

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

Alex Alsina (U. Pompeu Fabra)

This paper addresses the problem of periphrasis, specifically inflectional periphrasis, and proposes that periphrasis, together with inflectional morphology, can be adequately accounted for by means of the principles that map f-structure to c-structure. Whereas the correspondence between these two levels is standardly assumed in LFG to be unidirectional, with c-structure providing all the information for constructing f-structure, the assumption defended here is that this correspondence is bidirectional and that inflectional morphology takes f-structure information as the input for its rules, as argued in previous work. In the case of periphrasis, 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.

The central idea is that f-structure is the morphosyntactic representation that inflectional morphology needs for its word formation rules, playing the role of Anderson's (1992) MSR and Stump's (2001) morpho-syntactic properties (with the possibility of overriding; see below). Inflectional word formation does not take place in a component isolated from the syntactic levels of representation, as well as the other levels of representation, but takes place along with the various levels, including c-structure and f-structure. In this the present proposal differs from most approaches to morphology within LFG (e.g. Dalrymple et al 2019). Although there are purely morphological features, or m-features using Sadler & Spencer's (2001) term, such as inflectional classes, the use of m-features is restricted to those features that do not have a direct correspondence in f-structure, given that f-structure features are accessible to the morphology. Schematically, the relation between morphology, f-structure and c-structure, along with lexeme assignment, that is, the assignment of a particular lexeme to a particular f-structure, is shown in (1). The double-headed arrows indicate a mutually constraining relation between the various levels, as opposed to a unidirectional one.

(1) f-structure \longleftrightarrow lexeme assignment+morphology \longleftrightarrow c-structure

In this paper we will be concerned primarily with the claim that f-structure can constrain lexeme assignment (or the selection of a lexeme) and morphology. This idea is pivotal in accounting for inflectional periphrasis, as an option that arises along with inflectional morphology.

Verbal lexemes in Latin are classified into four conjugation classes, leaving aside irregular verbs. Conjugation classes are a purely morphological classification, with no direct effect on the syntax, and can be formalized as m-features. In addition, each conjugation class has two *subconjugations* (*subconj*, for short), which we may call A and P (mnemonic for active and passive). For most verbs the choice of subconj depends on whether the verb is active or passive. A small set of verbs, known as deponents, require subconj P, even though their syntax is that of active verbs. Another small set of verbs, known as semideponents, require subconj P only in the perfective aspect, using subconj A in the imperfective forms. Thus, we can say that the m-feature [SUBCONJ P] is selected (a) by verbs that lexically require it (deponents), (b) by verbs that lexically require it in combination with the syntactic feature [PERFECTIVE +] (semideponents), and (c) in passive forms (verb forms with an a-structure whose highest argument is suppressed). Elsewhere, if none of these conditions is met, the m-feature [SUBCONJ A] is selected and can thus be considered to be the default.

All verb forms in subconj A are synthetic, i.e. consisting of only one verb form; in contrast, in subconj P, only imperfective forms are synthetic, whereas perfective forms are periphrastic, consisting of two verb forms: an imperfective form of the copula *sum* 'be' and a form of the perfective participle of the verb under consideration. As argued persuasively by Börjars et al. (1997) and Sadler & Spencer (2001), the perfective forms of subconj P correspond to a cell in the inflectional paradigm of a verb, even though they are made up of two independent words at the level of c-structure, as do the imperfective forms of both subconjs and the perfective forms of subconj A. As an example, (2) shows the third person singular forms of the present indicative, an imperfective form, and of the past perfective indicative of both subconjs, for the verb *amo* 'love':

| | | |
|-------------------|-------------------------------|---------------------------------------|
| (2) | Present, imperfective | Past, perfective |
| Subconjugation A: | <i>amat</i> 's/he loves' | <i>amavit</i> 's/he loved' |
| Subconjugation P: | <i>amatur</i> 's/he is loved' | <i>amatus/-a est</i> 's/he was loved' |

