

The Syntax of Romanian Clitics

Louis-Harry Desouvrey

February 2019

First draft

louish8@gmail.com / louishd@outlook.com

This paper presents a syntactic account of Romanian clitic doubling and clitic clusters. It is shown that depending on feature specification of the argument, the direct object marker *pe* can behave either as a preposition or a case marker. If it is a preposition, the NP must be doubled by a clitic in order to satisfy the argument structure of the verb. If it is a case marker to the NP, the latter absorbs the case of the verb and satisfies its argument structure as well, hence precluding clitic doubling. As for clitic clusters, a sequence of two clitics must be compounded to form a new edge-bound element that can move under the OCP, without overloading the derivation. There are three types of compound rules, each of which operates in tandem with a series of phonological rules that alter their shapes. Ultimately these sandhi rules obliterate the boundary between the clitics, yielding a new word-like element. This fine-grained account, which relies heavily on grammatical features such as case, reference, π (tier-bound), ω (vector), and ψ (a series of traits like animate, definite, specific), runs like a clockwork. It makes it possible to predict all known peculiarities of clitic combinations, including the rigid dative-accusative ordering, the special behavior of the third feminine singular clitic, the lack of PCC effect, etc.

Keywords: Clitic doubling, clitic clusters, *pe*-marking, compound, OCP, features, movement, PCC, syntax, coreference, phonology.

1. Introduction

Romanian possesses a very complex pronominal system whose elements exhibit a great deal of surface forms. This leads to various hypotheses about their nature. In most analyses, they are taken to be phrasal affixes. In this paper, I show that Romanian clitics are syntactic elements, just like their French and Italian counterparts. Their behavior follows from the wealth of features they are specified for, unlike the NPs they replace. I show that most sequences of Romanian clitics are derived from at least

three types of morphosyntactic compounds which are sealed by a number of sandhi rules. Under this feature-based theory, all peculiarities result from the feature interaction and the few universal principles that regulate them. I proceed as follows. In the next section I briefly present the method and the principles which underlie the proposed analysis. Then in section 3, the feature structure of Romanian clitics and verbs are analyzed as well as the nature of the direct object marker and the clitic doubling it triggers under certain conditions. In section 4 and 5, I unravel every types of clusters, and then I conclude the article.

2. Theoretical assumptions

The analysis to be proposed here is embedded into the theory of features and constraints put to the forth in Desouvrey (2000) and further developed in subsequent works. In this theory, language variation is a matter of variation in feature specification across a paradigm. The interaction of features is ruled by unranked universal constraints, and not by parameters. In natural languages, parameters exist but in general they can explain nothing. For instance, the fact that English, unlike Romance, does not have clitic personal pronouns is due to different feature specification in their respective paradigm. However, the fact that Romance verbs are heavily conjugated, unlike Chinese verbs for instance, can be roughly attributed to some parameter.

This theory relies on features borne by lexical elements in general. So far the features that have the most considerable linguistic interest are cases (Nominative [N], Accusative [A], and Oblique [O]), $[\pi]$ (timing tier) and $[\omega]$ (scope), as well as referential features, noted with capital letters representing the virtualization of any real or imaginary world entity in the grammar. Syntactic features make of a tree and are laid out in a similar way to phonological features in non-linear phonology. Let us illustrate this with simple examples from French, where the pronominal paradigm shows the less complexities among Romance languages. Consider the derivation of the sentences in (1). It proceeds from a simple input structure, which reflects simple constituency analysis, as shown in (2) (French is an SVO language). The verb enters the derivation fully inflected for person and number and in addition it is specified for Nominative and Accusative; as for the pronoun *le*, it is specified for Accusative as well as person, number, gender, etc., which are set aside in this representation.

- (1) Marie le veut.
 Marie it wants
 'Marie wants it.'

- (2) a.
- ```

 N
 / \
 [Marie] [veut le]]
 | |
 A A

```
- b.
- ```

      N
     / \
  [Marie] [veut le]]
     |   |
     le  |
     |   |
     A   A
  
```
- c. Marie=le veut (Output)

What happens next to this input is due to the properties of the representation. The subject *Marie* does not bear any grammatical case, so it is assigned N case by the verb, a spreading process represented by a dash line. However, since the object pronoun bears A case, it correctly pairs up with the relevant case of the verb. One can say that the pronoun absorbs the case of the verb which in turn licenses the pronoun. Now a grammatical constraint, which I take to be the syntactic analog of the Obligatory Contour Principle (3), prohibits the occurrence of two identical case features in this syntactic domain, VP. The pronoun moves to a parallel timing tier, from where it attaches via an association line to the first element out of the VP domain, here the subject *Marie*. This attachment takes place at the skeletal level so that the rightmost segment of [mari] is linked up with the leftmost segment of [lə]. The Linearization Convention (4) ensures its proper spell out.¹

(3) Obligatory Contour Principle (OCP)

- a) Case-specified arguments must exit the complement domain of the verb (OCP-1).
- b) No element can intervene between two elements with identical features (OCP-2).

¹ As a result, the adjoined element becomes a phrasal affix, to use a familiar parlance in the clitic literature. This type of adjunction, which always puts the adjunct to the right of the host, is obviously superior to the traditional one where the adjunct can be either to the left or to the right of the host, according to the will of the analyst. Note that the linearized host-adjunct sequence is attached with an equal sign, as in (2c).

(4) **Linearization convention**

Embed the adjunct tier into the main tier and to the right of the host element.

