Cliticization in European Portuguese using parallel morpho-syntactic constraints

Berthold Crysmann
Graduiertenkolleg Kognitionswissenschaft
Computerlinguistik
Universität des Saarlandes
crysmann@coli.uni-sb.de

Proceedings of the LFG97 Conference
University of California, San Diego
Miriam Butt and Tracy Holloway King (Editors)

1997

CSLI Publications

http://www-csli.stanford.edu/publications/

Abstract[1]

In this paper I shall propose an account of morphosyntactic phenomena like European Portuguese (henceforth: EP) cliticization in terms of interacting morphological and surface-syntactic constraints. In particular, I suggest that morphological constraints are partially underspecified with respect to morpheme linearization, information that will be added monotonically by surface syntactic constraints. Moreover, I will argue that, first, a purely lexicalist treatment of the morphosyntax of clitics in EP is quite difficult, if not impossible, to formulate, and, second, that the use of order domains provides us with a convenient representation for modelling the surface-syntactic constraints that determine the exact positioning of the clitics in the morphosyntactic complex. Finally, a lexically-constrained morpho-syntax interface is defined which enables us to express the basic intuitions of the lexicalist hypothesis.

Contents

1 Introduction

Within current surface-oriented lexical frameworks, the standard approach to cliticization phenomena in Romance assumes that affixation of pronominal clitics to their verbal hosts should be conceived of as an instance of lexical alternation. Consequently, under the analysis proposed by e.g. Miller and Sag (1995) these phenomena are dealt with in terms of lexical rules. This assumption, which conforms to the Lexical Integrity Principle (Bresnan and Mchombo, 1995), appears to be well-motivated, given that clitic constructions in Romance typically display a variety of properties

[1]This work has been financed by a scholarship from the *Deutsche Forschungsgemeinschaft* 'German Science Foundation' (DFG). I am gratefully indepted to my informants Antonio Branco (DFKI Saarbrücken, Universidade de Lisboa) and Otília Dias (Universität Hamburg). I would also like to thank the people at the Department of Computational Linguistics at University of the Saarland and at DFKI, in particular Antonio Branco, Wojciech Skut, and Hans Uszkoreit for discussing various aspects of the work reported here. Of course, all remaining errors are mine.

generally attributed to the lexical level, among them morphophonological and semantic idiosyncrasies.

European Portuguese (EP), however, poses a challenge for the Lexical Integrity Hypothesis and, furthermore, for the principle of "phonology-free syntax" (Pullum and Zwicky, 1988) in that morphophonology, semantics and most of morphotactics suggest an affixal status of these elements, whereas the linearization of the clitics with respect to their host depends on entirely syntactic factors, i.e. the presence vs. absence of certain syntactic constituents in a circumscribed surface-syntactic domain. I shall argue in particular that the grammar of EP clitic placement cannot be satisfactorily accounted for in a purely lexicalist fashion. Rather, morphology and syntax should be conceived of as mutually constraining modules which interact across a very narrow, lexically restricted interface. The analysis will be carried out in the framework of HPSG (Pollard and Sag, 1987, 1994) and, more specifically, the linearization approach argued for by Kathol and Pollard (1995), Pollard *et al.* (1993), and Reape (1994). As has been shown by Manning (1996), analyses formulated in "linearization HPSG" can easily be translated into current LFG.

The paper is organized as follows: In the next section, I shall present the evidence in favor of the lexical status of EP pronominal clitics (sec. 2.1), followed by an examination of the syntactic factors which partially determine clitic placement (sec. 2.2). The section will conclude with a critical perspective on the possibility to provide a purely lexicalist account of the data (sec. 2.3). In section 3, I shall outline an analysis in terms of parallel morpho-syntactic constraints. First, I shall propose to capture the surface-syntactic factors of clitic placement by means of linearization constraints (sec. 3.1). Next, I shall address the representation of clitic morphology suggesting an underspecified formulation of morpheme order (sec. 3.2). In a third step, an interface will be set up which defines a very restricted communication channel between morphology and syntax (sec. 3.3).

