, the teacher clarified the concept.'

The head of an infinitival complement can be coordinated together with its own phrasal complements, (22a), but the result of coordinating the infinitive of the periphrastic past perfect construction with its phrasal complements is considerably degraded, (22b).

(22) a. *Provarà d' [aclarir el concepte] i [explicar*
try.FUT.3SG DE [clarify.INF the concept] and [explain.INF

el resultat].
the result]
'S/he will try to clarify the concept and explain the result.'

b. ?*Vaig [aclarir el concepte] i [explicar el resultat]*.
VA.1SG [clarify.INF the concept] and [explain.INF the result]
'I clarified the concept and explained the result.'

Depending on the subordinating verb, an infinitival complement can either be left out and interpreted as coreferential with an appropriate expression in the discourse, (23a), or can be pronominalized by means of a pronominal "clitic" such as *ho*, (23b). But neither of these options is available for the infinitive in the past perfect periphrasis, (23c).

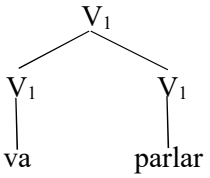
(23) a. *No vols comprar-te aquest cotxe o no pots* ∅?
not want.2SG buy.INF-2SG this car or not can.2SG
'You don't want to buy this car or you can't?'

b. *No aclariràs la idea, si no ho proves*.
not clarify.FUT.2SG the idea if not PRO.3SG try.2SG
'You will not clarify the idea, if you don't try (to do so).'

c. **Volia aclarir la idea, però no (ho) va*.
want.IPFV.3SG clarify.INF the idea but not (PRO.3SG) VA.3SG
'S/he wanted to clarify the idea, but s/he didn't (do so).'

These facts suggest that, as indicated in rule (16), the past perfect periphrasis is a V^0 consisting of two V^0 nodes each of which maps onto the same f-structure. To illustrate this idea, we can represent the c- and f-structures of the periphrasis *va parlar* 'spoke' as follows:

(24)



$$\begin{bmatrix} \text{TENSE} & \text{PST.PERF} \\ \text{PRED} & \text{'speak} < \text{ARG1}_2 > \text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \begin{bmatrix} \text{PERS} & 3 \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix}_2 \end{bmatrix}_1$$

Each verb form, as an independent word, undergoes morphological operations. One of these operations can be "clitic" attachment. Since the two words in the periphrasis map onto the same f-structure, the f-structure features that can license a particular "clitic" attached to the auxiliary can also license the same "clitic" attached to the infinitive. This accounts for the alternative position of "clitics" illustrated in (17)-(18).

Each V position can include a V-adjoined particle, whether postverbal (such as *pas*) or preverbal (such as *ni*). *Pas* is an emphatic negation particle that always follows a verb. We can assume that it adjoins to the right of a V node. As the auxiliary *va* occupies a V node, it can be followed by *pas*, as shown in (19a). Focus particles like *ni* and *fins i tot* are left-adjoined to the focused constituent, which can be a V such as the infinitive in the past perfect periphrasis, as in (19b).

The infinitive in this periphrasis is a V^0, which allows for the possibility of a coordinated V^0. A conjunction of two V^0s is a V^0: [$_V$ [$_V$ entrar] [$_C$ i] [$_V$ sortir]]. This accounts for the situation illustrated in (20), which clearly argues for the claim that the periphrasis is not a word.

The fact that phrases cannot appear between the two components of the periphrasis, as seen in (21b), follows from the c-structure of the periphrasis and the assumption that an XP cannot appear inside an X^0. The same idea accounts for the fact that a phrase cannot be part of the coordinated infinitive, as in (22b). The claim that the infinitive in the periphrasis does not head a complement of the auxiliary accounts for the observation that the infinitive cannot be left out as a null anaphora or expressed by means of a pronominal element, as in (23c).

5 Conclusions

Although lexical encapsulation, the LEH (4), is one of the design principles of the framework of LFG that most contribute to giving it its distinctive features, the evidence from languages with rich inflectional systems compellingly argues for abandoning this principle. The rules of inflectional morphology have to have access to syntactic information, information that LFG represents in its syntactic structures, namely, c- and f-structures. Attempts to account for complex inflectional morphology within LFG adopting a realizational non-morphemic approach to morphology have shown that preserving the LEH is only possible if syntactic information is copied in a large-scale fashion into the lexicon under the guise of m-features. This not only devoids the LEH of any empirical substance it may have had, but incurs a massive violation of Occam's Razor, or the simplicity criterion, as the same information is redundantly represented in different parts of the theory.

Abandoning lexical encapsulation does not entail abandoning lexicalism, the LIP (3). One can still assume that there is a fundamental difference between words and phrases, between the rules that govern the structure of words and those that govern the structure of phrases, and the units that make up words and the units that make up phrases, which is the essence of lexicalism, while rejecting the idea that the structure of words is governed by rules that are not informed by syntax. A lexicalist framework that is not constrained by lexical encapsulation has to assume that the rules of word formation operate alongside the rules of the syntax. Under this view, words are not inserted in the syntax with an array of f-structure annotations that are the syntactic counterpart of their inflectional morphology. Rather, the inflectional morphology of a word (the phonological side of this morphology) is computed on the basis of rules that are sensitive to the categorial and f-structure information available to the word. The rules of word formation require a word to have a specific phonological property (such as an affix appended to a stem) if certain syntactic and morphomic conditions are met. (By "morphomic" we refer here to properties that are not syntactic but may affect the form of words, such as inflectional classes.)