Consider now sentence (5), in which the object argument is an interrogative pronoun. Just like the preceding example, an OCP effect takes place, triggering the movement of the interrogative from the input (6a). Unlike *le*, *que* is specified for the features $[\pi]$ and $[\omega]$. The latter related to scope, which is irrelevant in this example.² An element bearing the feature $[\pi]$ is tier-bound, i.e., it cannot adjoin to another element in the structure, which would put it in a different tier as the head. Instead, it has to make a long movement to the left edge of the structure, as seen in (6b). This movement to an edge is referred to as outbound movement, as opposed to inbound movement of elements not specified for $[\pi]$. Now, the subject comes to be sandwiched between two case-specified elements, it is squeezed out of its normal position, adjoining to the verb, as seen in (6c). The constraint that ensures this outcome is referred to as OCP-2 (3b), which is analogous to the gemination effect in phonology. Thus, it appears that clitics must move from the verbal constituency under OCP-1, but they must still be adjacent to the verb by OCP-2.

- (5) Que voulez-vous?
 What want.2pl you
 What do you want?

- (6) a. $\begin{array}{cc} N & N \\ | & | \\ [vous & [voulez & que]] \\ & | & | \\ & A & A \end{array}$
- b. $\begin{array}{cc} & N & N \\ & | & | \\ Que & [vous & voulez] \\ | & & | \\ A & & A \end{array}$
- c. Que voulez=you?

2 A scopal element must be understood as an element bearing the feature $[\omega]$. The effect of this feature, namely a superiority effect, can only be seen in contexts where there are at least two scopal elements. Thus in (6) it has no effect, the *wh*-movement being triggered by case, not by scope.

Under this theory, clitic pronouns and *wh*-elements are subject to the same constraints which act on their feature structure. The features $[\pi]$ and $[\omega]$ are frequently found in *wh*-elements (and negation) across languages, but it is an empirical question as to whether or not *wh*-elements in some language can exist without these features. In any event, these features are not proper to *wh*-elements and negation, since they are found in personal pronoun paradigms across Romance. Indeed, as I will show, the feature $[\pi]$ is a key player in Romanian clitic system, just like in Italian and Portuguese (cf. Desouvrey 2018, 2008b).

3. The features of Romanian pronominal clitics

Romanian pronominal system is very complex although it counts less clitics than French and Italian, for instance. The complexities arise from the various sandhi rules that alter their shapes, as I will show. An overview of the system is given in Table 1.

Table 1

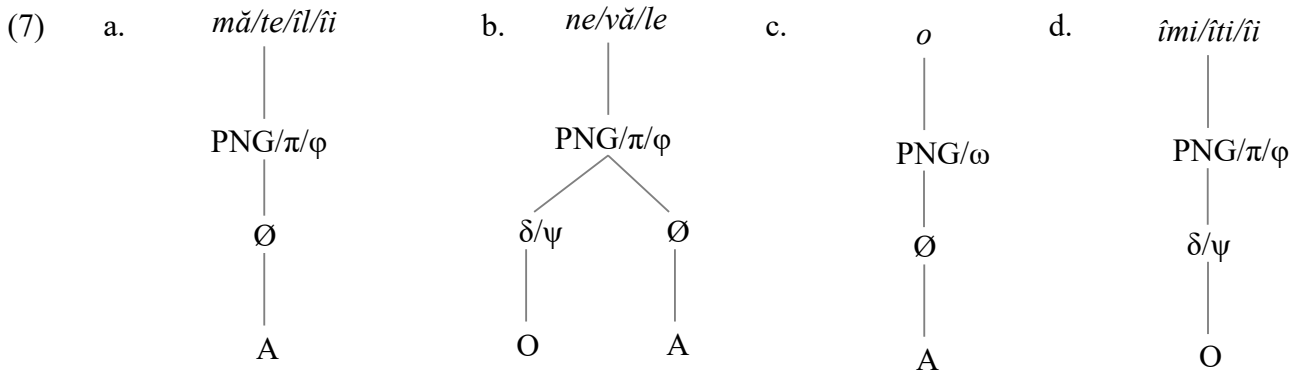
		Unstressed (full)			Stressed	
Number	Person	Subj	DO	IO	DO	IO
Sing.	1 st	eu	mă [A, π]	îmi [ψ, δ(O)]	mine	mie
	2 nd	tu	te [A, π]	iți [ψ, δ(O)]	tine	tie
	3 rd m.	el	îl [A, π]	îi [ψ, δ(O), π]	el	lui
	3 rd f.	ea	o [A, ω]		ea	ei
Plural	1 st	noi	ne [ψ, A, δ(O), π]		noi	nouă
	2 nd	voi	vă [ψ, A, δ(O), π]		voi	vouă
	3 rd m.	ei	îi [A, π]	le	ei	lor
	3 f.	ele	le [ψ, A, δ(O), π]		ele	lor
3 rd reflexive			se [A]	își [ψ, δ(O), π]	sine	sieși

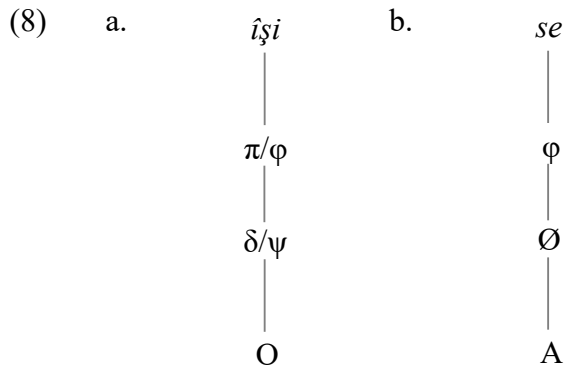
The clitics to be discussed are shown in their full forms. For ease of reference, their set of hypothesized features, to be discussed throughout, are shown in brackets, which include case (A and O), π , ω , and δ . The latter is a thematic feature (dative), a class node to which the oblique case is attached, hence δ (O) (see below). Similarly, Accusative (as well as Nominative) are dependent of an unspecified (dummy) class node, \emptyset (A), although this is not shown in the Table. In addition, Romanian IO clitics includes a new-found feature, namely $[\psi]$, which is a variable collection of traits such as

definiteness, specificity, humanness or perhaps animacy.

