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

ALEX ALSINA, UNIVERSITAT POMPEU FABRA

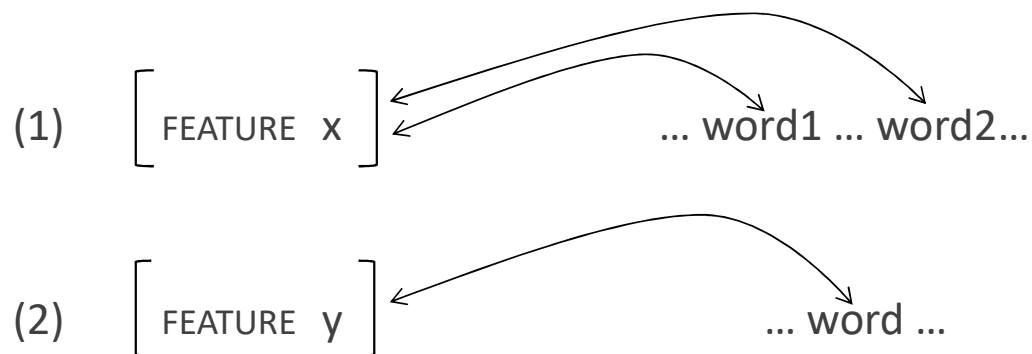# Lexicalism and the syntax-morphology relation

Two related questions:

◦ Can the syntax feed the morphology? I.e., can word formation use information present in the syntax?

◦ Can lexicalism be preserved if we allow the syntax to feed the morphology? Are we forced to abandon the lexical integrity principle if the syntax can feed the morphology?

   ◦ The lexical integrity principle (Bresnan and Mchombo 1995: 182): the morphemic structure of words differs from the c-structure of phrases both in constituents and principles of combination.

The standard answer in LFG is to enforce lexicalism by disallowing the syntax to feed the morphology: words are formed without using syntactic information.

My answer is yes to both questions: yes, the syntax feeds the morphology and yes, we can preserve lexicalism in a framework where the syntax feeds the morphology. The morphology is a module specifically dedicated to word formation that uses syntactic information in its principles.

The problem of periphrasis

Periphrasis is a notorious problem for the standard LFG conception of the relation between syntax and morphology. If a given lexeme is realized as a single word form for a set of f-structure features and is realized as two word forms for another set of f-structure features, we say that the latter is a periphrasis (or a periphrastic form). Schematically, a periphrasis can be depicted as in (1), showing the correspondence between word forms in the c-structure and features in the f-structure. By contrast, (2) shows a one-to-one correspondence between word and f-structure.

(1) $\begin{bmatrix} \text{FEATURE} & \text{x} \end{bmatrix}$ … word1 … word2…

(2) $\begin{bmatrix} \text{FEATURE} & \text{y} \end{bmatrix}$ … word …

The problem of periphrasis

If both (1) and (2) are possibilities for a given lexeme, it is a problem for the standard LFG conception of the relation between word formation and syntax.

The lexicon, or word formation component, cannot output two words, if the input is one lexeme.

Even if we allowed the lexicon to do this, it would need to access information in the f-structure, in order to know when to map a lexeme onto one word form and when to map it onto two word forms.

The standard assumption in LFG is that the features in the f-structure are licensed or contributed by the word forms that appear in the c-structure (as well as by rules): the c-structure provides all the information needed for constructing the f-structure. The f-structure features are not available for the application of word formation principles.

Main claim of the paper

The main claim of the paper is that inflectional periphrasis is a special case of inflectional morphology.

This builds on the idea that the correspondence between c-structure and f-structure is bidirectional and that inflectional morphology takes f-structure information as the input for its rules. See previous work: Alsina (2020; 2022).

Inflectional morphology is part of the mapping between f-structure and c-structure.

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.

Outline of the paper

Central ideas

Proponents of inferential realizational approaches to morphology posit a morphosyntactic representation that inflectional morphology needs for its word formation rules (Anderson's (1992) MSR, Stump's (2001) morphosyntactic properties, Darlymple et al.'s (2019) m-features).

