

MORPHOLOGICAL AND
SYNTACTIC WELL-FORMEDNESS:
THE CASE OF EUROPEAN PORTUGUESE PROCLITICS

Ana R. Luís and Ryo Otaguro
(University of Coimbra/University of Essex)

Proceedings of the LFG05 Conference
University of Bergen
Miriam Butt and Tracy Holloway King (Editors)
2005
CSLI Publications

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

Abstract

European Portuguese proclitics illustrate a mismatch between inflectional status and syntactic separability which is challenging to lexicalist theories of syntax. On the one hand, they form morphologically complex clitic clusters and realise verbal properties; on the other, they may be separated from the verb by lexical items, showing no sign of being morphologically attached to it. The question then is how to account for the partly inflectional and partly phrasal behaviour of proclitic affixes in a theory of syntax that prohibits elements smaller than words from being syntactically visible. In defence of the principle of Lexical Integrity (Bresnan 2001:92), Luís&Sadler (2003) take the view that proclitic affixes may not be assigned a c-structure position. In this paper, we also endorse the view that morphology and phrase structure constitute separate levels of analysis, but explore an alternative analysis.

1. Introduction

The problem posed by pronominal proclitics in European Portuguese (EP) arises from the fact that they exhibit both inflectional and syntactic properties¹. On the one hand, they form morphologically complex clitic clusters and realise verbal properties (exactly like their enclitic counterparts); but on the other, they may be separated from the verb by lexical items, showing no sign of being morphologically attached to it. These features suggest to Luís (2004) that proclitic affixes in EP should be analysed as phrasal affixes (i.e., verbal affixes with phrasal status). However, at the level of c-structure, it is not entirely clear how phrasal affixes can be accommodated in a theory that assumes lexical integrity.

The same problem has been addressed in Luís&Sadler (2003), within LFG, who argue that proclitic affixes may not be assigned a c-structure position on the grounds that such an analysis constitutes a violation of the Principle of Lexical Integrity (Bresnan 2001:92). Luís&Sadler (2003) sketch a proposal in which the proclitic affix is represented as pronominal f-structure information associated with a phrasal V-VP node. The affix itself however does not appear in the c-structure. Because of the somewhat unconventional model of c-structure adopted in that analysis, this paper aims to explore an alternative approach. We formulate a mapping between morphology and c-structure which assigns a c-structure

¹ We are grateful to Miriam Butt, Mary Dalrymple, Ron Kaplan, Tracy H. King, Gergana Popova, Louisa Sadler and Andrew Spencer for helpful comments in the early stages of this work and throughout. Remaining errors are solely our own.

position to proclitic affixes without making the assumption that incomplete morphological strings may be represented in the syntax (Luís&Otoguro 2004, to appear).

Section 2 surveys the basic facts about the EP data. Section 3, argues that mismatch phenomena in LFG pose problems to the principle of Lexical Integrity and that enough supporting evidence has been provided in the literature to justify the search for an alternative morphology/c-structure mapping. Section 4 presents the Morphological Token analysis which assumes a revised interface between morphology and syntax. Section 5 provides a short conclusion and outlines avenues for further research.

2. Overview of the data

In this section, we survey morphological and syntactic evidence in support of the claim that proclitics in EP constitute phrasal affixes. We show that proclitics are formally and semantically exactly identical to enclitics. However, while enclitics behave like genuine verbal suffixes, proclitics display phrasal properties.

2.1 Inflectional properties

It is well-known that European Portuguese, like other Romance languages, has two types of pronominal clitics. Depending on whether clitics precede or follow the verbal host, they may be enclitic to the verb, as in (1a), or proclitic, as in (1b).

Luís (2004) shows that enclitics display a significant number of affix properties such as fusion (1a), syncretism (3a), and cluster-internal allomorphy (5a), in addition to rigid ordering and idiosyncratic co-occurrence restrictions. Proclitics show exactly the same range of cluster-internal allomorphy and rigid ordering, as the examples in (1b), (3b) and (5b) illustrate.

Illustrating these properties in more detail, portmanteau forms appear when 3rd person accusative clitics follow either 1st/2nd person singular or 3rd person plural dative clitics, as in (1). A partial inventory of opaque clitic clusters is given in (2).