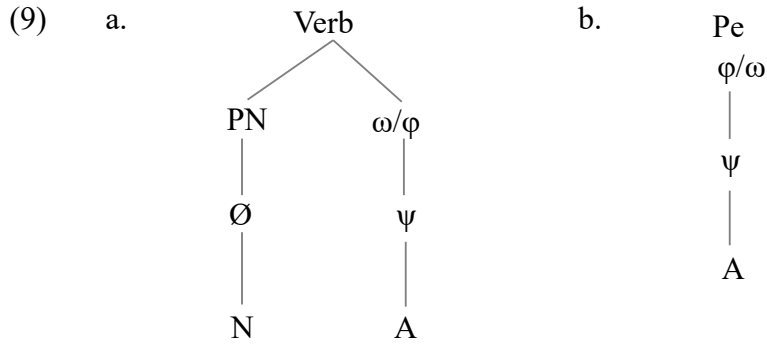
Setting aside the subject and the object stressed forms, the object clitics exhibit a number of syncretism. Specifically 1st and 2nd person plural as well as 3rd feminine plural have the same shape for both DO and IO. As a matter of fact, they are specified for both Oblique and Accusative. However, even though 3rd singular IO and 3rd plural masculine DO are identical in shape, they have to be considered as two distinct elements with respect to feature specification (see below). Unlike their plural counterparts, 1st and 2nd person singular clitics as well as 3rd masculine *îl* and 3rd feminine *o* are each specified for a single case, Accusative (DO) or Oblique (IO). Also, the clitic *o* is specified for ω (vector), while all other clitics are ϕ -specified (scalar). It is thus expected to present certain special behaviors (see below). Finally in this paradigm, all clitics but reflexive *se* and 3rd feminine *o* are π -specified, which makes them incompatible with inbound movement, that is, they have to move to the left edge of the structure (outbound movement).

The features shown in Table 1 are hierarchically organized into tree structures. All clitics in Table 1 can be derived from either of the tree structures in (7). Each clitic tree comprises a root node which includes person, number, and gender [PNG], [π], [ϕ]/or [ω]. The root node develops into one or more thematic nodes to which the case features are linked. One can thus obtain, by fixing the value for PNG, DO clitics *mă*, *te*, *îl*, and *îi* from (7a), IO *îmi*, *îti*, and *îi* from (7d), and the plural clitics *ne*, *vă*, and *le* from (7b). The tree in (7c) represents a single clitic namely *o*. As for the reflexive clitics, their trees are shown in (8).





At this point, it should be recalled that these personal pronouns are clitics because they are heavily specified for syntactic features, the most prominent of which is case, as opposed to morphological features like PNG. Cliticization is an interaction between two elements, the clitic and its host. In my perspective, the host is the element that selects the clitic, namely the verb. To unravel the cliticization process, a clear picture of verb features is needed. Clitics and their hosts must share the same features, although the overall sets of features of both obviously may not be the same. Identical features of both elements must pair up to make up a tier, while different features clash and yield the rejection of the whole structure. Feature clashes are generally avoided by the hierarchical structure of the feature tree, whereby features under similar class nodes interact locally. The feature structure of two-argument verbs includes two branches generally referred to as Agr, one for each argument, as shown below in (9a). Actually, AgrS is the projection of person and number feature (PN), while AgrOs (DO and IO) may include either φ , ω or a dummy feature. This intends to show that the subject agrees with the verb in person and number, but not the object. Nominative is a dependent of PN via the intermediate unspecified thematic node, whereas Accusative is a dependent of φ via the ψ node. The latter requires that certain complements (human, specific) be introduced by what is referred to as a direct object marker (DOM), namely *pe*. It must be the case that *pe* is a supplier of ψ , which is absent in nominal and accusative clitics, as shown in (9b). It must be the case that it includes in its root node a φ - and an ω -node, since it is compatible with both vector and scalar elements. However, it appears that the feature ψ is a dependent of the φ -node, while the ω -node is bare. The interplay between various types of argument and this feature, specified in DO verbs, but open to the speaker's interpretation on which argument requires it, is arguably what gives rise to various nuances of meaning, which are set aside here (on this, see Chiriacescu and von Heusinger 2010, Hill ????, etc.).



3.1 Clitic doubling

Clitic doubling is closely linked with *pe*-marking. What is interesting in Romanian is that not all *pe*-marked arguments need to be doubled with a clitic. Given the above assumptions, I show that the DOM may in fact function either as a preposition or a case marker, depending of certain features of the NP. The crucial difference to be made here is that the case marker becomes a part of the NP which satisfies the argument structure of the verb, while the preposition is autonomous of the complement to which it assigns case by spreading. Thus, the *pe*-phrase, as a PP, is not consistent with the argument structure of the verb which must be rescued by an argumental clitic. Put another way, the case of the preposition is not available for the verb, since it is absorbed by its complement.³

The DOM functions as a preposition when its complement is a ϕ - or a \emptyset -specified element, i.e., a scalar or a neutral element. Let us assume that unstressed pronouns are ϕ -specified, while NPs are \emptyset -specified. Thus in (10a), *pe* perfectly aligns with its ϕ -complement to which it assigns the feature ψ and by the same token accusative case. Both the head and the complement being ϕ -specified, there is no room for optionality: a relevant clitic must then be used to absorb the case of the verb and to satisfy its argument structure. In (10b) however, the NP is neutral with a \emptyset -node, which means it can pair up with either node of *pe*. If it pairs up with ϕ -*pe*, it gets the feature ψ and A, and hence clitic doubling is obligatory. If instead it pairs up with the bare ω -*pe*, it gets no feature at all, which precludes clitic doubling, (10c).⁴

3 The distinction between case markers and preposition is the core of Jaeggli (1982) account of clitic doubling in Spanish. In the present theory, this follows from the very nature of the representation, whereas it has to be stipulated on the GB apparatus.