I propose that the f-structure **is** this morphosyntactic representation. We don't need a different representation in addition to the f-structure.

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.

Schematically, the relation between morphology, f-structure and c-structure, along with lexeme assignment (or lexeme choice):

(3)   f-structure ⟷ lexeme assignment+morphology ⟷ c-structure

The Latin synthetic-periphrastic alternation

A problem studied by Börjars et al. (1997) and Sadler & Spencer (2001).

The Latin conjugation has a mix of synthetic and periphrastic forms. Many verbs have an "active" and a "passive" conjugation. All verb forms in the active conjugation are synthetic, but, in the passive conjugation, imperfective forms are synthetic and perfective forms are periphrastic:
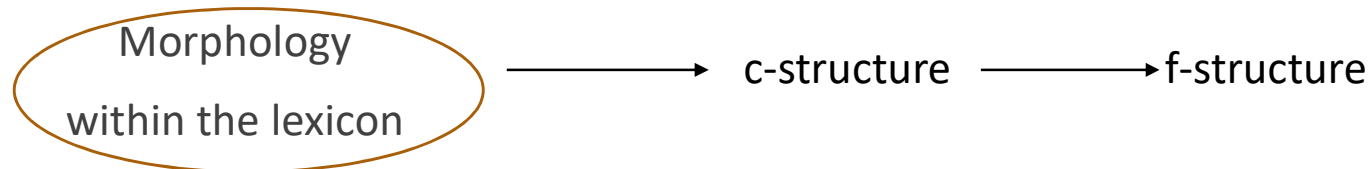
(4)	Present, imperfective	Past, perfective
Active:	*amat* 's/he loves'	*amavit* 's/he loved'
Passive:	*amatur* 's/he is loved'	*amatus/-a est* 's/he was loved'

The periphrastic form *amatus est* is functionally equivalent to the synthetic *amatur*, except that the former is perfective and the latter is imperfective. 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*.

The Latin synthetic-periphrastic alternation

There is a consensus that the periphrastic forms cannot be handled by the principles that combine words into phrases in the standard LFG conception of the relationship between morphology and syntax and, within syntax, between c-structure and f-structure:

(5)   Standard LFG view of the relation between modules and levels:

Morphology

within the lexicon   ⟶   c-structure   ⟶   f-structure

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?

The two Latin subconjugations

Verbal lexemes in Latin are classified into four conjugation classes, a purely morphological classification, with no direct effect on the syntax, formalized as m-features, morphological features with no direct correlate at f-structure.

In addition, each conjugation class has two *subconjugations* (*subconj*, for short): subconj A and subconj P (mnemonic for active and passive).

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, the default m-feature [SUBCONJ A] is selected.

The m-feature [SUBCONJ P] can be part of the information of a lexeme, with deponents and semideponents, or can be assigned by rule, in passive forms. *Loquor* 'speak' and *amor* 'be loved' are both [SUBCONJ P]: in the former, because of a lexical requirement; in the latter, by a rule.
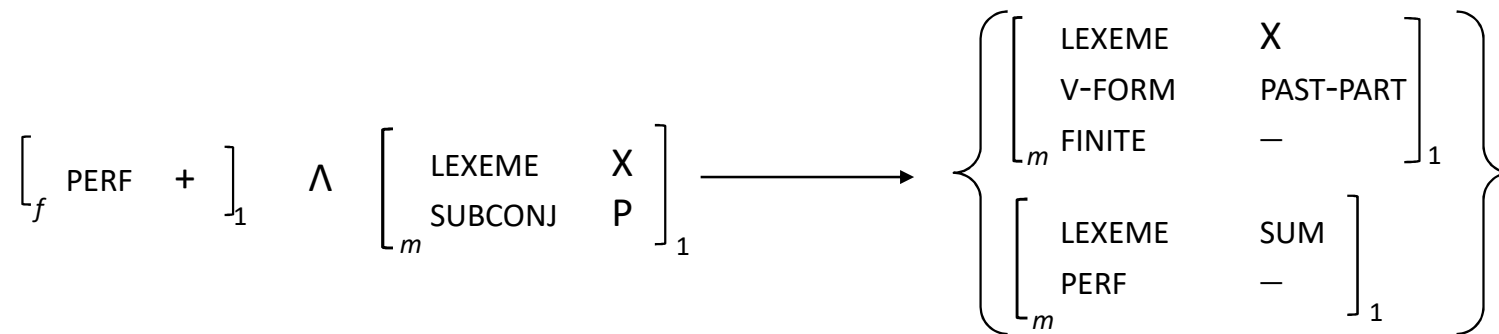
F-to-c-structure mapping