- (1) a. disse-**mo** (*me-o)
said-DAT.1SG-ACC.3SG.M
's/he said it to me'
- b. ... que **mo** disse (*me-o)
... that DAT.2PL-ACC.3SG.M-said
'...that s/he said it to me'

(2)

	3sg.masc.acc	3sg.fem.acc	3pl.masc.acc	3pl.fem.acc
1sg.dat	mo (= me+o)	ma (= me+a)	mos (= me+os)	mas (= me+as)
2sg.dat	to (= te+o)	ta (= te+a)	tos (= te+os)	tas (= te+as)

Table 1

When 3rd person dative clitics co-occur with 3rd person accusative clitics, the plural features on the dative forms are neutralised giving rise to syncretism, as shown in (3), where *lho* can either mean ‘V it to him’ or ‘V it to them’. The complete set of syncretic forms is provided in (4).

(3) a. deu-**lho** (*lhe-o)

gave-DAT.3SG/PL-ACC.3SG.M

‘s/he gave it to him/them’

b. ... que **lho** deu (*lhe-o)

... that DAT.3SG/PL-ACC.3SG.M-gave

‘...that s/he gave it to him/them’

(4)

	3Acc.Masc.Sg	3Acc.Fem.Sg	3Acc.Masc.Pl	3Acc.Fem.Pl
3Dat.Sg	lho	lha	lhos	lhas
3Dat.Pl				

Table 2

Cluster internally, object pronouns generally exhibit phonological alternation when 3rd accusative pronouns (*o*, *a*, *os*, *as* ‘him, her, them.masc, them.fem’) are preceded by a 1st/2nd person plural dative pronoun. The dative clitics loses its final consonant and an 3rd person accusative allomorph surfaces (i.e., *lo*, *la*, *los*, *las*).

(5) a. deu-**no-lo** (*nos-o)

gave-DAT.2PL-ACC.3SG

‘s/he gave it to us’

b. ... que **no-lo** disse

... that DAT.2.PL-ACC.3SG.M-said

‘...that s/he said it to us’

The complete inventory of clusters combining 1st/2nd person plural datives with 3rd person accusatives is shown in (6).

(6)

	3sg.m.acc	3sg.f.acc	3pl.m.acc	3pl.f.acc
1pl.dat	(nos+o)→ no-lo	(nos+a)→ no-la	(nos+os)→ no-los	(nos+as)→ no-las
2pl.dat	(vos+o)→ vo-lo	(vos+a)→ vo-la	(vos+os)→ vo-los	(vos+as)→ vo-las

Table 3

The morphophonological changes taking place inside the cluster suggest that a morphological analysis of EP pronominals should be preferred. To capture the fact that enclitics and proclitics are formally and semantically exactly identical, Luís (2004) develops an inflectional analysis within a revised version of Paradigm Function Morphology (Stump 2001) which generates enclitics and proclitics through one and the same realisation rule (e.g., $R \{ \text{Case:Dat, Nmb:Sg, P:3} \}_{\text{def}} \langle lhe \rangle$). Such realisation rule R defines affixes as ‘ambifixal’ exponents, that is as affixes which may either attach as prefixes or as suffixes (cf. Stump 1993 on Fula). A morphological alignment function is provided which places the clitic to the left or to the right of the host.

2.2. Enclitic suffixes

Shape variations found at the boundary between verbs and enclitics suggest that enclitics constitute verbal suffixes. For example, pronominal allomorphy is found when 3rd person accusative pronouns, i.e. *-a*, *-o*, *-os*, *-as*, are preceded by verbs ending in *-r*, *-s* or *-z* (7a) or by 3rd person plural verb forms (7b). In the first context, accusative clitics surface in their *l*-form, as *-lo*, *-la*, *-los*, *-las*; in the second context they appear in their *n*-form, as *-no*, *-na*, *-nos*, *-nas*.

- (7) a. Levamo **-la** (not: levamos-a)
take -acc.3.sg.fem
‘We will take her’
- b. Os meninos levam **-nos** (not: *levam-os)
the boys take -acc.1.pl
‘The boys take us’

Enclitics also trigger phonological changes on the verb. In particular they induce word-final consonant deletion in the two following contexts: a) when *l*-initial 3rd person accusative clitics are preceded by verb forms ending in *-s*, *-z* or *-r* (7a), and b) when 1st/2nd person plural clitics, i.e. *-nos* and *-vos*, follow 1st person plural verb forms (8).