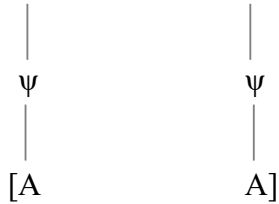
4 Regarding sentences like (10c), which has a proper name, Cojocaru points out: "For nouns designating defined / identified persons the anticipation [i.e., clitic doubling] is considered optional. However, in contemporary standard Romanian there is a strong tendency towards the anticipation, and sentences without the anticipation of this direct object are perceived as archaic or incorrect." (p.100). Thus a structure like (10c) would be perfectly acceptable with a common noun instead of a proper noun.

- (10) a. El ne-a așteptat pe noi. (Cojocaru 2003: 100)⁵
- | | | | | |
|---|----|----|------|----|
| ϕ | ϕ | ϕ | ω/[ϕ | ϕ] |
| | | | | |
| | ∅ | ψ | ψ | ψ |
| | | | | |
| | [A | A] | A | |
- he 1pl-has waited pe us
'he waited for us.'
- b. Am așteptat-o pe Maria o oră.
- | | | | |
|---|----|------|----|
| ϕ | ϕ | ω/[ϕ | ∅] |
| | | | |
| | ψ | ∅ | ψ |
| | | | |
| | [A | A] | A |
- c. Am așteptat pe Maria o oră. (archaic or incorrect)
- | | | |
|---|------|----|
| ϕ | ϕ/[ω | ∅] |
| | | |
| | ψ | ψ |
| | | |
| | [A | A] |
- have.1sg waited for-3sg.fem Maria one hour.
'I waited for Maria for one hour.'

Let us now consider the case where *pe* is obligatorily construed as a case marker, which sharply opposes (10a). This essentially occurs when the complement is ω-specified (vectors). Indeed, *wh*-operators, which are genuine vectors, are incompatible with clitic doubling. As can be seen in (11a), *pe* cannot spread its feature ψ to the operator, since the latter does not have a ϕ-node. The Accusative case is not assigned, resulting in the reanalysis of *pe* as a case marker. Thus the *pe*-phrase absorbs the case of the verb, excluding the use of a clitic. Since indefinite quantifiers cannot be clitic doubled, it must be the case that they are vectors in Romanian. The same analysis holds, as seen in (11b).

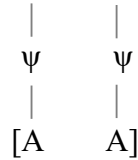
⁵ In the examples borrowed from Cojocaru (2003), the glosses are mine.

- (11) a. Pe cine ai invitat la cină? (Cojocaru: 88)
 $\varphi/[\omega]$ ω φ/ω



Pe who have.2sg invited to dinner
 'Whom did you invite to dinner?'

- b. Ai invitat pe cineva la cină?
 φ $\varphi/[\omega]$ ω



have.2sg pe anyone to dinner
 'Did you invite anyone to dinner?'

If the argument gets the feature ψ from the definite article, neither clitic doubling nor *pe*-marking is necessary, as illustrated in (12) (cf. Chiriacescu and von Heusinger 2010).

- (12) Doctorul examinează băiatul.
 φ \emptyset
 ψ ψ
 \vdots \vdots
 A _

However, Chiriacescu and von Heusinger point out that modified definite complements require the use of *pe*. Compare the acceptability of the sentences in (13). The fact of the matter is that the modifier must agree with the head in Romanian: any abstract syntactic feature passed to the head must be passed to the modifier as well. The nature of the representation and the above assumptions on feature-specification make it possible to capture this subtle difference in grammaticality. Consider (14a), the representation of (13b). Since the modifier does not have a ψ -node, it cannot get Accusative case at the same time as the head in a one-to-two spreading. It cannot get it either from the ψ -node of the verb, since line crossing is banned in the representation. Therefore, the structure may not be perfect.

(13) a. Doctorul îl examinează pe băiatul bolnav.
 Doctor.DEF CL examines PE boy.DEF sick
 “The doctor examines the sick boy”.

b. ? Doctorul examinează băiatul bolnav.
 Doctor.DEF examines boy.DEF sick
 “The doctor examines the sick boy”. (Chiriacescu and von Heusinger 2010: 4)

- b.
- | | | | | |
|---------------|---------------|---------------|---------------|---------------|
| Doctorul | l-examinează | pe | băiatul | bolnav. |
| \varnothing | \varnothing | \varnothing | \varnothing | \varnothing |
| ψ | ψ | ψ | ψ | ψ |
| A | A | A | A | A |

It appears that the nature of the representation forces the reanalysis of *pe* as a case-marker. The harmonization process extends the reanalysis to another type of argument, namely indirect object. In this case, a reverse process apparently takes place: a genuine case marker is reanalyzed as a post-position, which triggers a clitic doubling construction. Thus in (15b) the dative NP is reanalyzed as [părinți lor], where *lor* is a case-assigning post-position. Notice that (15a) is equally good as (15b), which is nevertheless preferred in modern Romanian, as Cojocaru points out.

- (15) a. Am scris părinților o scrisoare.
 have.1SG written parents.DAT a letter
- b. Le-am scris părinților o scrisoare.
 3PL-have.1sg written parents.DAT a letter
 'I wrote my parents a letter.' (Cojocaru: 102)

In this section I have presented the feature structure of verbs and clitics in Romanian, and the nature of the direct object marker *pe*. In the next section, I turn to the core of this research, namely the account of clitic clusters.