2 Morpho-syntactic Properties of EP Clitics

2.1 EP object clitics: Evidence for lexical status

Clitics in European Portuguese share a significant number of properties characteristic of "special clitics" found in other Romance languages, such as French or Italian: they cannot be topicalized, they cannot be substituted by full pronouns, they cannot be coordinated (cf. (1)) and they cannot be modified (cf. (2)):

(1) a. * *eu vi o e Paulo*
I saw him and Paul

b. * *eu não o e a conheço*
I not him and her know

(2) a. * *eu não os todos conheço*
I not them all know

b. * *eu conheço-os que estão aqui*
I know-them that are here

In addition to these properties, EP clitic order differs from the surface order of their non-clitic counterparts: while with ordinary object NPs, the direct object precedes the indirect object, the order is inverted with pronominal clitics.

Clitics in EP exhibit a relatively high degree of selection with respect to their host: with the exception of some northern dialects where the negative marker *não* 'not' may intervene between a preverbal clitic and its host, clitics are always positioned adjacent to the verb. Whenever the verb is affected by some syntactic process (like e.g. inversion), the clitic will be taken along.

Although these properties may also be reconciled with a syntactic approach to cliticization, they clearly suggest that these elements differ rather drastically from other syntactic constituents. Furthermore, under a lexical perspective, an account of these properties falls out directly.

Just as in Italian and French (cf. e.g. Grimshaw, 1982), EP witnesses lexicalized reflexive clitics (inherent pronominal and ergative) which cannot be given a compositional semantic interpretation. According to Zwicky and Pullum's (1983) criterion D, these semantic idiosyncracies clearly suggest an affixal status.

2.1.1 Morphology

Cliticization in EP is characterized by morphophonological alternations which do not apply across word boundaries and which have to be considered as highly item-specific.

Preceding 3rd person accusative clitics (*-o,-a,-os,-as*), oral coronal continuants (/s/, /z/, /r/) are replaced by [l] (cf. (3a-d)).

(3) a. *comprar + o → comprá-lo*
buy it buy-it
'to buy it'

b. *ela faz + o → ela fa-lo*
she does it she does-it
'she does it'

c. *tu fazes + o → tu faze-lo*
you do it you do-it
'you do it'

d. *nos + o + dão → no-lo dão*
us it give(3pl) us-it give(3pl)
'they give it to us'

e. *todos os alunos*
all the students
'all students'

It should be noted that this is an item-specific process which may therefore not be attributed to (automatic) surface phonology. In particular, it is not found with the homophonous definite article (cf. (3e)), thereby weakening the evidence for the kind of incorporation analysis favored by current transformational accounts (cf. e.g. Duarte *et al.*, 1995; Madeira, 1992). Furthermore, the process may not apply across word boundaries.

Similar evidence is provided by nasal epenthesis. Following nasal vowels, 3rd person accusative clitics surface with a nasal consonantal onset. Again, as can be seen from the contrast in (4b,c), this process does not apply across word boundaries.

(4) a. *eles conhecem + o/a → eles conhecem-no/-na*
they know him/her they know-him/her
'they know him/her'

b. * *eles não no/na conhecem*
they not him/her know

c. *eles não o/a conhecem*
they not him/her know
'they don't know him/her'

Perhaps the most striking argument for the lexical status of EP clitics comes from *mesoclisis*: with future and conditional tense forms, both accusative and dative clitics may be infixed between the tense morpheme and the stem.

Although, historically, these tense forms derive from a syntactically transparent construction, this is no longer true for contemporary EP. First, nothing except clitics may intervene between the stem form and the tense morpheme. Second, the particular ordering of the verbal stem and the tense/agreement morpheme does not reflect the overall word order found with true auxiliary + verb constructions. Moreover, future and conditional morphemes sometimes select irregular stem forms, as evidenced in (5).

(5) *fazer* 'do'

a. *fará + o → fá-lo-á*
do will(3sg) it do-it-will(3sg)
'(s)he will do it'

b. *faríamos + o → fá-lo-íamos*
do would(1pl) it do-it-would(1pl)
'we would do it'

From what has been said above, I conclude that EP clitics display a variety of properties generally attributed to the level of lexical morphology.

