

# The Morphosyntax of Romanian Cliticization

Paola Monachesi

University of Utrecht, Uil-OTS

## Abstract

In this paper, motivations are presented to argue in favor of the affixal status of Romanian pronominal clitics. It will be suggested that they should not be considered lexical items, i.e. 'signs', which are located in a special position by rules of syntax, but a complex of syntactic and semantic information which is provided in the lexicon for the morphophonological realization of the cliticized verb form. Appropriate constraints are proposed to relate this bundle of features which is present on verbs to the actual phonological realization of the clitics. Clitics are thus the 'spell-out' of certain morphosyntactic features of the verb. It will be shown that this morphophonological approach can provide an answer to certain peculiar facts about Romanian clitic climbing. The analysis presented thus provides a comprehensive account of Romanian cliticization since it deals both with the morphophonological properties and with the syntax of Romanian clitics.

## 1 Introduction

The status of pronominal clitics is an issue which is still under discussion; no definite answer has been provided in the literature to the question of whether pronominal clitics should be treated as independent syntactic forms or as affixes.

Within the generative tradition it is often assumed that they are syntactically independent elements, hence *words*, whose properties should be accounted for in syntax. Approaches in this direction are, among others, those of Kayne (1975) and subsequent work, Rizzi (1982), Burzio (1986), Sportiche (1993) and Dobrovie-Sorin (1994) which deals mainly with Romanian clitics. A shortcoming of these works is that they have focussed on the syntactic behavior of clitics and have neglected their morphological and phonological properties. On the other hand, morphological approaches such as that of Bonet (1991) have presented an analysis of certain crucial problems related to the morphology of cliticization such as that of the synthetic clitic clusters or that of the co-occurrence restrictions in clitic combinations. However, they do not provide a clear account of the syntactic contribution of cliticization.

In this paper, I propose an analysis that takes both the morphophonological and the syntactic aspect of cliticization into consideration. I argue that Romanian clitics are best treated as *affixal elements* on a par with French clitics (Miller 1992a) and Italian clitics (Monachesi 1996). They will be not considered lexical items, i.e. *signs*, but featural information which is provided in the lexicon and used in morphology and phonology for the realization of the cliticized verb form. The crucial issue is then how to relate this featural information to the actual phonological realization of the clitic. I propose that appropriate constraints are responsible for the realization of the clitics. In this view, clitics are considered the 'spell-out'

of certain morphosyntactic features of the verb. This morphophonological analysis shares similarities with *realizational* approaches to morphology such as that of Anderson (1992) or Stump (1992). Furthermore, I show that under this analysis, certain puzzling facts about clitic climbing can be accounted for.

The paper is organized as follows: section 2 discusses the status of Romanian pronominal clitics and arguments are provided in favor of their affixal behavior. A lexical analysis of cliticization is proposed in section 3; it is cast within the framework of Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag 1987, Pollard and Sag 1994). In section 4, I discuss clitic climbing which is triggered by auxiliary verbs and its peculiarities. Section 5 contains some speculations about enclitic placement on hosts which are different from the verb, while section 6 contains some concluding remarks.

## 2 The affixal status of pronominal clitics

There is convincing evidence that Romanian pronominal clitics behave as affixes. The arguments to support this position are mainly based on Zwicky and Pullum (1983) and Zwicky (1985) in which tests are proposed for distinguishing affixes from nonaffixes (or words from nonwords).

### 2.1 Rigid order

The order of pronominal clitics in the cluster is fixed. Romanian has only accusative and dative object clitics and they occur in the order *dative-accusative*, as shown in (1a). The opposite order would be ungrammatical as can be seen in (1b):

- (1) a. Alexandru mi-l trimite astăzi.  
 Alexandru cl.(dat-1sg) cl.(acc-3msg) sends today  
 'Alexandru sends it to me today.'
- b. \* Alexander îl-mi trimite astăzi.  
 Alexander cl.(acc-3msg) cl.(dat-1sg) sends today

As pointed out by Anderson (1992), there is a clear similarity between clitics and affixes, which also exhibit rigid order. Independent words are, on the other hand, usually allowed a certain degree of free ordering.

### 2.2 Coordination

Coordination also seems to argue in favor of affixal status for pronominal clitics. The crucial test in this respect is whether clitics can have wide scope over a coordination of hosts or have to be repeated on each conjunct.

If clitics have word status one would expect it to be possible for them to have wide scope over coordination since this is the case for syntactic words. On the other hand, if they must be repeated on each conjunct they must have affixal status.<sup>1</sup>

<sup>1</sup>See Miller (1992b) for a formulation of these intuitions in the *Coordination Criteria*.

In Romanian, if two verbs are coordinated which share the same clitic, the clitic has to be repeated within each conjunct:

- (2) el o dorea și o căuta.  
 he cl.(acc-3fsg) desires and cl.(acc-3fsg) looks for  
 'He desires her and looks for her.'

It cannot have wide scope over coordination:

- (3) \* el o dorea și căuta.  
 he cl.(acc-3fsg) desires and looks for  
 'He desires her and looks for her.'

The data thus seem to support the affixal status of Romanian pronominal clitics.

### 2.3 Arbitrary gaps

Zwicky and Pullum (1983) point out that arbitrary gaps in the set of combinations can occur with affixes. This is the case in inflectional paradigms, where verbs might lack certain forms.