4. Clitic clusters

4.1 Why dative must precede accusative

We have seen that simple clitics must exit the VP domain under OCP, just like in other Romance languages, including French, Italian, Spanish, and Portuguese (cf. Desouvrey 2018, 2008, 2000). When two clitics are in a cluster, an OCP effect takes place as well. In addition, Romanian imposes an important restriction on clitic clusters: the dative clitic, irrespective of its position with respect to the verb, has to precede the accusative clitic in every cluster. The obvious question is: Why is this the case?

In this theory, the verb merges with a first clitic forming a constituent which is then merged with the second clitic. Which clitic is merged first depends in general on the feature specification of each clitic and ultimately on the interaction of the overall set of features. For instance, in French both orders are attested (16). In the Romanian case, I claim the required order follows from the presence of the feature ψ (animate, even human) in the paradigm, specifically in the direct object branch of the verb, which induces a confusion regarding the Animacy Hierarchy (AH): if both arguments are sensitive to the feature humanness/animacy, the default state of the grammar by virtue of which animate arguments take precedence over non-animate ones is not applicable. Therefore the AH must be enforced by a rigid word order in favor of the IO argument.⁶ This is indeed ensured by the very principles of the structure building, which forces the verb to first compute the best match (the argument which it shares most features with) in its argument list. This is succinctly schematized in (17a), where the bundle of possible

⁶ It appears that certain grammatical features and principles that regulate them will stay in a latent state if they do not interact with other similar features. Another obvious example is provided by the feature ω , which defines an element as a vector (sensitive to scope) as opposed to a scalar. When two such elements interact, a vector (or superiority) effect takes place; otherwise such a feature may not be isolated.

thematic features is appended to the verb and the clitics in curly brackets. (17b) is excluded since the first clitic is only compatible with the verb by lack of features, \emptyset being a dummy node.

- (16) a. Elle va me le dire.
 She goes me.DAT it.ACC tell.
 'She will tell it to me.
- b. Elle va le lui dire.
 She goes it.ACC him/her.DAT tell.
 She will tell it to him.
- (17) a. $[[Verb_ \{\delta, \psi\} \quad CL1_ \delta] \quad CL2_ \emptyset]$
 b. $*[[Verb_ \{\delta, \psi\} \quad CL_ \emptyset] \quad CL_ \delta]$

4.2 π -effect

As seen above, under OCP object clitics must exit the VP domain, where they are generated. In French, they simply adjoin to the first available element they run into. In Romanian, however, clitics are specified for the feature $[\pi]$, and therefore they must reach the right edge of the structure, just like *wh*-elements. Let us illustrate this with the derivation of (18a). It starts from the natural input (19a), where the full forms of the clitics are computed in turn, the dative first and then the accusative. Notice that if a single clitic is used, it normally appears in its full form (18b). Both clitics must move under OCP to the left edge of the structure. The clitic which is closer to the target appears to have more features. Therefore, it must move first (19b), and then the second clitic moves past it (19c). This structure is incorrect, since it violates the strict order imposed by the ψ -hierarchy. The dative clitic could regain the initial position by a further movement, as in (17d). This step is forbidden, however, because it overloads the derivation.⁷ We may note that all clitic clusters in Romanian are doomed to run into a similar problem, except those containing third person feminine *o* (see below).

- (18) a. Mi-l trimite astăzi.
 1SG.IO-3SG.DO s/he.sends today.
 'S/he sends it to me.'

⁷ The number of derivational steps may be at most 3, consistent with the Derivation Equivalence Number ($DEN \leq 3$); see Desouvrey (2007, 2008, 2018).

- b. îmi/*mi trimite trandafirul astăzi.
 'S/he sends me a rose today.'

- (19) a. [[[trimite îmi] îl] astăzi]
 b. îmi [[[trimite] îl] astăzi]
 c. îl [îmi [trimite astăzi]]
 d. *îmi [îl [trimite astăzi]] (overload derivation)

To avoid this ill-derivation, Romanian systematically resorts to compounding, just like Italian, as I argue. By this process, two clitics are combined so that they come to behave as a single element, where every pair of any feature is reduced to a single occurrence. This is insured by the following morpho-syntactic rule (20), where F stands for feature and # and + indicate a word and a morpheme boundary respectively.

(20) **Romanian Type-1 Clitic Compound**

$$\begin{array}{c}
 \text{CL1} \# \text{CL2} \rightarrow \text{CL1} + \text{CL2} \\
 \begin{array}{ccc}
 | & | & \diagdown \quad \diagup \\
 \text{F} & \text{F} & \text{F}
 \end{array}
 \end{array}$$

Thus, (20) applies to the sequence in (19a) and returns a compound which is phonetically identical to the input, as seen in (21a). Such a result is not learnable, and in fact is ruled out by the ban on vacuous operations. To rescue the compound, the sandhi rule in (22) alters it by rendering it unrecognizable from its parts (21b). It should be clear that this rule deletes the vowel *î* anywhere within the compound and it seals the latter by obliterating the boundary. As a result, the sequence yields a new word-like element (*mil*).

- (21) a. îmi # îl → *îmi + îl
 b. *îmi + îl → mi-l (mil)

(22) **Romanian Type-1 Sandhi**

$$\hat{i} \rightarrow 0 / [\text{ ______ }]_{\text{compound}}$$

Clearly, this type of compounds is automatically formed in the input sequence, prior to the movement under OCP. Subsequent sandhi rules do not count toward the load of the derivation, under

the natural assumption that phonological processes are running in parallel to the syntactic derivation. Similarly the following clusters are correctly compounded and obliterated under (22). The hyphenation is required by the orthography, but in our view the clitics form a single word, as shown in parentheses.