2.2 Clitic placement

EP cliticization exhibits an alternation between preverbal (proclitic) and postverbal (enclitic/mesoclitic) realization. As opposed to Italian, this alternation cannot be attributed to a "mere" finiteness distinction. Rather, clitic placement is sensitive to the presence of so-called "trigger" items in a certain *syntactic* domain. Presence of a trigger in a domain preceding the clitic host enforces proclisis (cf. (7)), without such a trigger, enclisis or mesoclisis has to be used (cf. (6)).

(6) Enclisis/Mesoclisis

a. *O João compra -o.*
John buys it
'John buys it.'

b. * *o João o compra*

c. *O João comprá-lo-á.*
John buy-it-will
'John will buy it.'

d. * *o João o comprará*

(7) Proclisis

a. *O João não o comprará.*
John not it buy will
'John will not buy it'.

b. * *o João não compra-lo-á*

c. *que os amigos lhes deram livros*
that the friends to them gave books
'that the friends gave them books'

d. * *que os amigos deram-lhes livros*

e. *as pessoas a quem o contámos*
the people to whom it told(1pl)
'the people we told it'

f. * *as pessoas a quem contámo-lo*

g. *Que mentira lhe contaste?*
which lie him told(2pl)
'Which lie did you tell him?'

h. * *que mentira contaste-lhe?*

i. *Todos os alunos lhe telefonaram.*
all the students to him phoned
'All the students phoned him.'

j. * *todos os alunos telefonaram-lhe*

k. *O João raras vezes me dá razão.*
John seldom me gives reason
'John seldom agrees with me.'

l. * *o João raras vezes dá-me razão*

Proclisis triggers comprise a fairly heterogeneous set of elements, among them negation, complementizers, relatives, wh-expressions, (some) quantified NPs, focussed NPs, (some) prepositions and (some) temporal adverbials.

It is worth noting that the surface position of trigger items seems to play a much more important role than e.g. scope properties: with inverted quantified subjects, enclisis/mesoclisis obtains (compare (8a,b) to (7i,j)), despite the fact, that the postverbal trigger unambiguously has scope over the main predicate.

(8) a. *Telefonaram -lhe todos os alunos.*
phoned to him all the students
'All the students phoned him.'

b. * *lhe telefonaram todos os alunos*

c. *Os alunos telefonaram -lhe todos.*
the students phoned to him all
'The students all phoned him.'

d. * *os alunos lhe telefonaram todos*

e. *Ele lê -o raras vezes.*
he reads it seldom
'He seldom reads it.'

f. * *ele o lê raras vezes*

The same holds for floating quantifiers (cf. (8c,d)) and postverbal temporal adverbials (compare (8e,f) to (7k,l)).

What makes an account of the proclisis/enclisis alternation even more tricky is the fact that non-argument clitics (inherent pronominal and ergative *se* 'self'; cf. (9) and (10)) are not exempt from this phenomenon. Recall from section 2.1 that these elements are generally considered as being lexicalized with the verb they are attached to, primarily owing to the fact that they cannot be given a compositional semantic interpretation.

(9) Intrinsic *se*

a. *Nós rimos-nos da Maria.*
we laughed of-the Maria
'We laughed at Maria.'

b. * *nós nos rimos da Maria*

c. *Todos nós nos rimos da Maria.*
all we laughed of-the Maria
'We all laughed at Maria.'

d. * *todos nós rimos-nos da Maria*

(10) Ergative *se*

a. *O gelado derreteu-se com o calor.*
the ice-cream melted in the heat
'The ice-cream melted in the heat.'

b. * *o gelado se derreteu com o calor*

c. *O gelado não se derreteu com o calor.*
the ice-cream not melted in the heat
'The ice-cream did not melt in the heat.'

d. * *o gelado não derreteu-se com o calor*

To conclude my presentation of EP clitic placement, let us briefly take a look at the examples in (11):

(11) a. *Vi os meninos a abraçarem-se.*
saw(1sg) the children to embrace-REFL
'I saw the children embracing each other.'

b. * *vi os meninos a se abraçarem*

c. *Não vi os meninos a abraçarem-se.*
Not saw(1sg) the children to embrace-REFL
'I did not see the children embracing each other.'

d. * *não vi os meninos a se abraçarem*

e. * *vi os meninos a não abraçarem-se*

f. *Vi os meninos a não se abraçarem.*
saw(1sg) the children to not REFL embrace
'I saw the children without them embracing each other.'