Arbitrary gaps are also present in Romanian when clitics combine with the verb. Farkas and Kazazis (1980) mention that not all *dative-accusative* person and number combinations are grammatical in Romanian. In particular, it is not possible to have the first person accusative clitic *mă* together with a dative one:

- (4) a. \* ți m-au dat de nevastă numai pentru că  
 cl.(dat-2sg) cl.(acc-1sg) have given as wife only because  
 ai insistat.  
 have insisted  
 'They gave me in marriage to you only because you have insisted.'
- b. \* am auzit că părinți mei vor să i mă  
 have heard that parents mine want that cl.(dat-3sg) cl.(acc-1sg)  
 dea de nevastă.  
 give as wife  
 'I have heard that my parents want to give me in marriage to him.'

Similarly, a first person singular or plural clitic together with a second person plural one does not yield a grammatical result:<sup>2</sup>

- (5) a. \* vor să mi vă omoare.  
 want that cl.(dat-1sg) cl.(acc-2pl) kill  
 'They want to kill you on me.'

<sup>2</sup>In these sentences, the dative clitic should be interpreted as an ethical dative. A translation of these 'ethical dative constructions' is nearly impossible to achieve.

- b. \* vor să ni vă omoare.  
 want that cl.(dat-1pl) cl.(acc-2pl) kill  
 'They want to kill you on us.'

Farkas and Kazazis (1980) make an attempt to explain the ungrammaticality of the sentences in (4) in terms of topicality. They assume two topicality hierarchies:

- (6) a. *Ethical > Goal > Theme*  
 b. *I person > II person > III person*

They claim that the ungrammatical sentences in (4) are the consequence of a conflict in the two hierarchies in (6) which arises when the clitic sequence involves the extremities of the hierarchies.<sup>3</sup> However, the topicality hierarchies cannot explain the ungrammaticality of the sentences in (5). Farkas and Kazazis admit that they should postulate an additional arbitrary constraint in order to rule them out.

This fact suggests that it is not possible to provide a principled account of the ungrammatical clitic sequences suggesting that they should be considered cases of arbitrary gaps. This is evidence in favor of the affixal status of clitics and it supports an analysis of these data in morphology which is the appropriate locus for dealing with this kind of arbitrariness.

## 2.4 Clitic doubling

Perhaps the most convincing evidence in favor of the affixal status of Romanian pronominal clitics comes from the fact that they can co-occur with full complements behaving in these cases virtually as agreement markers:

- (7) a. Maria îi dă președintelui un buchet.  
 Maria cl.(dat-3sg) gives president-the a bouquet  
 'Maria gives the president a bouquet.'
- b. Ion m-a văzut pe mine.  
 Ion cl.(acc-1sg) has seen PE me  
 'Ion saw me.'

In example (7a), the dative clitic *îi* co-occurs with the indirect object *președintelui*, while in (7b) the object clitic co-occurs with the direct object. It should be noted that the doubled NP is usually preceded by the marker *pe*. However, a direct or indirect object can be expressed simply by means of a clitic:

- (8) a. Maria îi dă un buchet.  
 Maria cl.(dat-3sg) gives a bouquet  
 'Maria gives him a bouquet.'

<sup>3</sup>It should be noted that no independent motivation is given for the order in (6), which needs to be modified in order to account for the fact that certain additional sequences are ungrammatical in languages such as French and Modern Greek.

- b. Ion m-a văzut.  
 Ion cl.(acc-1sg) has seen  
 'Ion saw me.'

The examples in (8) are therefore similar to those found in other Romance languages such as Italian or French, where clitics and full complements are in complementary distribution.

## 2.5 Morphophonological idiosyncrasies

It is usually assumed that the phonological shape of affixes can be affected by the phonology of the stem or of other affixes with which they combine. Romanian pronominal clitics behave as affixes in this respect since in specific contexts they can undergo certain changes. Optional vowel deletion occurs if a clitic ending in *ă* precedes a verb beginning with unstressed *a* or *o*:

- (9) a. mă așteaptă.  
 cl.(acc-1sg) waits  
 'He waits for me.'
- b. m-așteaptă.  
 cl.(acc-1sg) waits

However, if the verb is an auxiliary, vowel deletion is obligatory.

- (10) a. m-a invită.  
 cl.(acc-1sg) has invited  
 'He has invited me.'
- b. \*mă a invită.  
 cl.(acc-1sg) has invited

Similarly to clitics ending in *ă*, the clitic *se* undergoes optional vowel deletion in front of a verb which begins with unstressed *a* or *o*:

- (11) a. se așeză.  
 cl.(acc-3sg-refl) sits  
 'He sits.'
- b. s-așeză.  
 cl.(acc-3sg-refl) sits

On the other hand, there is no vowel deletion if the clitics *te*, *ne* or *le* occur in a similar context:

- (12) a. te așteaptă.  
 cl.(acc-2sg) waits  
 'He waits for you.'

- b. \* t-așteaptă.  
cl.(acc-2sg) waits

It seems that the morphophonological changes mentioned above could be naturally interpreted if clitics are considered affixes, while they would be unexpected if they are considered independent words.