- (11) a. ... acho que ela **o** ainda não disse.
 ... think that she ACC.3SG.MASC yet not told
 ‘... I think that s/he hasn’t told it to him/her/them yet’
- b. ... embora eu saiba que **a** já tens
 ... although I know that ACC.3SG.FEM already have
 em grande dose.
 in big portion
 ‘... although I know that you already have tons of it (= patience)’

What the data shows is that the difference between enclitics and proclitics is not just a question of right/left linearisation to the host. Based on the above evidence, Luís (2004) accounts for the asymmetry between enclitics and proclitics by analysing enclitics as verbal suffixes and proclitics as phrasal affixes. This proposal elaborates on the well-known distinction between word-level affixation and phrasal-affixation, formulated originally by Klavans (1985) and developed more recently by Anderson (1992), Legendre (2000), Spencer (2000), Spencer & Luís (to appear).

In section 2.1 we alluded to the fact that enclitics and proclitics should be derived through an inflectional realisation rule *R* (cf. *lhe*, in cf. 2.1). In addition, the fact that enclitics and proclitics constitute the same exponent is accounted for by deriving both through the same realisation rule and by formulating an alignment function which positions the clitic affix either to the left or to the right of the host. We have now seen that the difference between enclitics and proclitics is not merely positional: it is not just enough to determine the direction of attachment of the clitic affix but it is also necessary to define the nature of the host the clitic affixes attaches to. Hence, in Luís (2004), the alignment function is formulated so as to allow clitics to attach to the right edge of a verbal stem (for enclitics) and to the left of a phrasal node (for proclitics). The asymmetric placement accounts for the difference in status between stem-level suffixation and phrasal affixation.

Summarising: from the point of view of morphology, EP pronominal affixes are constructed within the morphology using a realisational architecture of Paradigm Function Morphology. The assumption is that proclitic affixes are assigned the ability to select their host in the syntax. The question we will address in the following sections is how to capture the phrasal status of proclitics at the level of c-structure.

3. Lexicalism and c-structure

Even though enclitics and proclitics contribute the same f-structure information to LFG c-structure (i.e., OBJ/OBJ2), it is not clear how to incorporate phrasal affixes into a lexicalist model of syntax. The essence of the problem may be summarised as follows: on the one hand, an approach that places the proclitic

affix and its immediately adjacent host under the same terminal node is theoretically in line with lexicalist assumptions but lacks empirical support; on the other hand, an approach that assigns phrasal status to proclitic affixes, at the c-structure level, is empirically correct but in violation with lexicalist assumptions.

3.1 Lexical Integrity

LFG treats morphology and syntax as independent levels of linguistic structure. A strong division is assumed between word-internal structures, on the one hand, and structures between words, on the other, with the underlying conviction that word-formation cannot take place in the syntax.

In a lexicalist theory of grammar the role of morphology is to process morphological operations (e.g. combining a root and affixes, changing stem forms and so forth) and to create fully inflected words. In LFG, those morphological operations are completely separated from syntactic ones, as defined in the principle of Lexical Integrity:

- (12) “morphologically complete words are leaves of the c-structure tree and each leaf corresponds to one and only one c-structure node” (Bresnan 2001:92).

Hence, at the level of c-structure a terminal node can only be instantiated by a single and morphologically complete word.

The only way of adjusting pronominal proclitics to this assumption would be to analyse them either as a) verbal prefixes or as b) fully-fledged words. As prefixes they would attach to the verb and surface as part of an inflected word; as words, they would themselves constitute their own c-structure node. The problem, however, is the lack of empirical evidence supporting these analyses.

There is no data suggesting that proclitics are morphologically attached prefixes, simply because proclitics do not select the category of the word they are adjacent to. In this respect, the representation in (13) would be correct for pronominal enclitic in EP (or for enclitics and proclitics in Italian, Monachesi 1999), but not for EP proclitics:

- (13)
- $$\begin{array}{c}
 \text{VP} \\
 | \\
 \uparrow = \downarrow \\
 \text{V} \\
 | \\
 \text{vêem-nos} \\
 (\uparrow \text{OBJ PRED}) = \text{PRO} \\
 \text{'they see us'}
 \end{array}$$