The f-to-c mapping principles involve the selection of a lexeme and the application of the morphological rules to it. If the lexeme *amo* is selected in its active form, the morphological rules generate a word form with the morphology that corresponds to a first conjugation word and to subconj A. If it is selected in the passive form, the passive a-structure of the verb triggers the assignment of the m-feature [SUBCONJ P]. If the word is associated with this m-feature and the f-structure to which the word is assigned has the feature [PERF +] (for perfective aspect), a specific lexeme assignment principle is activated which makes the f-structure in question

map onto two words: the past participle of the verb associated with the feature [SUBCONJ P] and a form of *sum* in the imperfective aspect. A formalization of this principle is given in (3):

$$(3) \quad \left[\begin{array}{c} \text{PERF} \\ \text{+} \end{array} \right]_f \wedge \left[\begin{array}{cc} \text{LEXEME} & \text{X} \\ \text{SUBCONJ} & \text{P} \end{array} \right]_m \longrightarrow \left[\begin{array}{cc} \text{LEXEME} & \text{X} \\ \text{V-FORM} & \text{PAST-PART} \\ \text{FINITE} & \text{--} \end{array} \right]_m \left[\begin{array}{cc} \text{LEXEME} & \text{SUM} \\ \text{PERF} & \text{--} \end{array} \right]_m$$

The left-hand subscript on the feature structures in (3) distinguishes f-structures, with subscript *f*, from morphological structures, with subscript *m*. Coindexation of structures signals correspondence between levels: the f-structure on the left of the arrow in (3) is in a mapping or correspondence relation with the coindexed morphological structures. The two morphological structures on the right of the arrow in (3) undergo the rules of the morphology that yield the appropriate inflected forms of the two lexemes involved.

The rules of the morphology have access to the information 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 morphology accesses this information and assigns the appropriate inflections to the verb in question. In the case of an f-structure that maps onto two verb forms –a periphrasis–, such as the structure that results from the application of rule (3), the form of the copula, if finite, reflects the features of the subject. The non-finite past participle form of the main verb does not show person features. But for this form the morphology needs to access the gender and number features of the subject (as well as case), as the past participle has the morphology of an adjective and thus reflects the gender, number, and case features of the subject. This accounts for contrasts such as the following:

- (4) a. *Discipulus amatus / *amatae est / *sunt.*
 student.NOM.M.SG loved.NOM.M.SG / *loved.NOM.F.PL be.PRES.3.SG / *be.PRES.3.PL
 ‘The (male) student was loved.’
- b. *Discipulae amatae / *amatus sunt / *est.*
 student.NOM.F.PL loved.NOM.F.PL / *loved.NOM.M.SG be.PRES.3.PL / *be.PRES.3.SG
 ‘The (female) students were loved.’

Crucially, although the morphology has to have access to the f-structure features, the access to a particular feature is blocked in case that same feature is specified with a different value in the morphological structure, as a result of a principle of the f-structure-morphology mapping such as (3). Thus, for example, the form of the lexeme *sum* in the output of this rule is assigned the feature [PERF –]: this feature is incompatible with the feature [PERF +] in the corresponding f-structure and consequently overrides the latter. As a result, the *be* form of the construction is imperfective, as seen by the present forms *est* or *sunt* in (4), but the semantics corresponds to the perfective aspect, with a past tense interpretation, as required by the f-structure perfective aspect feature. A similar situation arises with finiteness. If the f-structure has the feature [FINITE +], the form of *sum* is finite, as no conflicting specification arises through the application of rule (3); but this rule does specify the feature [FINITE –] on the form of the main verb. In the application of the morphological rules, this feature overrides the [FINITE +] in the f-structure and the main verb is treated as a non-finite verb form morphologically. Thus, the morphology regularly accesses f-structure information, but this information is overridden by conflicting information introduced in the f-to-c-structure mapping.

A periphrasis-licensing rule like (3) maps an f-structure onto two verb forms. In order for this not to cause 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 verb in the construction, it is selected with its PRED feature. When it co-occurs with another verb, the option without the PRED feature is chosen. In order to exclude using a semantically empty verb, such as the auxiliary *sum*, unless required by a rule such as (3), we can appeal to Bresnan et al.’s (2016: 90) Economy of Expression. In this way we can assume there is a single lexeme *sum*, in its use both as an auxiliary and as a main verb.