## 2.6 Morphophonological properties of Romanian pronominal clitics

The morphophonological idiosyncrasies that I discussed above suggest that different allomorphs should be postulated in the case of pronominal clitics. Three different classes can be distinguished:<sup>4</sup>

Table 1: Different classes of clitics

Person	1		2		3			4	5	6			3,6	
	D	A	D	A	D	A		DA	DA	D	A		D	A
						m.	f.				m.	f.		
Class 1	îmi	mă	îți	te	îi	îl	o	ne	vă	le	îi	le	își	se
Class 2	mî-	m-	ți-	te-	i-	l-	o-	ne-	v-	le-	i-	le-	și-	s-
Class 3	-mî	-mă	-ți	-te	-i	-l	-o	-ne	-vă	-le	-i	-le	-și	-se

Class 1 includes those clitics that occur in front of any verbs (*proclitics*) while in Class 2 are those clitics that can occur only before verbs that begin with *a* or *o*. Those pronominal clitics that follow the verb (*enclitics*) are grouped in Class 3. Certain changes and unexpected forms appear also in the combination of two clitics. This fact seems to suggest that pronominal clitics are not realized one at a time, but they cluster together forming a unit. If clitics were realized one at a time, one would expect forms such as *ne le* or *ne îl, le le* where the dative form is the one that occurs either before a vowel or a consonant. However, this is not the case, as can be seen in the table below, which gives the possible combinations of clitics: It can be noted that there is a special form of the dative clitic *ni, vi, li* which surfaces only when it is in combination with another pronominal clitic. However, these clitics cannot occur in combination with the third person feminine accusative form *o*. It is not clear how this behavior could be explained; it does not seem possible to postulate appropriate phonological rules that could clarify it.

The data above seem to support the hypothesis that the combination of two clitics constitutes a new unit and does not result from the composition of two single forms.

Additional motivation in support of the idea that pronominal clitics form a unit is provided by the following facts. Romanian clitics can phonologically cliticize to

<sup>4</sup>The hyphen which is present in the case of Class 2 and Class 3 clitics illustrates an orthographic convention that requires these forms to be written attached to the verb. However, in certain cases, its presence reflects a different pronunciation (cf. Avram 1986).

Table 2: Possible combinations of two clitics

	3 Acc Sg Ms	3 Acc Sg F	3 Acc Pl Ms	3 Acc Pl F
1 Dat Sg	mi-l	mi-o	mi-i	mi le
2 Dat Sg	ți-l	ți-o	ți-i	ți le
3 Dat Sg	i-l	i-o	i-i	i le
1 Dat Pl	ni-l	ne-o	ni-i	ni le
2 Dat Pl	vi-l	v-o	vi-i	vi le
3 Dat Pl	li-l	le-o	li-i	li le

a host different from the verb (as will be discussed in section 5); the example below shows that the clitic can be enclitic on negation:

- (13) nu-l                      dă.  
 neg. cl. (acc-3msg) gives  
 'He doesn't give it.'

In this case the enclitic form *-l* is used and not the proclitic form *îl*. However, if negation is present, as well as two pronominal clitics, the clitics cluster together and not with negation:

- (14) a. \* nu-ne                      îl                      dă.  
 neg. cl.(dat-1pl) cl. (acc-3msg) gives  
 b. nu ni-l                                      dă.  
 neg. cl. (dat-1pl) (acc-3msg) gives  
 'He doesn't give it to us.'

It should be noted that the dative form that surfaces is the one which occurs in combination with another pronominal clitic (*ni*) and not the one that occurs as enclitic (*ne*). The pronominal clitics form a unit also in a configuration in which they occur together with negation and an auxiliary verb:

- (15) a. \* nu-ne                      l-a                      dat.  
 neg. cl.(dat-1pl) cl. (acc-3msg) has given  
 b. nu ni-l                                      a dat.  
 neg. cl. (dat-1pl) (acc-3msg) has given  
 'He hasn't given it to us.'

Evidence that the two pronominal clitics cluster together can be found in the fact that the dative form which surfaces is the one that occurs in combination with another clitic. It is not the case that one clitic is enclitic on negation and the other is proclitic on the verb. The most crucial evidence in favor of the hypothesis that pronominal clitics cluster together is provided by the combination of two clitics which are enclitic on the verb:

- (16) a. \* dă-ne-l  
           give cl.(dat-1pl) cl. (acc-3msg)  
       b. dă-ni-l  
           give cl. (dat) (acc)  
           ‘Give it to us !’

The same combination of clitics which surfaces before the verb occurs also after the verb and not the composition of the enclitic forms. However, the latter is what one would expect if clitics were realized one at a time. To summarize: the morphophonological idiosyncrasies in which Romanian clitics are involved suggest that different classes of clitics should be postulated. In addition, the evidence presented above indicates that the clitics are not realized one at a time, but that they cluster together forming a new unit.

### 3 The analysis

In the previous section I have provided ample motivation in favor of the affixal status of Romanian pronominal clitics. Clitics will not be considered lexical items, i.e. *signs*, but featural information which is provided in the lexicon and used in morphology and phonology for the realization of the cliticized verb form. I will assume that cliticization is a lexical operation which has both a syntactic/semantic effect and a morphophonological one (cf. also Monachesi 1996, Miller and Sag 1997).

The syntactic/semantic contribution of cliticization is reflected in the fact that clitics satisfy the subcategorization requirements of the verb they are an argument of, as shown by examples (8) repeated below:<sup>5</sup>

- (8) a. Maria îi           dă   un bouquet.  
       Maria cl.(dat-3sg) gives a bouquet  
       ‘Maria gives him a bouquet’  
       b. Ion m-a           văzut.  
       Ion cl.(acc-1sg) has seen  
       ‘Ion saw me.’

Within HPSG, a lexical rule can be proposed to achieve this effect; it relates two sets of words.<sup>6</sup>

<sup>5</sup>Romanian clitics can also co-occur with the related complement, as shown by the clitic doubling examples in (7). These cases can receive a straightforward account under the lexical analysis defended here which assimilates clitics to affixes. In the doubling configuration the relevant complement is not removed from the complement list, provided it shares agreement and case information with the clitic (which thus acts as agreement marker) and it satisfies certain pragmatic and semantic conditions.