Luís (2004) also makes a strong case against analysing proclitics as words, more precisely as non-projecting X^0 units. Empirically, the strong resemblance between enclitic clusters and proclitic clusters (cf. section 2) can only be insightfully captured if these sequences are effectively generated through the same inflectional mechanisms. Differentiating between clusters that are proclitic and clusters that are enclitic entails the assumption that *lho* or *se-lhe* would be analysed as sequences of affixes in enclitic position but as lexical units in proclitic position, even though they are formally, semantically and morphotactically exactly the same. In addition, if we did differentiate between lexical clusters and inflectional clusters, other problematic questions would arise about proclitic clusters, in particular: a) would the internal structure of *se-lhe* be analysed as a sequence of two function words or as an opaque unit? If proclitic clusters are regarded as sequences of function words, then how would the many co-occurrence restrictions and morphophonological idiosyncrasies be accounted for? Likewise, if proclitic clusters are treated as an opaque forms, how could one explain that the clitic *se* can co-occurs productively (and agglutinatively) with many other clitic forms, as in *se-me*, *se-lhes*, *se-nos*, etc.

Supposing that there are technical answers to all these questions, one would still need to explain, as alluded to above, why the mechanisms for the derivation of proclitic clusters must be different from those applied in the derivation of enclitic clusters, considering that clusters in either position are formally and semantically exactly identical.

These and other questions suggest to Luís (2004) that the treatment of proclitics as function words – even though technically possible – is not tenable and that clusters should be uniformly analysed as complex inflectional exponents. It would also be unsound to rule out the theoretical status of phrasal affixation solely on the grounds that it challenges Lexical Integrity. Instead, it would seem to be more correct to explore ways of solving the problem of phrasal affixation without violating the integrity of words (cf. section 4 for proposal).

3.2 Morphology-syntax mismatches in LFG

In this section, we briefly survey the analysis developed by Wescoat (2002) for the treatment of morphology-syntax mismatches. Wescoat (2002) provides evidence to support the claim that well-formed morphological words do not always correspond to one and only one terminal node. English non-syllabic auxiliaries are among the phenomena examined by Wescoat.

The claim that non-syllabic auxiliary forms are morphologically attached to the (subject) pronoun was originally formulated by Spencer (1992). Luís (1997) provides empirical evidence which shows that the auxiliary-pronoun combination does effectively behave phonologically, morphologically and syntactically like one single word. Adopting Zwicky & Pullum's criteria for affixation (Zwicky & Pullum 1983), Luís (1997) points out, among other aspects, that word-internal

phonological rules, such as vowel laxing, apply to the non-syllabic auxiliary, reducing a bimoraic unit into a monomoraic one.

(14) (Luís 1997)

He'll { /hi:l/ → /hɪl/ } go

We'll { /wi:l/ → /wɪl/ } go

You've { /ju:v/ → /juv/ } been watching tv.

Luís also shows that non-syllabic auxiliary forms trigger non-productive allomorphy on the pronominal host, as illustrated in (15).

(15) (Luís 1997)

you /yu:/ but *you're* /jɔ:/

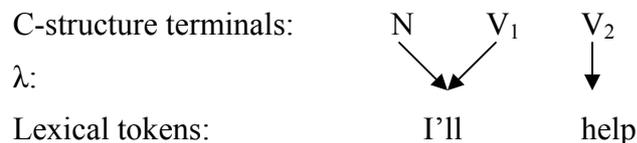
we /wi:/ but *we're* /wɜ:/ (i.e. same as 'were')

they /ðei/ but *they're* /ðɛ:/ (i.e. same as 'there')

(Other affix properties include narrow scope and high degree of selectivity).

The problem with inflected pronouns such I'll [aɪl], as Wescoat (2002) observes, is that they appear to be composed to two syntactically accessible parts. To capture this insight, 'lexical-sharing trees' are proposed which allow two or more 'terminal' nodes to share the same morphological object. The 'lexical sharing' relation is schematically illustrated below:

(16) (Wescoat 2002, p.5)



The mapping developed by Wescoat presupposes a new model of phrase-structure tree in which the Single Root Condition is not obligatory (Partee et al 1993:437-44). The analysis, thus, shows that a more complex approach to the interface between morphology and phrase-structure is necessary, one in which the well-formedness of phrase structure may have to be sacrificed. The question we would like to address now is whether the analysis may be adopted for EP phrasal affixation.

The principle of 'homomorphic lexical integrity, which requires shared nodes to be immediately adjacent, rules out any attempt at applying the analysis to EP phrasal affixes. As alluded to before, proclitic affixes do not attach morphologically to the verb and need not be adjacent to it. What this effectively means is that the proclitic-verb combination does not constitute a single word form. However, it is defined by the morphology as a well-formed inflectional

string for the following reasons: a) the sequence corresponds to a cell in the inflectional paradigm of the lexeme VER ‘see’ (Luís 2004) and b) the clitic affix realises features associated with that lexeme.