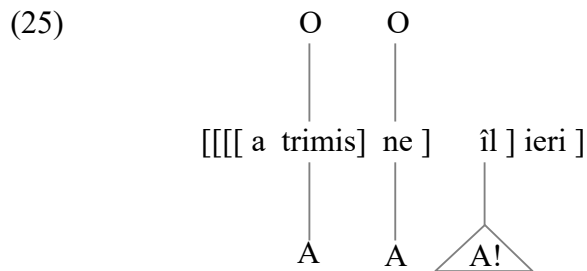
- (23) a. $\hat{m}i + le \rightarrow mi-le$ (mile)
 b. $\hat{i}i + \hat{i}l \rightarrow i-l$ (il)
 c. $\hat{i}i + \hat{i}i \rightarrow i-i$ (ii)
 d. $\hat{i}i + le \rightarrow i-le$ (ile)

4.3 Case-licensing effect

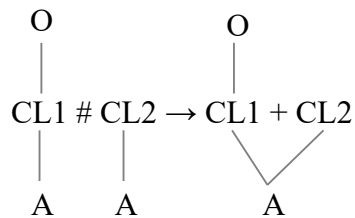
Besides the compounds just discussed, I show that there is a further type of compounds in Romanian. It occurs when the first clitic in a cluster is specified for both accusative and oblique case, as is the case of clitics *ne*, *vă*, *îi*, and *le*. Being in the first argument slot, these clitics absorb both cases of the verb, preventing the second one from being licensed, just as in Italian (Desouvrey 2018). Let us illustrate this type of cluster with the derivation of the sentence shown in (24).

- (24) a. Ni l- a trimis ieri.
 1PL.IO 3SG.MASC.DO has sent yesterday.
 'S/he sent it to us yesterday.'

As can be seen in the input (25), the first clitic saturates both cases of the verb, and therefore the accusative case of the second one cannot pair up with the verb. In the present perspective, this structure can be either rejected or repaired. A reordering of the clitics would solve the problem, but this is not possible under the rigid order imposed by the feature ψ . The only way out for the grammar is to create a new type of compound (26), which is sealed with two further sandhi rules (27). It is important to notice that both types of compound overlap, since the clitics mentioned in (25) are also π -specified. It appears that type-1 compounds is more general, and its sandhi rule is expected to apply in the type-2 compound as well, as far as there is no conflict. This is indeed the case, as shown in (28).



(26) **Romanian Type-2 Compound**



(27) **Romanian Type-2 Sandhi**

- a. $\check{a} \rightarrow i / [\text{ } +]$
 b. $e \rightarrow i / [\text{ } +]$

- (28) a. $ne \# \hat{i}l \rightarrow ne + \hat{i}l$ (Type-2 compound)
 b. $ne + \hat{i}l \rightarrow ni-l$ (nil) (Type-2 and type-1 sandhi)

One can check that these three sandhi rules correctly account for the following clusters which are compounded under (27).

- (29) a. $ne + \hat{i}i \rightarrow ni-i$ (nii)
 b. $ne + le \rightarrow ni - le$ (nile)
 c. $v\check{a} + \hat{i}l \rightarrow vi - l$ (vil)
 d. $v\check{a} + \hat{i}i \rightarrow vi - i$ (vii)
 e. $v\check{a} + le \rightarrow vi - le$ (vile)
 f. $le + \hat{i}l \rightarrow li - l$ (lil)
 g. $le + \hat{i}i \rightarrow li - i$ (lii)
 h. $le + le \rightarrow li - le$ (lile)

4.4 The clitic-auxiliary/verb cluster

Let us turn now to the sequences involving a clitic and an auxiliary beginning with a vowel. The clitic

and the auxiliary form a prosodic unite, as often mentioned in the literature (e.g. Popescu 2000). In my perspective, the auxiliary must make a compound with the clitic. This appears to be the third type of compound, which is stated in (30). A compound formed under (30) must be obliterated by two further sandhi rules, (31), which states that *ă* must be deleted when it appears within a compound and next to a morpheme boundary, and vowel *e* must glide in the same context.

- (30) **Romanian Type-3 Compound (T3C)**
 $CL \# V_{aux} \rightarrow CL + V_{aux}$ (V stands for vowel)
 (or $CL1 \# CL2 \rightarrow CL1 + CL2$)

- (31) **Romanian Type-3 Sandhi**
 a. $\check{a} \rightarrow 0 / __ +]$
 b. $e \rightarrow \text{e} / __ +]$

Consider the contrast in (32), where the clitic *vă* undergoes two different transformations (*v* vs. *vi*). In (32a), the clitic has moved to the edge and comes to precede the auxiliary. Then T3C applies, and as a result the vowel of the clitic is deleted under (31a), as shown in (33a). In (32b), where the compound has been formed while the clitics were in situ, a conflict seemingly appears between T2C and T3C. If the latter applies, its sandhi rule must delete the vowel of the clitic, which would be incorrect. Actually this conflict does not arise because the sandhi rules enter in an Elsewhere relationship. The compound rules and their respective sandhi rules operate in tandem. When a compound is formed, its own rule takes precedence over others that have the same structural description. Thus, non-local rules are relegated elsewhere in the grammar. All of the sandhi rules seen above mostly apply in a similar context, hence they are context-free, which make them compatible with the three types of compound whenever their structural description is met. However, when two such rules are in a competition to apply to a structural description, the local one thwarts the application of the more general one. Thus in (32b), T2C normally applies, yielding the desired result, as shown in (33b).

- (32) a. **V-a** trimis trandafirul ieri.
 2PL.IO-has sent rose.the yesterday.
 'she/he sent a rose to you yesterday.'

- b. **Vi** **I-a** trimis ieri.
 2PL.IO 3SG.DO-has sent yesterday
 'She/he sent it to you yesterday.'