<sup>6</sup>There are two possible ways in which lexical rules can be interpreted in HPSG: either as meta-descriptions relating *lexical entries* (Calcagno 1995) or as descriptions relating *word objects* (Meurers and Müllen 1997). In this work I will assume the latter formalization. Note that in the lexical rule, the input and output descriptions are connected via “ $\mapsto$ ”, while in the case of the implicative constraints introduced in the following section “ $\rightarrow$ ” is used. In the rule above,  $\circ$  is the shuffle operator defined in Reape (1994).

## (17) Complement Cliticization Lexical Rule (CCLR)

$$\left[ \begin{array}{l} \textit{word} \\ \text{HEAD} \\ \text{VAL} \mid \text{COMPS} \quad \boxed{1} \quad \circ \quad \boxed{2} \\ \text{CLTS} \\ \textit{elist} \end{array} \right] \quad \rightarrow \quad \left[ \begin{array}{l} \text{VAL} \mid \text{COMPS} \quad \boxed{1} \\ \text{CLTS} \quad \boxed{2} \end{array} \right]$$

The rule relates verbs which subcategorize for certain complements to other ones with the same properties except that their subcategorization list is reduced. In other words, the effect of the rule is that those complements which are in the COMPS list are removed and added as members of the CLTS list. This list contains syntactic and semantic information about those complements which will be realized as clitics.

Consider a verb such as *văd* which subcategorizes for one complement:

$$(18) \quad \left[ \text{COMPS} \langle \text{NP}[\textit{acc}] \rangle \right]$$

As shown by the example below, this complement can be realized as a clitic:

$$(19) \quad \begin{array}{l} \textit{te} \quad \quad \textit{văd}. \\ \text{cl. (acc-2sg) see} \\ \text{'I see you.'} \end{array}$$

It is possible to account for this fact given the lexical rule in (17), which relates the description in (18) to the one below:

$$(20) \quad \left[ \text{COMPS} \langle \quad \rangle, \text{CLTS} \langle \text{NP}[\textit{acc}]_{[2\textit{sg}]} \rangle \right]$$

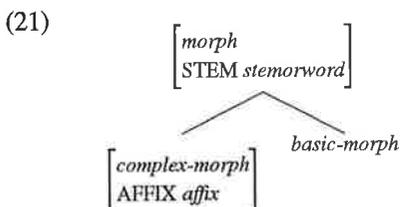
The direct object is removed from the COMPS list and added as member of the CLTS list. The crucial issue is then how this information can be used in order to spell out the cliticized verb form in phonology. In the following sections, I propose specific constraints to relate the information contained in the CLTS list to the actual phonological realization of the clitic.

### 3.1 The signature

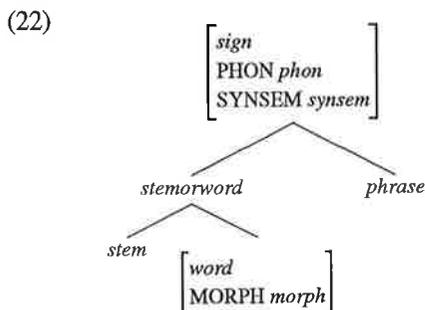
Before presenting the relevant constraints, I will introduce the type hierarchy and the appropriateness conditions assumed (i.e. the signature). In particular, I suggest that the type *word* has MORPH as additional appropriate attribute, with value *morph*. Furthermore, the type *morph* should be partitioned into two subtypes, which are *complex-morph* and *basic-morph*, depending on whether the value contains a stem and affix or just a stem.<sup>7</sup> Therefore, the attribute STEM is defined as appropriate for *morph* and it is inherited by both of its subtypes, while the attribute AFFLIX is appropriate only for *complex-morph*.<sup>8</sup>

<sup>7</sup>The type *complex-morph* is relevant in the treatment of inflection and of morphologically complex forms and it plays a role in the analysis of cliticization developed here.

<sup>8</sup>Similar types have been proposed by Bird and Klein (1994).

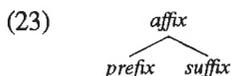


The value of STEM is *stemorword* which is a subtype of *sign*:



In the case of inflection, the attribute STEM will have *stem* as value, while in the case of cliticization its value will be *word*. The same apparatus can be employed to account both for inflection and for cliticization. This is a desirable result, given the similarities between the two. I follow Zwicky (1990) in assuming that stems constitute the input for inflectional processes while cliticization has an inflected word as stem and produces another inflected word. Cliticization constitutes therefore an outer layer of inflectional morphology.

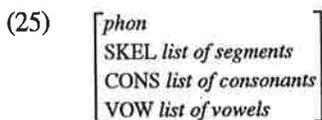
The type *affix* is further partitioned in order to distinguish *prefixes* from *suffixes*:



In addition, it should be mentioned that the only appropriate attribute for *affix* is PHON:



Following Bird and Klein (1994), I will assume that *phon* will have certain appropriate features which are necessary to distinguish the segmental structure:



Since affixes (and pronominal clitics) have only phonological information associated with them, it follows that they are not considered *signs*. This approach shares

insights with realizational approaches to morphology such as those of Anderson (1992) or Stump (1992) that assume that morphemes do not exist as lexical entries, but only as realization of certain morphosyntactic properties of the host.