In what follows, we will try to develop an analysis which shares with Wescoat (2002) the claim that there is enough supporting evidence in favour of a revised view of the relationship between words and phrase structure.

4. Proposal

In this section, we present the Morphological Token analysis. This analysis, which is broadly outlined in Luís&Otoguro (2004) and in Luís&Otoguro (to appear), assumes that morphological well-formedness and integrity are defined solely in the morphology, through morphology-internal principles, and that morphological strings cannot be inserted directly into c-structure. Additional structure mediates between the level of morphology and the level of c-structure.

4.1 Analysis

At the interface between morphology and c-structure, we put morphological tokens in correspondence with syntactic atoms.

- (17) a. *Morphological token*: each morphological token corresponds to a well-formed stem-affix string that are defined by morphology-internal principles.
- b. *Syntactic atom*: syntactic atoms are leaves on c-structure trees; each leaf corresponds to one and only one terminal node; the insertion of syntactic atoms into c-structure is subject to standard phrase structure constraints, such as linearisation, immediate dominance, and instantiation.

The mapping between morphological tokens and syntactic atoms, as shown in (18), takes as input morphological tokens and delivers labelled syntactic atoms. In the labelling function given below, the variables y and z stand for the affixes and H represents the verbal base:

$$(18) \quad [x-H-y] \Rightarrow x_{CL} H-y_I$$

These minor alterations help us formulate the core idea of our analysis: we prohibit morphological strings from being inserted directly into phrase structure and define the ‘integrity’ of words as a condition over morphological tokens. Under this view, complete morphological strings constitute morphological tokens

which are mapped onto c-structure. Only complete strings will be properly mapped. To make our proposal clearer, we will show how the analysis works.

a) Morphological well-formedness

Within Generalised Paradigm Function Morphology (GPFM) (Luís&Spencer 2005, Spencer ms.), the well-formedness of each stem-affix string is determined as follows: the Paradigm Function PF takes the pair $\langle VER, \sigma \rangle$ (i.e., the lexeme VER and a set of morphosyntactic features σ associated with the lexeme) and delivers two complete stem-affix combinations: $v\hat{e} \langle me$ and $me \langle v\hat{e}$. Each inflectional string is the well-formed realisation of a pair $\langle VER, \sigma \rangle$.

(19) PF analysis ($v\hat{e}$ - me / me $v\hat{e}$ ‘sees me’)

- a. $PF(VER, \sigma) =_{\text{def}}$
 - i. $S(VER, \sigma) = v\hat{e}$
 - ii. $R... = me$
 - iii. $L = v\hat{e} \langle me$

- b. $PF(VER, \sigma) =_{\text{def}}$
 - i. $S(VER, \sigma) = v\hat{e}$
 - ii. $R... = me$
 - iii. $L = me \langle v\hat{e}$

Clarifying in more detail the Paradigm Function PF in (19), we note that the PF defines a) the selection of the stem S , b) the realisation of the affix R and c) the linearisation of the affix with respect to the stem L . Both PFs yield the same stem $v\hat{e}$ and the same exponent me . Only the linearisation differs: the affix follows the stem in (19a) and precedes it in (19b) (see Luís&Otoguro 2004 for an analysis of the morphosyntactic contexts triggering preverbal positioning).

Adopting Generalised Paradigm Function Morphology (Luís&Spencer 2005, Spencer ms), our morphological analysis factors out the realisation of affixes from their linearization, allowing us to capture the idea that the same affix may be subject to different linearization constraints.

Finally, the PF delivers the complete morphological strings $me \langle v\hat{e}$ and $v\hat{e} \langle me$ which constitute two distinct morphological tokens.

b) At the morphology/c-structure interface

The correspondence between morphological tokens and c-structure nodes is mediated through the algorithm in (20) which takes as input morphological

tokens and delivers labelled syntactic atoms that are inserted into c-structure as instantiations of terminal nodes².