- (33) a. $v\check{a} \# a \rightarrow v\check{a} + a$ (T3C) $\rightarrow v-a$ ($\check{a} \rightarrow 0$) (va)
 b. $v\check{a} \# \hat{i}l \rightarrow v\check{a} + \hat{i}l$ (T2C) $\rightarrow vi-l$ ($\check{a} \rightarrow i$; $\hat{i} \rightarrow 0$) (vil)

It should be noted that the fronted cluster in (32b) may not make a further compound with the auxiliary. The reason is that the compound rules apply to existing elements in the paradigm and not to derived ones. Thus, a compound may be resyllabified with an adjacent word under general rules of syllabification of the language. Indeed, the sequence *vil-a* (CVC-V) is resyllabified in *vi-la* (CV-CV) (also see below).

Let us now turn to the case where a clitic makes a compound with a lexical verb. This occurs with gerundives and imperatives, where the OCP fails. In these moods, the OCP cannot take the clitic anywhere without incurring a string vacuous movement.⁸ Thus, whenever the OCP fails, the clitic is automatically affixed to the verb, which triggers the application of the sandhi rules. In imperative verbs the process is straightforward, as shown in (34). Notice that in (34d) the sandhi rule does not apply (see below).

- (34) a. Arată-mi/*îmi pantoful!
 Show me the shoe
 b. Arată-i/*îi cartea!
 Show him/her the book
 c. Arată-ni-l! (ne + îl)
 Show it to us!
 d. Arată-ne-o!
 Show it to us!

With negative imperatives (35), the OCP can move the clitics out of the VP domain. The derivation proceeds as in (36). The compound is automatically created and then obliterated. It moves to the left edge of the structure, (36b), which triggers the leftward movement of the negation marker,

⁸ Given a verb V and its complement O, a grammar may compute them as either VO or OV. Thus if from the simple structure [V O], O moves past V, the result would be OV. Such a derivation is pointless, since it results in an option freely given by the grammar, assuming that a movement comes with a cost.

(36c). (See below further considerations on imperatives.)

(35) Nu mi-l arăta!
Don't show it to me!

(36) a. [nu [[arăta îmi] îl]] → [nu [arăta mi-l]]
b. mi-l [nu arăta]
c. nu [mi-l arăta]

With gerund verbs, a vowel is inserted at the boundary by the rule in (37). This rule serves to the resyllabification of the compound, but also for the obliteration of the boundary. This is illustrated in (38). Notice that (37) predicts that if the gerund verb is not in a compound, *u*-epenthesis may not occur. This is borne out, since the clitic *o* need not be compounded (38d) (see below).

(37) $0 \rightarrow u / \text{gerund}] __ +$

(38) a. văzându-mă
seeing me
b. văzându-i
seeing them
c. dându-mi-l
giving it to me
d. văzând-o
Seeing her

If the rationale of (37) were only the resyllabification, the epenthetic vowel would not appear when the clitic begins with (or is reduced to) a vowel, as in (38b). If the gerund verb and the clitic were not in a compound, there would be no reason for *î*-deletion to apply. Thus, it is incorrect to state that both rules feed each other (cf. Luis????). Notice that, as can be seen in (39), the structural description of each rule is independent from the other so that there is no need for the application of the rules to be ordered.

(39) a. văzând # îi (input) → văzând + îi (compounding)
b. văzându + i (u-epenthesis and î-deletion)

4.5 Optional compounds

Up to now, we have seen that compounding is an obligatory process involving clitics, whether pronouns or auxiliaries. It appears that in certain cases, this process is quite optional. Consider the following sentences (adapted from Cojocaru 2003). As can be seen, the particles *și* 'and' and *nu* 'not' may form a single prosodic unit with the clitic argument, which the orthography indicates by hyphenation. The question now is to account for this non-mandatory processes in which *î*-deletion comes to take effect out of its original domain, namely T1C.

- (40) a. Mă ascultă **și îmi** / **și-mi** răspunde.
'He is listening to me and answering to me.'
- b. **Nu îl** / **nu-l** ascultă și **nu îi** / **nu-i** răspunde.
'He is not listening to him and he is not answering to him.'
- c. **Nu o** / **n-o** ascultă și **nu îi** / **nu-i** răspunde.
'He is not listening to her and he is not answering to her.'
- d. **Nu a** / **n-a** ascultat-o și **nu i-a** /***nu-i a** răspuns.
'He did not listen to me and did not answer me.'

Three of the phonological rules discussed above delete the vowels *e*, *ă* and *î* in the compounds. As a result, the number of syllables in the compound is reduced. This is a welcome result even though it is not the rationale of the compounding process. This prosodic gain, I suggest, leads to the relaxation of the morphological rules so that the pronominal clitics may freely enter into a compound with another elements outside their paradigm.

Thus in (40a) and (40b), the clitics *îmi*, *îi* and *îl* enter in a compound with *și* and *nu*, which triggers the *î*-deletion rule. In (40d), where the auxiliary forms a compound with negation, there is no dedicated sealer. Thus the grammar turns to a simple elision rule ($u \rightarrow 0/ \text{---} +$) that deletes the vowel of negation near a boundary. This type of rules, which deletes a vowel next to another one, is highly productive across languages of various families. As for the *n-o* unit in (40c), it may not be a morphosyntactic compound of the types discussed above, since clitic *o* is never compounded (see below). Instead, it results from its adjunction to negation with which it shares the feature ω . In the complex thus formed ($nu=o$), the elision rule may still apply, yielding an acceptable result.

The over-generalization of the compound rules may well induce some grammatical illusion. This seems to be the case with verbs beginning with an unstressed vowel, as in the following examples (Luis

????). What seems to happen here is that the unstressed syllable is reanalyzed as a clitic, which prompts a compound rule and then its obliteration with the *ă*-elision rule (41a). Since there is no *e*-elision rule, this process fails in (41b).

- (41) a. Mă așteapta / M- așteapta
acc.2sg waits
's/he waits for you'
- b. Te așteapta / *T- așteapta
acc.2sg waits
's/he waits for you.'