### 3.2 The realization of the clitics

On the basis of the signature introduced in the previous section, appropriate constraints are formulated that relate the information contained in CLTS to the actual phonological realization of the pronominal clitic. In the case of the example (19), which represents the cliticized verb form *te văd* 'I see you', a constraint like the following will be relevant:<sup>9</sup>

$$(26) \quad \left[ \begin{array}{l} \text{complex-morph} \\ \text{STEM} \mid \text{SS} \mid \text{L} \mid \text{C} \mid \text{CLTS} \langle \text{NP}[\text{acc}]_{2\text{sg}} \rangle \end{array} \right] \rightarrow \left[ \begin{array}{l} \text{AFFIX} \left[ \begin{array}{l} \text{prefix} \\ \text{PHON} \mid \text{SKEL} \langle \text{te} \rangle \end{array} \right] \end{array} \right]$$

The constraint says that if there is a STEM with a CLTS list with an element which is a second person, singular, accusative NP, then it must be a pronominal clitic whose phonological realization is *te*. Similar constraints can be expressed in order to account for the different classes of clitics shown in the table in (1). Therefore if the host starts with a vowel and it is not an auxiliary, both the clitics from class 1 and those from class 2 can precede it, as can be seen from the examples in (9) repeated below:

- (9) a.    *mă*            *așteaptă*.  
           cl.(acc-1sg) waits  
           'He waits for me.'
- b.    *m-așteaptă*.  
           cl.(acc-1sg) waits

A constraint like the following can account for this fact:

<sup>9</sup>It should be mentioned that within HPSG, constraints determine which linguistic objects are grammatical and which are not. Implicative constraints like the ones presented in this section are interpreted as stating that every feature structure which satisfies the antecedent will also have to satisfy the consequent.

$$(27) \left[ \begin{array}{c} \text{complex-morph} \\ \left[ \begin{array}{c} \text{word} \\ \text{STEM} \left[ \begin{array}{c} \text{PHON} \left[ \begin{array}{c} \text{VOW} \langle \text{[1]} a \vee o \rangle \oplus \text{[2]} \\ \text{SKEL} \langle \text{[1]} \rangle \oplus \text{[3]} \end{array} \right] \\ \text{SS | L | C} \left[ \begin{array}{c} \text{HEAD verb [AUX -]} \\ \text{CLTS} \langle \text{NP[acc]}_{1\text{Sg}} \rangle \end{array} \right] \end{array} \right] \\ \left[ \begin{array}{c} \text{AFFIX} \left[ \begin{array}{c} \text{prefix} \\ \text{PHON | SKEL} \langle m\check{a} \rangle \vee \langle m \rangle \end{array} \right] \end{array} \right] \end{array} \right] \rightarrow$$

It says that if there is a lexical verb which begins either with the vowel *a* or *o* and it has one element in the CLTS list which is the first singular accusative NP, this can be realized either as the clitic *mă* or as *m*. A similar constraint can be formulated to account for the fact that if the verb begins with a consonant or a vowel which is different from *a* or *o*, only the clitic *mă* is possible, as shown by the following example:

- (28) a. *mă vede.*  
 cl.(acc-1sg) sees  
 'He sees me.'
- b. \**m vede.*  
 cl.(acc-1sg) sees

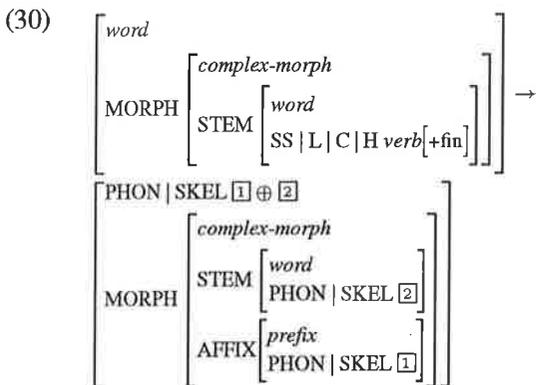
It should be noted that the combination of two pronominal clitics can be expressed by means of the same mechanism. The following is an example of the realization of the dative, first person plural clitic and the accusative, masculine, third person plural clitic:

$$(29) \left[ \begin{array}{c} \text{complex-morph} \\ \left[ \begin{array}{c} \text{STEM | SS | L | C | CLTS} \langle \text{NP[dat]}_{1\text{Pl}}, \text{NP[acc]}_{3\text{MPl}} \rangle \\ \left[ \begin{array}{c} \text{AFFIX} \left[ \begin{array}{c} \text{affix} \\ \text{PHON | SKEL} \langle ni-i \rangle \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \rightarrow$$

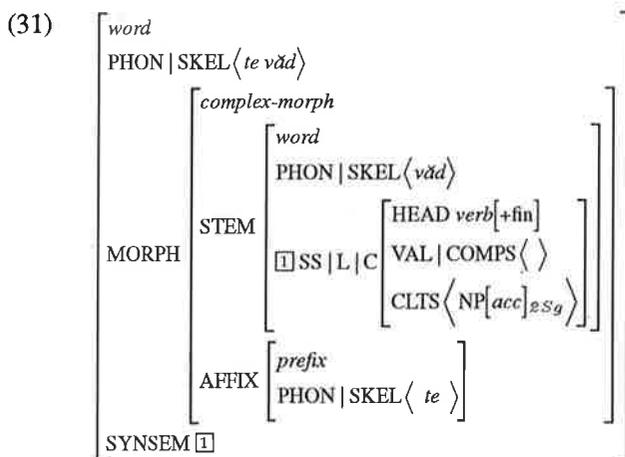
Recall that in this case it is not the form *ne* that surfaces, but the exceptional form *ni*. Given the present approach, it is quite easy to account for this peculiarity. The information about the two elements is available at the same time and the combination of two clitics is treated just like a new form.<sup>10</sup>

<sup>10</sup>As already mentioned in section 2.6, an approach that realizes the clitics one at a time would run into a series of problems. For example it should justify why forms such as *ni*, *vi* and *li* cannot occur alone and must be realized in combination with another clitic (except with the clitic *o*).