As can be observed from the contrasts between (11c,e) and (11d,f), respectively, linear precedence of a trigger is always relative to a local domain.

2.3 Problems with Lexicalism

A purely lexical account of these phenomena is far from being trivial: while one might try to represent modifiers, wh-phrases, relatives, negation, and quantifiers at the lexical level already (e.g. using techniques like type-raising (Manning *et al.*, 1995; Kim and Sag, 1995), extraction lexical rules (Sag and Fodor, 1994), and lexicalized quantification (Manning *et al.*, 1995)), such an account will still face severe problems. First, the representation of trigger elements will be scattered across different parts of the lexical entry (e.g. valence features, nonlocal features, quantifier storage) making a unique cliticization rule quite difficult to state. Second, it will be hard to explain without appeal to surface configurations why postverbal floating quantifiers never trigger proclisis, even if they unambiguously quantify over the preverbal subject. In quantificational terms, preverbal and postverbal quantifiers would then be indistinguishable. The same argument holds for certain temporal adverbs (like *sempre* 'always' or *raras vezes* 'seldom') which take scope over the event denoted by the finite verb, regardless of surface position. Likewise, these adverbs only trigger proclisis when used preverbally. Third, it would be necessary to introduce complementizers and prepositions on the lexical entry of the verbal head whose sentential projection they embed. Finally, it remains unclear how to model effects of surface intervention where a preposed constituent makes mesoclisis possible despite a complementizer triggering proclisis (12).

(12)

Acho	*que*	*ao*	*João,*	*far*	*-lhe*	*-ia*	*bem*	*ir*	*a*	*festa.*
think(1sg)	that	to	John	do	to him	would	well	go	to the	party

'I think that it would be good for John to go to the party.'

An alternative strategy to preserve the syntactic opacity of morphologically derived forms would be to encode information about clitic placement by means of verbal types, i.e. classifying derived lexical verbs as *proclisis-verb*, *mesoclisis-verb*, or *enclisis-verb*. Apart from being rather ad hoc, this strategy can hardly obscure the fact, that it is indeed internal morphophonological information which is made available to the syntactic level: although cliticization may then be represented as a lexical phenomenon, such an approach runs counter to the spirit of the integrity hypothesis.

3 Parallel Morphological and Syntactic Constraints[2]

I shall propose an alternative account in terms of parallel morphological and syntactic constraints. The separation of morphological and syntactic information, along with a lexically constrained interface across which they establish a very restricted communication channel will enable us to preserve the basic notions of lexical integrity while still being able to account for the data under discussion. The approach to be described is to a certain extent inspired by Sadock's (1991) work on morpho-syntactic mismatches.

In order to achieve this goal, three steps are necessary: first, the syntactic determinants of clitic placement are modelled by virtue of linearization constraints. Second, a fragment of the morphophonology of EP cliticization is implemented which shows how unification of partial morphological schemata may determine the set of possible EP clitic-verb complexes. In a third step, then, the morpho-syntax interface is defined which performs the task of mediating between syntactic and morphological linearization constraints as a kind of homomorphism constraint.

3.1 Linear Syntax

As we have seen above, cliticization in EP poses severe problems for a lexicalist approach to cliticization phenomena. This is mainly due to the fact that linearization of the clitics relative to the verbal head involves pure surface phenomena like the order of trigger items. The elements found at the left periphery of the sentence typically include the subject, modifiers, fillers, and complementizers, elements which are standardly assumed to be licensed by a variety of different ID-Schemata, thereby introducing an extremely layered phrase structure. Thus, under a phrase-structural perspective, some additional machinery is due in order to provide flat domains in which linearization constraints can apply.