All verb forms in subconj A are synthetic, i.e. consisting of only one verb form. 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.

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 of a first conjugation word and subconj A. If it is a passive form, the passive a-structure triggers the assignment of the m-feature [SUBCONJ P].

If the word is associated with the m-feature [SUBCONJ P] and with an f-structure specified as [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 (6):

Periphrasis licensing rule

(6)

$$\begin{bmatrix} \text{PERF} & + \end{bmatrix}_{f1} \wedge \begin{bmatrix} \text{LEXEME} & \text{X} \\ \text{SUBCONJ} & \text{P} \end{bmatrix}_{m1} \longrightarrow \left\{ \begin{bmatrix} \text{LEXEME} & \text{X} \\ \text{V-FORM} & \text{PAST-PART} \\ \text{FINITE} & - \end{bmatrix}_{m1} \begin{bmatrix} \text{LEXEME} & \text{SUM} \\ \text{PERF} & - \end{bmatrix}_{m1} \right\}$$

The left-hand subscript on the feature structures in (6) 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 (6) is in a mapping or correspondence relation with the coindexed morphological structures (or words). The two morphological structures on the right of the arrow in (6) undergo the rules of the morphology that yield the appropriate inflected forms of the two lexemes involved.

Simultaneous access to different levels

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 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 (6), 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, which is an adjective, reflects the gender, number, and case features of the subject. This accounts for contrasts such as the following:
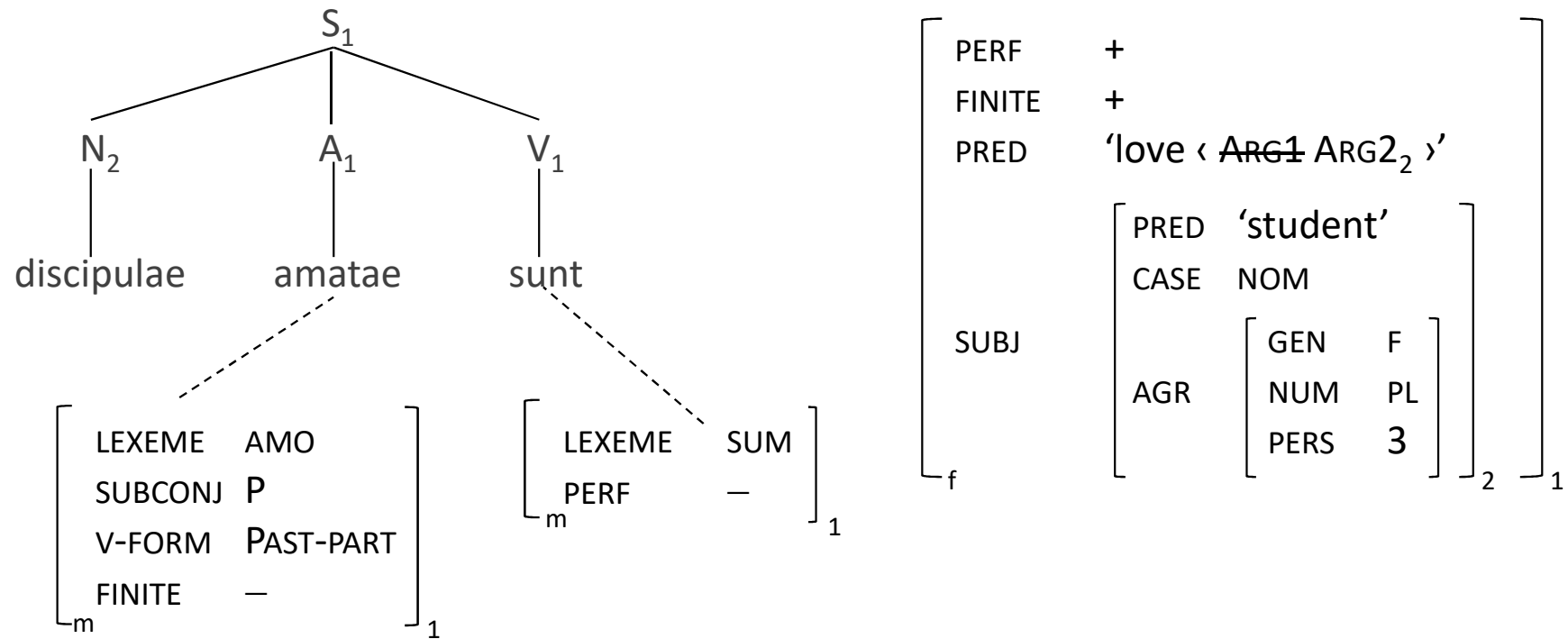
Conflicting specifications