Certain generalizations can be expressed within this system, e.g. in the ordering constraints that regulate the enclitic or proclitic position of the clitic with respect to the host. In particular, a constraint like the following states that clitics should precede finite verbs. This is due to the fact that the phonology of the clitic indicated by tag [1] should precede that of the verb, indicated by tag [2]:



Therefore, it is through the interaction of the lexical rule and the constraints in (26) and (30) that a cliticized verb form like *te văd* 'I see you' is licensed:



The description above states that the verb has an empty COMPS list, while the information about the direct object is contained in the CLTS feature which triggers its spell-out as the clitic *te*. Since the verb *văd* is a finite one, the clitic will precede it.

#### 4 The status of auxiliaries and clitic climbing

The analysis I have sketched accounts for the linearization of pronominal clitics in the presence of verbs in simple tenses. However, if an auxiliary verb is present, the

clitic does not combine with the verb that subcategorizes for it, but must attach to the auxiliary, as shown in the following example:

- (32) le-am văzut.  
 cl.(acc-3fpl) have seen  
 'I have seen them.'

In Romanian, auxiliaries can be found in the present perfect, conditional and future paradigms. In the case of the present perfect, the auxiliary combines with a past participle, in the other two cases it combines with a bare infinitive (e.g. an infinitive without the particle *a*). The following table summarizes the different forms of auxiliaries:

(33)

	1 Sg	2 Sg	3 Sg	1 Pl	2 Pl	3 Pl	
Aux Psp.	am	ai	a	am	ați	au	past participle
Aux Cond.	aș	ai	ar	am	ați	ar	bare infinitive
Aux Fut.	voi	vei	va	vom	veți	vor	bare infinitive

I suggest that the status of auxiliaries in Romanian is similar to that of auxiliaries in other Romance languages. They are words and it is for this reason that they can function as hosts for pronominal clitics. I will show that an analysis in terms of argument composition can account for the linearization of pronominal clitics in (32) (cf. also Monachesi (1996) for Italian and Miller and Sag (1997) for French).

Argument composition is a lexical operation according to which the auxiliary inherits the complements of the embedded verb, including those which might be realized as clitics. In the case of the auxiliary *am*, the following lexical entry will be associated with it:

- (34)
- $$\left[ \begin{array}{l} \text{PHON | SKEL} \langle am \rangle \\ \left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} verb \\ AUX + \end{array} \right] \\ \text{SUBJ} \langle \boxed{1} NP \rangle \\ \text{VAL} \left[ \begin{array}{l} \text{COMPS} \langle \left[ \begin{array}{l} w-ss \\ \text{HEAD} verb \\ \text{VAL} \left[ \begin{array}{l} \text{SUBJ} \langle \boxed{1} NP \rangle \\ \text{COMPS} \boxed{2} \end{array} \right] \end{array} \right] \rangle \oplus \boxed{2} \end{array} \right] \end{array} \right] \end{array} \right]$$

In the description above, the complements of the auxiliary verb are identified with those of the embedded verb, through the operation of argument composition. This is indicated by the presence of the tag  $\boxed{2}$  both in the COMPS list of the auxiliary verb and in that of the embedded verb. The auxiliary verb will thus inherit the complements of the past participle which can be realized as pronominal clitics along the lines of the analysis sketched in the previous sections. So, the two verbs act as a unit, as far as clitic placement is concerned. Given a sentence like the following:

- (35) le-am văzut.  
cl.(acc-3fpl) have seen  
'I have seen them.'

the auxiliary subcategorizes for the past participle and its complements; in this case the direct object NP which will be eventually realized as clitic:

- (36)  $\left[ \text{COMPS} \left\langle v \left[ \text{COMPS} \langle \square \rangle \right], \square \text{NP}[\text{acc}]_{[3fpl]} \right\rangle \right]$

The Complement Cliticization Lexical Rule in (17) can then apply to license cliticized verbs. Its effect is that the relevant complement will be removed from the COMPS list and added as value of the CLTS list:

- (37)  $\left[ \text{COMPS} \left\langle v \left[ \text{COMPS} \langle \square \rangle \right] \right\rangle, \text{CLTS} \langle \square \text{NP}[\text{acc}]_{[3fpl]} \rangle \right]$

A constraint like the following will then be responsible for the spell out of the cliticized verb form:

- (38)  $\left[ \text{complex-morph} \left[ \text{STEM} | \text{SS} | \text{L} | \text{C} | \text{CLTS} \langle \text{NP}[\text{acc}]_{3fpl} \rangle \right] \rightarrow \left[ \text{AFFIX} \left[ \begin{array}{l} \text{prefix} \\ \text{PHON} | \text{SKEL} \langle le \rangle \end{array} \right] \right] \right]$

It states that if the verb has an NP accusative, third person plural feminine in its CLTS list, it will be realized as the clitic *le*.

As already mentioned, clitic climbing is obligatory with auxiliary verbs. However, the third person feminine singular clitic *o* constitutes an exception since it must attach to the past participle and cannot precede the auxiliary.<sup>11</sup>

- (39) a. am văzut-o.  
have seen cl.(acc-3fsg)  
'I have seen her.'
- b. \*o-am văzut.  
cl.(acc-3fsg) have seen

This is the case also for the conditional. The clitic must attach to the bare infinitive and not to the auxiliary:

- (40) a. aş 'vedea-o.  
would see cl.(acc-3fsg)  
'I would see her.'