Two approaches are conceivable: we can either postulate entirely flat constituent structures, or, alternatively, decouple the domain in which linearization constraints apply (phenogrammar) from constituent structure (tectogrammar). Adopting the first approach confronts us with at least two disadvantages, one theoretical and one empirical in nature: assuming flat constituent structures

[2]Throughout this section, $\bigcirc$ will denote the shuffle or sequence union operation (cf. Reape, 1994), $\prec$ will stand for immediate precedence and $\frown$ will perform an ordinary concatenation.

which simultaneously license subject and adjunct daughters calls for a major revision of the SEMANTICS PRINCIPLE, a principle which determines the meaning of a phrase compositionally by virtue of the meaning of its daughters. Standard HPSG (Pollard and Sag, 1994) assumes that the semantic head of HEAD–SUBJECT structures is the syntactic head daughter, whereas the semantic head of HEAD–ADJUNCT structures is the non-head (i.e. adjunct) daughter. Furthermore, EP does not provide us with the kind of evidence found in "free word order" languages such as German where all major constituents, including complements, scramble rather freely.

I therefore conclude that EP has a layered constituent structure in general. Within the context of clitic placement, however, the order-relevant domain spans the entire preverbal field. Thus, a mechanism is called for to decouple order-relevant domains from immediate dominence. This task could be performed by adopting a mechanism of complex domain formation (cf. Kathol and Pollard, 1995; Pollard *et al.*, 1993). Under this conception of order domains, constituents can selectively be liberated from ordering within the immediately dominating node. Technically, this is achieved by the introduction of a list-valued feature DOM of domain objects and relational constraints on the formation of these lists. Complex order domains are then built up using *sequence union* (Reape, 1994), i.e. "shuffling" together elements from the DOM lists of the daughters while preserving the relative order imposed by linearization constraints on lower domains.

Compaction on a list of domain objects orders these objects according to the linearization constraints of the grammar and encapsulates them in a single domain object, thereby precluding further ordering in higher domains. Partial compaction, however, permits specific constituents to escape compaction into a single domain object. Instead, these domain objects will be inherited by higher order domains.

In contrast to entirely flat constituent structures, these approaches permit us to precisely define the upper and lower bounds of the relevant order domains. Under such an approach, the upper bound of the order domain is defined using compaction on both HEAD-MARKER- and HEAD-FILLER-Schemata, thus excluding a triggering complementizer from the relevant order domain just in case a filler intervenes (cf. example (12)). Thus, linear intervention may be accounted for by hierarchical intervention of an appropriate ID–Schema. Likewise, partial compaction on the HEAD-COMPLEMENT-Schema will preserve constituenthood of the VP, except for the head which participates in the preverbal field. As noted above, this is supported by the empirical observation that complements do not scramble into the preverbal field. In contrast, inversion of the verb may be regarded as further empirical evidence for assuming the liberation of the verbal head from its own projection.

Having restricted the domain of application, linearization constraints can now be formulated in a quite convenient way using regular expressions (cf. figs. 1 and 2).[3]

Note that, at this level, no distinction is drawn between mesoclisis and enclisis: as the alternation is purely morphologically induced, they can both be treated as syntactically postverbal.

3.2 Morphology

On the morphological side, the analysis is based on the framework of constraint-based HPSG-phonology formulated by Bird and Klein (1993) for which there is an efficient one-level imple-

[3]Of course, an analysis of EP clitic placement can not be considered complete until a much more precise characterization of trigger elements can be given. Yet, this task is clearly beyond the scope of this article. A first step in this direction, however, is taken in Crysmann (1997).

$$[]^* \left\{ \begin{bmatrix} \text{CONT} & \textit{quant} \end{bmatrix} \; \begin{bmatrix} \text{MARKING} & \textit{que} \end{bmatrix} \; \begin{bmatrix} \text{MOD} \ldots \text{REST} & \boxed{1} \\ \text{CONT}|\text{RESTR} & \neg \boxed{1} \end{bmatrix} \right\} [\text{CONT } \textit{pron}]^* \; []^* \; [\text{HEAD } \textit{verb}]$$