(7) 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 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 (6) 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.

Sunt in (7b) is morphologically imperfective, but the corresponding f-structure, and semantics, is perfective.

Correspondence among levels

S_1

N_2 A_1 V_1

discipulae   amatae   sunt

$$\begin{bmatrix} \text{LEXEME} & \text{AMO} \\ \text{SUBCONJ} & \text{P} \\ \text{V-FORM} & \text{PAST-PART} \\ \text{FINITE} & - \end{bmatrix}_{m} {}_{1}$$

$$\begin{bmatrix} \text{LEXEME} & \text{SUM} \\ \text{PERF} & - \end{bmatrix}_{m} {}_{1}$$

$$\begin{bmatrix} \text{PERF} & + \\ \text{FINITE} & + \\ \text{PRED} & \text{'love} \langle \text{ A\textsc{rg}1 A\textsc{rg}2}_2 \rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'student'} \\ \text{CASE} & \text{NOM} \\ \text{AGR} & \begin{bmatrix} \text{GEN} & \text{F} \\ \text{NUM} & \text{PL} \\ \text{PERS} & 3 \end{bmatrix} \end{bmatrix}_2 \end{bmatrix}_{f} {}_{1}$$

A single lexeme for aux. and main verb

A periphrasis-licensing rule like (6) 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 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*, 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.

Periphrastic past perfect in Catalan

The periphrastic past perfect in Catalan is composed of an auxiliary and an infinitive: see (8). 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* (9).
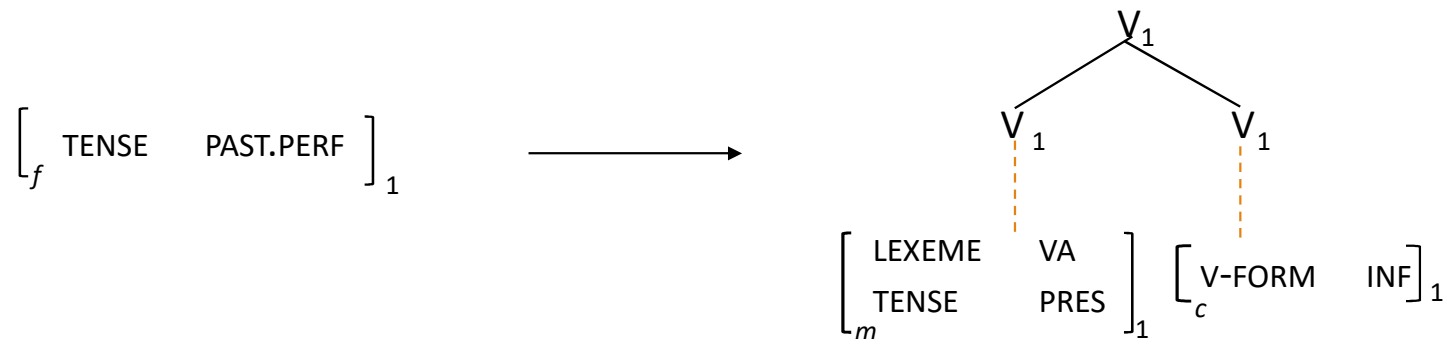
(8) a. *Va   parlar   l'   advocat.*
    VA.3SG speak.INF the   lawyer
    'The lawyer spoke.'