The algorithm may be formalised as in (20), where y and z are the affixes and H represents the verbal base. The morphological token is represented in square brackets, on left side of the arrow. The syntactic atoms, which appear on the right side of the arrow³.

$$(20) \quad [x-H-y] \Rightarrow x_{CL} H-y_I$$

In (21), the mapping function has been applied to the morphological tokens derived in (19).

$$(21) \quad \begin{array}{l} \text{a. } [me, v\hat{e}] \Rightarrow me_{CL} v\hat{e}_I \\ \text{b. } [v\hat{e}, me] \Rightarrow v\hat{e}-me_I \end{array}$$

In (19a), a single morphological token corresponds to two syntactic atoms, me_{CL} $v\hat{e}_I$. This mismatch, we claim, is what separates phrasal affixation from simple affixation at the level of c-structure. In most cases, a single morphological token corresponds to a single syntactic atom, thus in (19b) no mismatch is found and the correspondence is one-to-one. In other words, in simple affixation, one stem-affix string will be inserted under one single terminal.

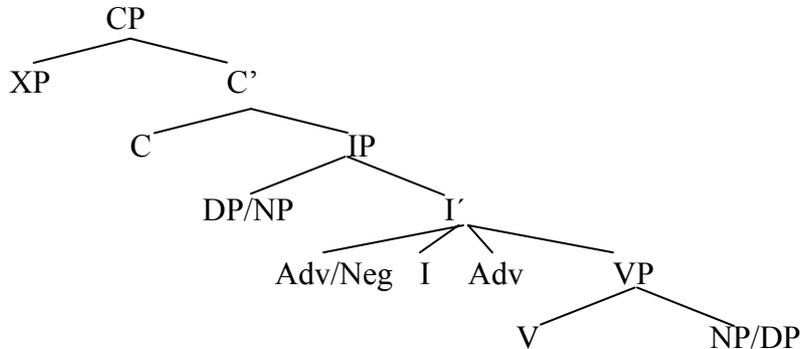
c) The c-structure

The insertion of syntactic atoms into phrase-structure is regulated by standard phrase structure principles (e.g., immediate dominance, linearisation and instantiation) in combination with PS rules. By the phrase structure in (23), proposed in Luís&Otoguro (2004) for EP, the example in (24a) has the c-structure representation in (24b).

² This process is similar to tokenisation in XLE (e.g., Kaplan&Newman 1997, Butt et al. 1999, Kaplan et al. 2004).

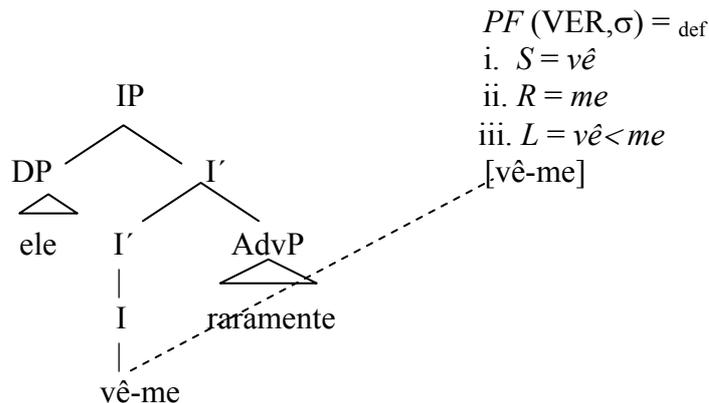
³ We position the finite verb in an I position, following the proposal in Luís&Otoguro (2004).

(23) (Luis&Otoguro 2004)



(24) a. O João vê-me raramente.
 the J. sees- ACC.1SG rarely
 'John sees me rarely'

b.



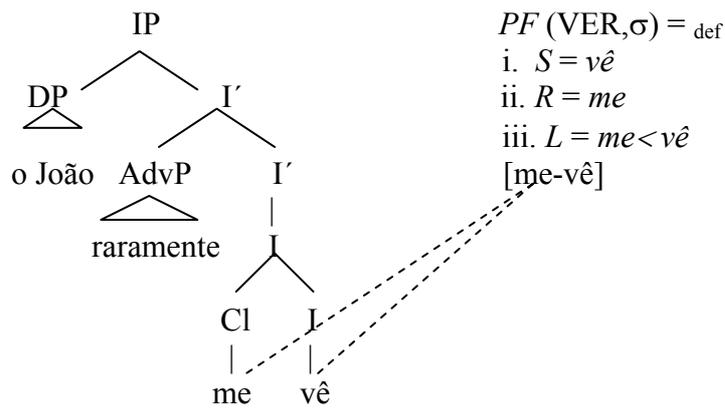
In (24b), the sequence *vê-me* is analysed as a single syntactic atom and, as predicted by the mapping in (21), it is positioned under one single c-structure node.