Figure 1: Proclisis linearization constraint

$$\neg \left\{ \begin{bmatrix} \text{CONT} & \textit{quant} \end{bmatrix} \; \begin{bmatrix} \text{MARKING} & \textit{que} \end{bmatrix} \; \begin{bmatrix} \text{MOD} \ldots \text{REST} & \boxed{1} \\ \text{CONT}|\text{RESTR} & \neg \boxed{1} \end{bmatrix} \right\}^* [\text{HEAD } \textit{verb}] \; [\text{CONT } \textit{pron}]^*$$

Figure 2: Enclisis linearization constraint

mentation by means of FSAs (Bird and Ellison, 1992). Morphophonological constraints are formulated using morphological schemata (cf. Riehemann, 1994), which are partially underspecified with respect to the exact linear position of the clitic cluster. In general, these schemata specify that the PHON value of the toplevel *sign* is an unordered concatenation (i.e. *shuffle*) of the PHON values of its parts (i.e. the elements on the MORPHS list). Morpheme order, as well as morphophonological alternations are imposed as constraints attached to the toplevel PHON value.

Fig. 3 shows how the morphophonological alternation illustrated in (3) can be incorporated into the morphological schemata in a completely declarative way (cf. Bird and Klein, 1993). Note that this alternation is stated without making reference to particular morphemes or morpheme orders. Whatever happens to appear to the immediate left of the clitic has to match either the first or the second description.

$$\left[\begin{array}{ll} \text{PHON} & \boxed{1} \bigcirc \boxed{2} \bigcirc \boxed{3} \wedge \boxed{2} \prec \boxed{1} \\ \text{MORPHS} & \boxed{4} \bigcirc \left\langle \left[\text{PHON} \quad \boxed{2} \frown \left\{ \begin{array}{l} \text{s} \\ \text{z} \\ \text{r} \end{array} \right\} \right] \right\rangle_{\textit{morph}} \bigcirc \left\langle \left[\begin{array}{ll} \text{PHON} & \boxed{1} \langle \text{lo} \rangle \\ \text{DOM} & \left\langle \left[\begin{array}{ll} \text{PHON} & \langle\rangle \\ \text{CONT} & \textit{pron} \end{array} \right] \right\rangle \end{array} \right] \right\rangle_{\textit{aff}} \end{array} \right]$$

$$\left[\begin{array}{ll} \text{PHON} & \boxed{1} \bigcirc \boxed{2} \bigcirc \boxed{3} \wedge \neg \left(\boxed{2} \left\langle \ldots \left\{ \begin{array}{l} \text{s} \\ \text{z} \\ \text{r} \end{array} \right\} \right\rangle \wedge \boxed{2} \prec \boxed{1} \right) \\ \text{MORPHS} & \boxed{4} \bigcirc \left\langle \left[\text{PHON} \quad \boxed{2} \right] \right\rangle_{\textit{morph}} \bigcirc \left\langle \left[\begin{array}{ll} \text{PHON} & \boxed{1} \langle \text{o} \rangle \\ \text{DOM} & \left\langle \left[\begin{array}{ll} \text{PHON} & \langle\rangle \\ \text{CONT} & \textit{pron} \end{array} \right] \right\rangle \end{array} \right] \right\rangle_{\textit{aff}} \end{array} \right]$$

Figure 3: The Clitic-Schema

Additional morphotactical constraints may be imposed, like e.g. future/conditional affixes forc-

ing all other morphemes, including the clitic cluster, to appear to their left (cf. fig. 4).

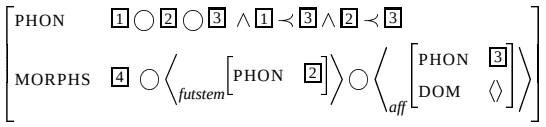

Figure 4: The Future Schema

Once these schemata (i.e. figs. 4 and 3) are unified, they denote the set of possible future tense verb forms with an accusative clitic. Note that further clitic schemata can be added monotonically, deriving increasingly complex forms. Restrictions on the linear order among the clitics (e.g. *dat* $\prec$ *acc*) can easily be expressed by linearization constraints on clitic phonology.