b. *Vam   /*anem   comprar les entrades.*
    VA.1PL /*VA.1PL buy.INF   the tickets
    'We bought the tickets.'

(9) a. *Va   endavant.*
    go.3SG   forward
    'It's going forward.'

b. *Anem /*vam   al   teatre.*
    go.1PL /*go.1PL to.the theater
    'We are going to the theater.'

Two different lexemes for the past perfect auxiliary and for the present of 'go'.

Past perfect periphrasis rule

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 (10):

(10)

$$
\left[\begin{array}{cc} \text{TENSE} & \text{PAST.PERF} \end{array} \right]_{f_1} \longrightarrow
\begin{array}{c}
V_1 \\
\diagup \quad \diagdown \\
V_1 \qquad V_1
\end{array}
$$

$$
\left[\begin{array}{cc} \text{LEXEME} & \text{VA} \\ \text{TENSE} & \text{PRES} \end{array} \right]_{m_1}
\left[\begin{array}{cc} \text{V-FORM} & \text{INF} \end{array} \right]_{c_1}
$$

One f-structure, two words

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 (10): 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, overriding any feature in the f-structure that is inconsistent with it.

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.

The position of "clitics"

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 (11):

(11) 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).'

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 are independent words: they 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':

(12) a. *No   li   va   pas   dir   que no.*

   not   3.SG.DAT   VA.3SG   EMPH-NEG say.INF that not

   '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. These elements 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:

(13) 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.*
      3.SG.DAT VA.1.SG [say.INF and repeat.INF] that bring.PAST.SUBJV.3SG the ID card
      'I told him/her over and over 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.

C-structure constraints on the periphrasis

The two components are separate words, but behave unlike Verb-complement sequences.

1. There is a strict linear precedence: the aux must precede the infinitive: (14a).

2. Phrases cannot appear between them: (14b).

3. Coordination of the infinitive cannot involve phrases: (14c).

4. The infinitive cannot be left out: (14d).

(14) a. (*Volien   parlar   i)   *parlar   van.*
        want.IMPF.3PL speak.INF and   speak.INF VA.3PL
        'They wanted to speak and they spoke.'

C-structure constraints on the periphrasis

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

c.??Va   [pelar   les   pomes] i   [tallar   les   cebes].
   VA.3SG peal-INF the   apples   and   cut-INF the   onions
   'S/he pealed the apples and cut the onions.'

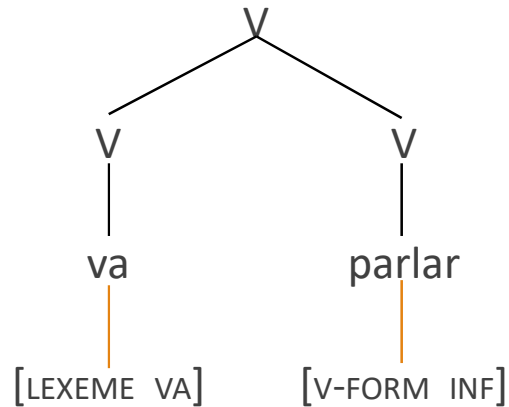d. (Deies   que   parlaries.)   *I tant si   vas!
   say.IMPF.2SG that   speak.COND.2SG.   Indeed   VA.2SG
   'You said you would speak. And indeed you did'

These facts suggest that the past perfect periphrasis is a V^0 and each of the component verbs is also a V^0.