The mismatch is illustrated in (26), where a proclitic-verb combination is represented at the level of c-structure. Given the analysis in (21a), which associates the stem-affix sequence *me-vê* to two syntactic atoms, the sequence *me-vê* appears under two separate nodes. The correct insertion of the proclitic and the verb under Cl and I', respectively, is defined by the annotated phrase-structure rule in (25). The combination of (23) with (25) yields the c-structure representations in (26b).

(25) I → Cl Adv* I
 ↑ = ↓ ↓ ∈ (↑ ADJ) ↑ = ↓

(26) a. O João raramente me vê
 the J. rarely ACC.1SG sees
 'John rarely sees me'

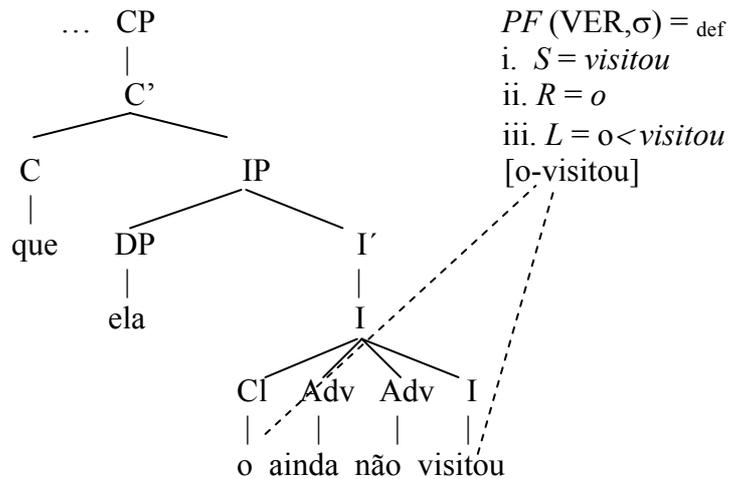
b.



In (27b), the proclitic is followed by interpolated adverbial particles: These are adjoined to I, following the proposal in Luís (2004). Again, by the annotated phrase-structure rule in (25), we represent (27a) as in (27b).

(27) a. Eu sei que ela o ainda não visitou.
 I know that she ACC.3SG.MASC yet not visited
 'I know that she has no visited him yet'

b.



4.2 Summary

We may summarise the assumptions underlying the Morphological Token analysis as follows:

- a) morphological strings are not allowed to be inserted directly into c-structure;
- b) the morphological well-formedness of words is defined in the morphology through morphology-internal principles such as the Paradigm Function which regulates the inflectional paradigm of a given language;
- c) at the interface between morphology and c-structure, a labelling algorithm takes as input morphological tokens and delivers labelled syntactic atoms;
- d) in c-structure, each syntactic atom instantiates a single syntactic terminal node, by general principles of phrase structure and PS rules.

The key goal of the analysis is to allow a single morphological token (i.e., stem-affix combination) to be mapped onto one or more syntactic atoms without incurring any violation of lexical integrity. In terms of the theoretical features of the analysis, we point out that our revised view of the interface between morphology and c-structure requires no changes in the formal model of c-structure trees, nor in the nature of the f-structure to c-structure mapping.

5. Conclusion and avenues for further research

In this paper, we have been concerned with the c-structure representation of proclitic affixes (i.e., phrasal affixes). What the above discussion has revealed is that it is possible to represent phrasal affixes without violating the integrity of words. Our claim is based on the view that ‘integrity’ is defined as a condition on morphological tokens (i.e., complete and well-formed stem-affix sequences defined through morphology-internal principles), rather than as a condition on the mapping between words and c-structure terminals.

The mapping we propose between morphological tokens and syntactic atoms finds theoretical support in the parallel linguistic structures of LFG grammar. Also, by assuming that each level is defined by its own set of well-formedness conditions, our proposal is in full harmony with the division of labour between morphology and syntax, one of the building blocks of lexicalist grammars.

In future research, we examine the scopal behaviour of proclitics in light of the c-structure representation provided in this paper. As alluded to in section 2, proclitics can take wide scope. Thus, any phrase structure representation should also accommodate these coordination properties. Also, further work will be

necessary to determine the different mismatch phenomena that our mapping theory can allow⁴.