The morpho-syntactic differences between ordinary affixes and clitics are captured by having the clitics specify a non-empty list of domain objects. When the schema is instantiated to a particular lexical entry, this list is sequence unioned with the DOM value of the toplevel *sign*. The resulting list of domain objects constitutes the lexical interface to syntax. The objects contain both categorial information as well as underspecified (kernel) phonology. In principle, the case where a lexical sign introduces a domain list with more than one domain object is just a subcase of the general organisation of HPSG *sign*s proposed by Pollard *et al.* (1993).

3.3 Interfacing Morphology and Syntax

The morpho–syntactic properties of EP cliticization are obtained by unification of partial syntactic and morphological constraints. As the entities on which these constraints may operate are quite distinct (e.g. domain objects vs. stems and affixes), a translation step is necessary. A homomorphism constraint is formulated which ensures a certain correspondence of the descriptions while permitting us to maintain a modular view of syntax and morphology.

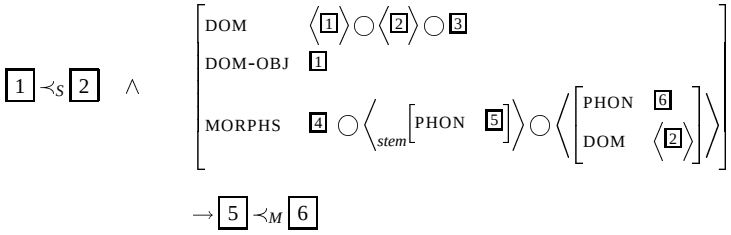

Figure 5: The Syntax-Morphology Homomorphism

This translation is performed by the syntax-morphology homomorphism constraint: as defined in fig. 5, this implication holds between a morphological linearization and a syntactic one just in case there is a lexical sign which provides at least two domain objects, one corresponding to the domain object as a whole (in our case the clitic-verb complex, 1 in fig. 5), the other to one of its parts (the clitic, cf. 2). If this description is met, i.e. if the clitic is syntactically required to appear last, it follows that clitic phonology (6) has to be ordered after the stem (5) in morphology.

Note that the particular setup of the interface does not affect ordering of the clitics relative to each other. Thus, morphology and syntax may specify conflicting linearization constraints, e.g. *acc* $\prec_S$ *dat* and *dat* $\prec_M$ *acc*.

To conclude this section, I will briefly take stock of what has been achieved thus far: a strict notion of lexicalism has been replaced by mutually constraining morphological and syntactic descriptions. The interaction of these levels enables us to deal with morphophonological processes cleanly in a purely morphological module, while dealing with surface induced order variation completely in terms of syntactic linearization constraints. The way the interface is set up ensures that we do not have to deal with morphological entities at the syntactic level. Furthermore, the interface is morpho-lexically constrained, a property which may capture the fact that these instances of morpho-syntactic mismatch are highly marked constructions cross-linguistically which tend to be eliminated in contemporary EP (cf. Duarte *et al.*, 1995).

4 Conclusion

Based on cliticization data from EP, an analysis in terms of parallel morphological and syntactic constraints has been proposed. A lexically constrained interface has been defined which allows for a very restricted interaction between the modules. While this approach clearly weakens the strict notion of lexical integrity where morphological objects and syntactic atoms are in a one-to-one correspondence, it nevertheless enables us to flesh out the basic intuitions behind the lexicalist hypothesis, i.e. a clear separation of morphological and surface-syntactic operations.

The approach described here can easily be extended to account for related phenomena in a variety of languages, e.g. it could be used to provide Gunji's (1995) linearization analysis of Japanese causatives with a treatment of the morphophonological aspects. Moreover, the assumption that words may introduce more than one domain object is indepently motivated by phenomena such as German separable prefix verbs (Kathol, 1996; Stiebels and Wunderlich, 1992).

Extending the analysis in minor respects will allow us to account for the fact that, in European Portuguese, preverbal clitics may take wide scope over conjoined predicates.