The *va*-INF periphrasis is a V^0

The c-structure of the periphrasis (example: *va parlar* 'spoke'):


```
            V
           / \
          V   V
          |   |
         va  parlar
          |   |
[LEXEME  VA] [V-FORM  INF]
```
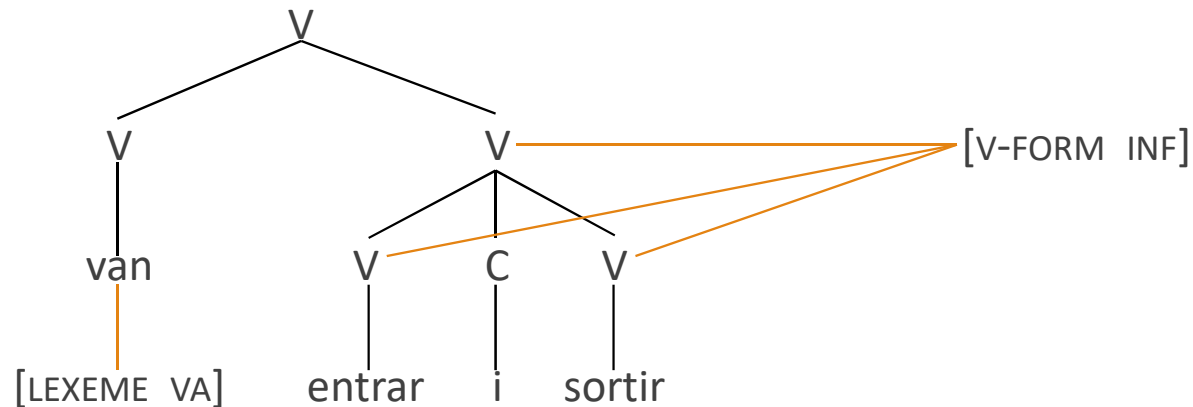
Each verb form, as an independent word, can undergo further morphological operations (e.g., "clitic" attachment).

Each V position can include a V-adjoined particle: postverbal (*pas*) or preverbal (*ni*).

The *va*-INF periphrasis is a V^0

The sister of the *va*-form is a V^0, allowing for the possibility of a coordinated V^0:

V
├── V
│   └── van
│   └── [LEXEME VA]
└── V ────────── [V-FORM INF]
    ├── V
    │   └── entrar
    ├── C
    │   └── i
    └── V
        └── sortir

But phrases cannot appear inside a V^0, accounting for why a phrase cannot separate the *va*-form and its sister and a phrase cannot be part of the coordinated infinitive.

Concluding remarks

Inflectional morphology is accounted for by assuming that the rules of inflection take as their input f-structure information and output fully inflected word forms in the c-structure.

Inflectional periphrasis is just a special case of this f-to-c mapping required by inflectional morphology. The difference is that, in inflectional periphrasis, a given f-structure maps onto two different words.

This view 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).

Thank you!

References

Alsina, Alex. 2020. Obligatory clitic expression, clitic omission, and the morphology-syntax interface. In Butt, Miriam & Toivonen, Ida (eds.), *Proceedings of the LFG'20 Conference*, 5–25. CSLI Publications, http://csli-publications.stanford.edu/

Alsina, Alex. 2022. Partitivity in Romance and the syntax-morphology connection. In Butt, Miriam & Findlay, Jamie Y. & Toivonen, Ida (eds.), *Proceedings of the LFG'22 Conference*, 1–20. CSLI Publications, http://csli-publications.stanford.edu/

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

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.

References

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.

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.