References

- Anderson, Stephen R. (1992) *A-Morphous Morphology*. Cambridge Studies in Linguistics 62. Cambridge: CUP.
- Bresnan, Joan (2001) *Lexical Functional Syntax*. Oxford: Blackwell.
- Butt, Miriam et al. (1999) *A Grammar Writer's Cookbook*. Stanford, CA: CSLI Publications.
- Kaplan, Ronald & Paula Newman (1997) Lexical resource reconciliation in the Xerox Linguistic Environment. In D. Estival, et al. *Proceedings of the ACL Workshop on Computational Environments for Grammar Development and Linguistic Engineering*, 54-61. Chicago: Chicago University Press.
- Kaplan, Ronald et al. (2004) Integrating finite-state technology with deep LFG grammars. In *Proceedings of the Workshop on Combining Shallow and Deep Processing for NLP*. ESSLI 2004.
- Klavans, Judith (1985) The Independence of Syntax and Phonology in Cliticisation. *Language*, 61, pp. 95-120.
- Legendre, Géraldine (2000a) Positioning Romanian Verbal Clitics at PF: an Optimality-Theoretic Analysis. In: Birgit Gerlach and Janet Grijzenhout (eds.) *Clitics in Phonology, Morphology, and Syntax*. John Benjamins Publishing Company: Amsterdam/Philadelphia, pp. 219-254.
- Miller, Philp and Ivan Sag (1997) French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15 (3), pp. 573-639.
- Monachesi, Paola (1999) *A Lexical Approach to Italian Cliticization*. CSLI, Stanford.
- Luís, Ana (2004) *Clitics as Morphology*. PhD thesis, University of Essex.
- Luís, Ana (1997) *English Reduced Auxiliaries*. Ma. Diss. University of Essex.
- Luís, Ana and Ryo Otaguro (to appear) Morphological tokens and phrase structure: towards a treatment of morphology-syntax mismatches. In Ryo Otaguro, G. Popova and A. Spencer (eds.) *Essex Research Reports in Linguistics*. University of Essex.
- Luís, Ana and Ryo Otaguro (2004) Proclitic contexts and their effect on clitic placement. In Miriam Butt and Tracy Holloway King (eds.) *Proceedings of the LFG04 Conference*. Stanford, CA: CSLI Publications.

⁴ In Luís&Otaguro (to appear) we apply the Morphological Token analysis to Japanese lexical compounds which illustrates one other type of morphology-syntax mismatch. In the case of Japanese, we argue that two morphological tokens are treated as one syntactic atom at the level of c-structure. The mapping thus is from many to one, rather than from one to many as in EP.

- Luís, Ana and Louisa Sadler (2003) Object Clitic and marked Morphology. In: Beyssade, C., et al. (eds.) *Empirical Issues in Formal Syntax and Semantics 4*. Paris: Presses de l'Université de Paris-Sorbonne, pp. 133-153.
- Luís, Ana and Andrew Spencer (2005) Paradigm Function Account of 'mesoclitisis' in European Portuguese. In: G. Booij and J. van Marle, eds. *Yearbook of Morphology*. Kluwer: Dordrecht. Miller, P. and I. Sag (1997) French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15 (3), pp. 573-639.
- Luís, Ana and Andrew Spencer (to appear) Udi clitics: a Generalized Paradigm Function Approach. In Ryo Otoguro, G. Popova and A. Spencer (eds.) *Essex Research Reports in Linguistics*. University of Essex.
- Spencer, Andrew (1992) *Morphological Theory*. Oxford: Basil Blackwell.
- Spencer, Andrew (2000) Verbal clitics in Bulgarian: A Paradigm-Function approach. In: Birgit Gerlach and Janet Grijzenhout (eds.) *Clitics in Phonology, Morphology and Syntax*. Amsterdam/Philadelphia: John Benjamins, pp. 355-386.
- Spencer, Andrew (ms.) Extended Paradigm Function Morphology. Unpublished manuscript, University of Essex. (Available at <http://privateweb.essex.ac/~spena/papers/epfm.pdf>)
- Stump, Gregory T. (1993) Position classes and morphological theory. In: Booij, G. et al. (eds.) *Yearbook of Morphology 1992*. Kluwer: Dordrecht, pp. 129-180.
- Stump, Gregory T. (2001) *Inflectional Morphology. A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Wescoat, Michael (2002) *On Lexical Sharing*. PhD thesis, Stanford University.

Ana Luís
 University of Coimbra, Portugal
aluis@fl.uc.pt

Ryo Otoguro
 University of Essex, UK
rotogu@essex.ac.uk