<sup>11</sup>This is an idiosyncratic property of Romanian which is not present in other Romance languages. For example, it does not seem possible to assimilate this peculiarity to other properties of the language. It does not seem plausible to attribute the postverbal position of *o* to the fact that it is a feminine clitic since other feminine clitics occur in front of the auxiliary. It should be noted that in old Romanian the clitic *o* could occur before the auxiliary.

- b. \* o-aş                      vedea.  
 cl.(acc-3fsg) would see

In order to account for the obligatoriness of clitic climbing in the presence of Romanian auxiliaries, past participles (and bare infinitives) should be excluded as possible hosts for pronominal clitics.<sup>12</sup> Given the constraints introduced in the previous section for the spell out of the cliticized verb form, it is quite straightforward to add an additional condition which will produce the desired results. The constraint (38) should be revised in the following way:

$$(41) \quad \left[ \begin{array}{l} \text{complex-morph} \\ \text{STEM} \mid \text{SS} \mid \text{L} \mid \text{C} \left[ \text{CLTS} \langle \text{NP}[\text{acc}]_{3fpl} \rangle \right] \right] \rightarrow \\ \left[ \begin{array}{l} \text{word} \\ \text{STEM} \left[ \text{SS} \mid \text{L} \mid \text{C} \left[ \text{HEAD} \left[ \text{VFORM} [\neg (\text{psp} \vee \text{bare-inf})] \right] \right] \right] \right] \\ \text{AFFIX} \left[ \begin{array}{l} \text{prefix} \\ \text{PHON} \mid \text{SKEL} \langle \text{le} \rangle \end{array} \right] \end{array} \right]$$

The additional condition says that the VFORM of the host must be different from past participle or from bare infinitive. In this way it is possible to account for obligatory clitic climbing in the general case. On the contrary, the clitic *o* cannot climb and must combine with a past participle (and a bare infinitive), as shown by the examples (39a) and (40a). These cases can be accounted for if no additional condition on the VFORM of the host (e.g.  $VFORM \neg (\text{psp} \vee \text{bare} - \text{inf})$ ) is added in the constraint that is responsible for the realization of the clitic *o*. As for the ill-formedness of the examples in (39b) and (40b) in which the clitic *o* has climbed, they will receive a phonological explanation. The clitic *o* cannot precede an auxiliary which begins with a vowel: a phonological constraint should be postulated to exclude this possibility.

It should be noted that in the case of the future paradigm, the auxiliary begins with a consonant and the clitic *o* can occur either after the infinitive (42a) or in front of the auxiliary (42b):

- (42) a.    voi vedea-o.  
           will see cl.(acc-3fsg)  
           'I will see her.'
- b.    o            voi vedea.  
       cl.(acc-3fsg) will see

Additional evidence for the phonological explanation is provided by the fact that the clitic *o* precedes the modal verb *a putea* 'can':

<sup>12</sup>Romanian behaves like French in this respect. Miller and Sag (1997) suggest that in French past participles cannot host clitics.

- (43) o pot vedea.  
cl.(acc-3fsg) can see  
'I can see her.'

In the case above, the modal begins with a consonant and the pronominal clitic must attach to it, like all the other object clitics.<sup>13</sup>

## 5 Speculations on enclitic placement

Romanian clitics have a peculiar property which sets them apart from other Romance counterparts. They can *optionally* attach to elements which are different from their morphosyntactic host (i.e. the verb). They can combine with negation (44a), complementizers (44b), nouns (44c) or wh-elements (44d):

- (44) a. Mihai nu-l așteapta.  
Michael neg cl.(acc-3msg) waits for  
'Michael doesn't wait for him.'
- b. Cred că-l vede.  
believe that cl.(acc-3msg) sees  
'I believe that he sees him.'
- c. Maria-i scrie des.  
Maria cl.(dat-3sg) writes frequently  
'Maria writes him frequently.'
- d. Unde-mi lasa fratele tau cheia.  
Where cl.(dat-1sg) leaves brother yours key the  
'Where does your brother leave me the key.'

Given the view argued for in this paper, which considers clitics not as *signs*, but as featural information encoded in the CLTS feature, it seems possible to provide an analysis for the data above.<sup>14</sup>

I will briefly sketch what a possible direction of explanation could be. The lexical rule in (17) accounts for the fact that even though the clitic attaches phonologically to a host different from the verb, it still satisfies its subcategorization requirements. The information about those complements which should be realized as clitics is thus encoded in the CLTS feature. A specific principle can be formulated

<sup>13</sup>It should be noted that if the clitic *o* can combine with a bare infinitive, one would expect a form such as \**pot vedea-o* 'I can see her', which is ungrammatical. I assume an analysis of *a putea* along the lines of that proposed for Italian restructuring verbs (Monachesi 1998b). The modal verb is therefore required to subcategorize for an infinitival which has not combined with clitics (i.e. a verb whose CLTS list is empty). I refer to Monachesi (1998a) for further details.

<sup>14</sup>These data might argue against the affixal view of clitics defended in this paper. According to one of the tests proposed by Zwicky and Pullum (1983) affixes should be selective with respect to the host they attach to. However, as discussed in Miller (1992a) the relevance of this test is questionable under the view that elements which show 'promiscuous attachment' such as possessive 's in English should be considered (phrasal) affixes (cf. Zwicky 1987).

to account for the percolation of the relevant information on the appropriate host. The morphophonological constraints will account for the realization of the correct clitic form. It should be noted that in the examples above, it is the enclitic form of the clitic that surfaces. Therefore, they cannot be analysed in terms of phonological restructuring since in that case, the clitic form which occurs before the verb would be expected.