References

Bird, S. and T. M. Ellison (1992), "One level phonology: Autosegmental representations and rules as finite state automata," Tech. Rep. 51, University of Edinburgh, Center for Cognitive Science. 10

Bird, S. and E. Klein (1993), "Enriching HPSG phonology," Research Paper EUCCS/RP-56, Centre for Cognitive Science, University of Edinburgh. 9, 10

Bresnan, J. and S. A. Mchombo (1995), "The lexical integrity principle: Evidence from Bantu," *Natural Language and Linguistic Theory* **13**, 181–254. 1

Crysmann, B. (1997), "(Im)proper quantifiers and clitic placement in European Portuguese," paper presented at the Colloque de Syntaxe et Sémantique de Paris (CSSP), Université de Paris. 9

Duarte, I., G. Matos and I. Faria (1995), "Specificity of European Portuguese clitics in Romance," in I. H. Faria and M. J. Freitas, eds., *Studies on the Acquisition of Portuguese*, Edições Colibri, Lisboa. 3, 12

Grimshaw, J. (1982), "On the lexical representation of Romance reflexive clitics," in J. Bresnan, ed., *The Mental Representation of Grammatical Relations*, pp. 87–148, The MIT Press, Cambridge, Massachusetts. 3

Gunji, T. (1995), "On lexical treatment of Japanese causatives," manuscript, Osaka University. 12, 12

Kathol, A. (1996), "Discontinuous lexical entries," manuscript, paper presented at the Third International Conference on HPSG, Marseille. 12

Kathol, A. and C. Pollard (1995), "Extraposition via complex domain formation," manuscript, Ohio State University. 2, 9

Kim, J.-B. and I. A. Sag (1995), "The parametric variation of French and English negation," in *Proceedings of WCCFL*, Stanford, CSLI publications. 7

Madeira, A. M. (1992), "On clitic placement in European Portuguese," in *Working Papers in Linguistics*, vol. 4, pp. 97–122, London, Department of Phonetics and Linguistics, University College London. 3

Manning, C. (1996), "Dissociating functor-argument structure from surface phrase structure: the relationship of HPSG order domains to LFG," manuscript, Carnegie Mellon University. 2

Manning, C., I. A. Sag and M. Iida (1995), "The lexical integrity of Japanese causatives," manuscript. 7, 7

Miller, P. H. and I. A. Sag (1995), "French clitic movement without clitics or movement," manuscript, Université de Lille 3, Stanford University. 1

Pollard, C. and I. Sag (1987), *Information–Based Syntax and Semantics*, vol. 1, CSLI, Stanford. 2

Pollard, C. and I. Sag (1994), *Head–Driven Phrase Structure Grammar*, CSLI and University of Chicago Press, Stanford. 2, 9

Pollard, C., R. Levine and R. Kasper (1993), "Studies in constituent ordering: Toward a theory of linearization in Head-Driven Phrase Structure Grammar," Grant Proposal to the National Science Foundation, Ohio State University. 2, 9, 11

Pullum, G. and A. Zwicky (1988), "The syntax-phonology interface," in F. J. Newmeyer, ed., *Linguistics: The Cambridge Survey*, vol. 1, pp. 255–280, New York, Cambridge University Press. 2

Reape, M. (1994), "Domain union and word order variation in German," in J. Nerbonne, K. Netter and C. Pollard, eds., *German in Head-Driven Phrase Structure Grammar*, no. 46 in Lecture Notes, pp. 151–197, CSLI Publications, Stanford University. 2, 8, 9

Riehemann, S. (1994), "Morphology and the hierarchical lexicon," manuscript, CSLI, Stanford. 10

Sadock, J. M. (1991), *Autolexical Syntax. A Theory of Parallel Grammatical Representations*, The University of Chicago Press, Chicago. 8, 8

Sag, I. A. and J. D. Fodor (1994), "Extraction without traces," in *West Coast Conference on Formal Linguistics*, vol. 13, Stanford University, CSLI Publications/SLA. 7

Stiebels, B. and D. Wunderlich (1992), "A lexical account of complex verbs," *Theorie des Lexikons. Arbeiten des SFB 282* **30**, 1–43. 12

Zwicky, A. and G. K. Pullum (1983), "Cliticization vs. inflection: English *n't*," *Language* **59**, 502–513. 3, 3