The second construction to be analyzed within the approach presented here is the periphrastic past perfect in Catalan, which is composed of an auxiliary 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 and references cited there), but the lack of complete homophony precludes assuming that in contemporary Catalan the past perfect auxiliary is a form of *anar*. Interestingly, a synthetic past perfect exists, but is only used in highly formal and literary registers in most areas including Barcelona. In the standard and colloquial registers there is a gap in the morphological paradigm. However, the f-structure paradigm (the set of possible

combinations of f-structure feature relative to tense, aspect, and mood) is complete, as there is a periphrastic form that fills this morphological gap.

The f-structure feature [TENSE PAST.PERF] triggers a rule (a lexeme assignment rule) that maps an f-structure containing that feature to two 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, which is notated by the left-hand subscript “v” on the corresponding morphological feature structures in the rule in (5):

$$(5) \quad [f \text{ TENSE PAST.PERF}]_1 \longrightarrow \left[\begin{array}{cc} \text{LEXEME} & \text{VA} \\ \text{TENSE} & \text{PRES} \end{array} \right]_1 \\ \left[\begin{array}{c} \text{v} \\ \text{V-FORM} & \text{INF} \end{array} \right]_1$$

Each lexeme undergoes the rules of the morphology, which produce the appropriate word form. The morphology accesses the f-structure features, except for those that are overridden by features specified in the morphological structure. This is the case of the tense feature in (5): 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 v-form feature INF, which overrides any feature in the f-structure that is inconsistent with it.

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 (6):

- (6) a. *El va llegir.* / *Va llegir-lo.*
 3.M.SG.ACC VA.3SG read.INF / VA.3SG read.INF-3.M.SG.ACC
 ‘S/he read it.’
- b. *Us vam esperar.* / *Vam esperar-vos.*
 2.PL VA.1.PL wait.INF / VA.1.PL wait.INF-2.PL
 ‘We waited for you (pl).’

This indicates that, from a morphological point of view, the two components of the past perfect are separate words. Being verbs, each one can host an affix of the “clitic” kind. Further evidence for the claim that these two components are independent words is that they can be separated by certain syntactic elements, such as the emphatic negative particle *pas*, the negative focus expressions *ni* and *ni tan sols* ‘not even’, and that the infinitive can be a coordinate structure. The two components are restricted to appearing in the strict order Aux-Inf and do not allow phrases to appear between them and the coordination of the infinitive cannot include any phrase. This suggests that they form a V⁰ in the c-structure.

The present proposal accounts for inflectional periphrasis as well as inflectional morphology by assuming that the rules of inflection take as their input f-structure information and output fully inflected word forms, with the possibility that a given f-structure may map onto two different words, as is the case of periphrasis. This requires abandoning the traditional LFG view that word formation, or morphology, takes place independently of syntax and is in a one-way feeding relation with f- and c-structure. Although this implies a weakening of certain interpretations of lexicalism, it maintains the lexical integrity principle, as stated in Bresnan and Mchombo (1995: 181).

References

- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press. Börjars, Kersti & Vincent, Nigel & Chapman, Carol. 1997. Paradigms, periphrases and pronominal inflection: a feature-based account. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 1996*, 155–180. Dordrecht: Kluwer. Bresnan, Joan & Asudeh, Ash & Toivonen, Ida & Wechsler, Stephen. 2016. *Lexical-Functional Syntax*. Chichester, UK: Wiley-Blackwell. Bresnan, Joan & Mchombo, Sam A. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13.2: 181–254. Cruschina, Silvio & Kocher, Anna. 2022. A surprise in the past: The historical origins of the Catalan go-past. *Catalan Journal of Linguistics* 21: 159–186. Dalrymple, Mary & Lowe, John J. & Mycock, Louise. 2019. The interface to morphology. In Dalrymple, Mary et al. (eds.), *The Oxford reference guide to Lexical Functional Grammar*, 436–470. Oxford: Oxford University Press. Sadler, Louisa & Spencer, Andrew. 2001. Syntax as an exponent of morphological features. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 2000*, 71–96. Dordrecht: Springer. Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.