## 6 Conclusion

In this paper I have provided motivations in favor of the affixal status of Romanian pronominal clitics. I have shown that cliticization is a lexical operation which has both a syntactic/semantic effect and a morphophonological one. In order to account for the morphophonology of Romanian clitics, specific constraints have been introduced which are able to deal with their idiosyncratic behavior. I have shown that the morphophonological analysis proposed can provide an answer to a peculiar property of Romanian clitic climbing: the fact that the clitic *o*, unlike the other clitics, does not climb in the presence of an auxiliary and attaches to the embedded verb. The analysis presented provides a comprehensive account of Romanian cliticization since it deals not only with the syntactic properties of Romanian clitics, but also with the morphophonological ones. In this sense, it is superior to previous accounts of Romanian clitics such as that of Dobrovie-Sorin (1994) which does not deal with the kind of morphophonological idiosyncrasies discussed in this paper. Furthermore, the analysis proposed here shares some crucial insights with that assumed for Italian clitics in Monachesi (1996) and for French clitics in Miller and Sag (1997). The same lexical approach can deal with the variation presented by different Romance data, which is a desirable result.

## Acknowledgements

I would like to thank the audience of the HPSG seminar in Tübingen, in particular Tilman Höhle and the audience of the EUROLAN summer school in Tusnad, especially Ana Maria Barbu and Emil Ionescu for comments and suggestions. Many thanks to all the people who helped me with the data: Agnes Bende Farkas, Edward Goebbel, Amalia Todiraşcu and Diana Zaiu. This work was partially supported by a grant from the *Netherlands Organization for Scientific Research (NWO)*.

## References

- Anderson, S. (1994), *A-Morphous Morphology*, Cambridge University Press, Cambridge.
- Avram, A. (1986), Sandhi phenomena in Romanian, in: H. Andersen (ed.) *Sandhi phenomena in the languages of Europe. Trends in Linguistics, Studies and Monographs 33*, Mouton de Gruyter, Berlin.
- Bird, S. and E. Klein (1994), Phonological analysis in typed features systems, *Computational Linguistics* 20.3.

- Bonet, E. (1991), *Morphology after Syntax: pronominal clitics in Romance*, MIT, PhD Thesis.
- Burzio, L. (1986), *Italian Syntax*, Reidel, Dordrecht.
- Calcagno, M. (1995), Interpreting lexical rules, *Proceedings of the Conference on Formal Grammar*, Barcelona.
- Dobrovie-Sorin (1994), *The syntax of Romanian. Comparative Studies in Romance*, Mouton De Gruyter, Berlin.
- Farkas, D. and K. Kazazis (1980), Clitic pronouns and topicality in Rumanian, *Proceedings of the 16th Regional Meeting of the Chicago Linguistic Society*, Chicago.
- Kayne, R. (1975), *French Syntax: the transformational cycle*, MIT Press, Cambridge.
- Meurers, D. and G. Minnen (1997), Computational Treatment of Lexical Rules in HPSG as Covariation in Lexical Entries, *Computational Linguistics* 24.4.
- Miller, P. (1992a), *Clitics and Constituents in Phrase Structure Grammar*, Garland, New York.
- Miller, P. (1992b), Postlexical cliticization vs. affixation: coordination criteria, in: C. Canakis, G. Chan and J. Denton (eds.), *Proceedings of the 30th Regional Meeting of the Chicago Linguistic Society*, Chicago.
- Miller, P. and I. Sag (1997), French clitic movement without clitics or movement, *Natural Language and Linguistic Theory*.
- Monachesi, P. (1996), *A grammar of Italian clitics*, Tilburg University, PhD Thesis.
- Monachesi, P. (1998a), On certain properties of Romanian auxiliary (and modal) verbs, *Proceedings of the conference on Formal, Head-driven and Categorical Grammar*, Saerbruecken.
- Monachesi, P. (1998b), Italian restructuring verbs: a lexical analysis, in E. Hinrichs, A. Kathol and T. Nakazawa (eds), *Complex Predicates in Nonderivational Syntax. Syntax and Semantics, vol 30*, Academic Press, San Diego.
- Pollard, C. and I. Sag (1987), *Information-based Syntax and Semantics. Fundamentals. Vol. 1*, CSLI, Stanford.
- Pollard, C. and I. Sag (1994), *Head-Driven Phrase Structure Grammar*, University of Chicago Press and CSLI, Stanford.
- 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*, CSLI Publications, Stanford.
- Rizzi, L. (1982), *Issues in Italian Syntax*, Foris, Dordrecht.
- Sportiche, D. (1993), *Clitic constructions*, ms. UCLA.
- Stump, G. (1992), On the theoretical status of position class restrictions on inflectional affixes, in G. Booij and J. v. Marle (eds.), *Yearbook of Morphology*, Kluwer Academic Press, Dordrecht.
- Zwicky, A. (1985), Clitics and Particles, *Language* 61.2.
- Zwicky, A. (1987), Suppressing the Zs, *Journal of Linguistics* 23.
- Zwicky, A. (1990), Inflectional morphology as a (sub)component of grammar, in W.D. et al. (eds.), *Contemporary Morphology. Trends in Linguistics, Studies and Monographs* 49, Mouton De Gruyter, Berlin.

Zwicky, A. and G. Pullum (1983), Cliticization vs. inflection: English n't, *Language* 59.3.