An LFG implementation of such rules may be viewed as part of the mapping of f-structure to c-structure, or f-to-c mapping. A specific set of f-structure features has a consequence on the phonology of a terminal node of the c-structure. Once we accept we have this type of rule, we can easily extend it to account for periphrasis. A periphrasis licensing rule has the effect of licensing a form of a lexeme: a given set of features (f-structure features and morphomic features) may require the presence of a form of a lexeme in the c-structure, as well as possibly requiring other features. This is what we see in the Latin periphrastic perfect and in the Catalan periphrastic past perfect. In the Latin case, the f-structure features perfective and finite coupled with the morphomic feature of the P conjugation require the word that maps onto these features to be a past participle form and to cooccur with an imperfective form of the lexeme SUM. As for the Catalan past periphrasis, the presence of the past perfect feature in an f-structure requires that the corresponding V node in the c-structure be composed of two V nodes: a form of the lexeme VA in the present tense morphology and an infinitive.

Crucially, the forms licensed by these f-to-c mapping rules undergo their own morphological rules and, in the application of these rules, an attribute specified in the morphological representation of a word takes precedence over, or blocks, the same attribute in the corresponding f-structure. For example, in the Latin periphrasis, the attribute PERFECTIVE is specified both in the f-structure and in the morphological representation of the form of SUM required in this periphrasis, with different values. The rules of the morphology only take into account the latter occurrence of the feature when spelling out the form of SUM.

To conclude, the rejection of lexical encapsulation leads us to assume that inflectional word formation rules, or rules of the morphology, are part of the mapping of f-structure to c-structure. If we allow this mapping not only to constrain the form of words, but also to require the presence of particular lexemes, then we account for periphrasis as part of the f-to-c mapping.

References

Alsina, Alex. 2020. Obligatory clitic expression, clitic omission, and the morphology-syntax interface. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG'20 Conference*, 5–25, Stanford, CA: CSLI Publications. https://typo.uni-konstanz.de/lfg-proceedings/LFGprocCSLI/LFG2020/lfg2020-alsina.pdf

Alsina, Alex. 2022. Partitivity in Romance and the syntax-morphology connection. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'22 Conference*, 1–20, Stanford, CA: CSLI Publications. https://lfg-proceedings.org/lfg/index.php/main/article/view/11/5

Alsina, Alex. In press. LFG and Romance languages. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.

Alsina, Alex & Eugenio M. Vigo. 2014. Copular inversion and non-subject agreement. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG'14 conference*, 5–25, Stanford: CSLI Publications. https://typo.uni-konstanz.de/lfg-proceedings/LFGprocCSLI/LFG2014/papers/lfg14alsinavigo.pdf

Alsina, Alex & Eugenio M. Vigo. 2017. Fixing LFG to account for direct-inverse agreement: the case of Plains Cree. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG'17 Conference*, 24–44, Stanford: CSLI Publications. https://typo.uni-konstanz.de/lfg-proceedings/LFGprocCSLI/LFG2017/lfg2017-alsina-vigo-2.pdf

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Bresnan, Joan & Sam A. Mchombo. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63, 741–782.

Bresnan, Joan & Sam A. Mchombo. 1995. The Lexical Integrity Principle. Evidence from Bantu. *Natural Language and Linguistic Theory* 13, 181–254.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-Functional Syntax*. Malden, MA: Wiley Blackwell.

Brown, Dunstan & Andrew Hippisley.2012. *Network morphology: A defaults-based theory of word structure*. Cambdrige: CambridgeUniversityPress.

Börjars, Kersti, Nigel Vincent & Carol Chapman. 1997. Paradigms, periphrases and pronominal inflection: a feature-based account. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1996*, 155–180. Dordrecht: Kluwer.

Cruschina, Silvio & Anna Kocher. 2022. A surprise in the past: The historical origins of the Catalan go-past. *Catalan Journal of Linguistics* 21: 159–186. https://revistes.uab.cat/catJL/article/view/v21-cruschina-kocher/379-pdf-en

Dalrymple, Mary. 2015. Morphology in the LFG Architecture. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG'15 Conference*, 64–83, Stanford, CA: CSLI Publications. https://typo.uni-konstanz.de/lfg-proceedings/LFGprocCSLI/LFG2015/papers/lfg15dalrymple.pdf.

Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. The interface to morphology. In Mary Dalrymple, John J. Lowe & Louise Mycock (eds.), *The Oxford reference guide to Lexical Functional Grammar*, 436–470. Oxford: Oxford University Press.

Julien, Marit. 2007. On the relation between morphology and syntax. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, pp.209–238. Oxford: Oxford University Press.

Luís, Ana & Louisa Sadler. 2003. Object clitics and marked morphology. In Claire Beyssade, Olivier Bonami, Patricia Cabredo Hofherr & Francis Corblin (eds.), *Empirical issues in syntax and semantics* 4, 133–153. Paris: Presses Universitaires de Paris-Sorbonne.

Mohanan, Tara. 1995. Wordhood and lexicality: Noun incorporation in Hindi. *Natural Language and Linguistic Theory* 13, 75–134.

O'Neill, Paul. 2016. Lexicalism, the Principle of Morphology-free Syntax and the Principle of Syntax-free Morphology. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology (Cambridge handbooks in language and linguistics)*, 237–271. Cambridge: Cambridge University Press.

Sadler, Louisa & Andrew Spencer. 2001. Syntax as an exponent of morphological features. In Geert Booij & Jap Van Marle (eds.), *Yearbook of Morphology 2000*, 71−96. Dordrecht: Springer. DOI: 10.1007/978-94-017-3724-1_4

Spencer, Andrew. 2013. *Lexical relatedness: a paradigm-based model*. Oxford: Oxford University Press

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Stump, Gregory T. 2006. Heteroclisis and paradigm linkage. *Language* 82.2, 279−322.

Stump, Gregory T. 2016. *Inflectional paradigms : content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.