In the next subsection, I turn to the clitic *o*, which departs from the other clitics in the paradigm with respect to the feature ω , which it is specified for, and the feature π , which it lacks.

4.6 The special status of clitic *o*

As can be seen in Table 1 above, it is assumed that the clitic *o* differs from the others with respect to two important features: (a) it is ω -specified, which makes it a vector, and (b) it is the only inbound clitic in the paradigm, a property it owes to its lack of feature $[\pi]$. Indeed, it displays some special behaviors, the most obvious of which is its encliticization to the verb when a compound perfect auxiliary is used. This is illustrated in (42). In (42a, c, e), *o* is in a preverbal cluster, while in (42b, d) it parts from the other clitic by staying in situ.

- (42) a. Mi-o trimite astăzi.
1SG.IO-3SG.DO s/he.sends today
- b. Mi-a trimis-o ieri.
1SG.IO-s/he.has sent-3SG.DO yesterday
- c. V-o trimite astăzi.
2PL.IO-3SG.DO s/he.sends today
- d. V-a trimis-o ieri.
2PL.IO-s/he.has sent-3SG.DO yesterday
- e. V-o va trimite mâine.
2PL.IO-3SG.DO s/he.will send tomorrow

It is important to recall that clitic clusters in Romanian must be compounded for two reasons: the dative clitic must always precede the accusative one and both clitics are specified for π . Because of this feature, the required order cannot be preserved if each clitic moves independently of the other. Since the clitic *o* is not π -specified, there is no reason for the cluster to be compounded. Thus in (42a) and (42e), the dative clitic normally moves first and then the accusative adjoins to it, forming a syntactic compound, as seen in the derivation (43). The cluster $\hat{m}i=o$ makes up a compound, ' = ' being a morpheme boundary, and is thus automatically obliterated. The derivation of (42e) is similar, except that the obliteration rule is, of course, different ($v\check{a} = o \rightarrow v-o$).

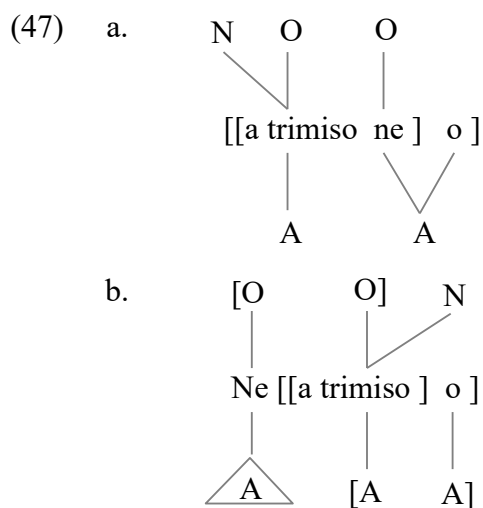
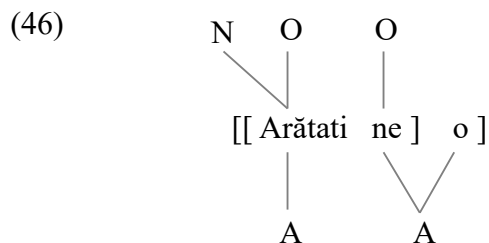
- (43) a.
- | | | |
|-------------|--------------|--------------|
| O | O | |
| | | |
| [[[trimite | $\hat{m}i$] | o] astăzi] |
| | | |
| A | | A |
- b. $\hat{m}i$ [[trimite o] astăzi]
- c. $\hat{m}i = o$ [trimite astăzi] \rightarrow $mi-o$ [trimite astăzi] (Obliteration)

Sentences (42b) and (42d) strongly support the compound analysis of Romanian clitics. If *o* were in a compound, it would not stay in situ, or else the cluster would stay in situ. In no way the clitics would be separated. Now, it must be explained why the auxiliary blocks the movement of *o*. The failure of the OCP is due to the feature specification of this tense of the auxiliary. I suggest that the auxiliary is a vector, i.e., ω -specified, just like the clitic. Therefore a superiority effect takes place, blocking the reordering that would take place if the clitic were to exit the domain of the compound tense verb. Similar phenomena can be found in the personal pronoun paradigm of Italian, Portuguese and Spanish at various degrees and places. With other compound tenses (e.g. 42e), *o* is head-adjoined to the dative clitic, and therefore only the context-free rules apply, as shown in (44).

- (44) a. $le = o \rightarrow le - o$
- b. $ne = o \rightarrow ne - o$
- c. $v\check{a} = o \rightarrow v - o$
- d. $\hat{m}i = o \rightarrow i - o$

Consider now a post-verbal cluster complement of an imperative verb, as in (45a). Since the first clitic takes both cases of the verb, *o* may not be licensed. Thus the obvious question is how this sentence can be grammatical. One can account for this fact by assuming that the grammar still takes advantage of the compound rule T2C, without its sandhi rules, to license the second clitic. Thus the latter is jointly licensed with the first one, given the case reduction in the compound. This is shown in (46). However, since this compound is not sealed, the clitics are still autonomous, and the first one can eventually break free from the other to satisfy the OCP, as in (45b). The derivation of the latter is given in (47), where each clitic correctly pairs up with the relevant case of the verb after the movement (47b). It is important to note that Nominative case is, by assumption, normally absorbed in personal moods by a dummy clitic (*pro*) so that it does not interfere with the extra case of *ne*.

- (45) a. Arătați-ne-o!
Show it to us!
- b. Ne a trimiso-o.
S/he sent it to us.



In any event, *o* may not behave like the other clitics, since it is ω -specified and it lacks the feature π , which allows it to move inbound. To deal with it, the grammar just uses the available apparatus largely used in the language, namely compounding. If a language does not have a productive compounding rule, (46) and (47) are expected to be sharply ungrammatical. French is a such language. In French imperative verbs, the ill-cluster *le-me* / *me-le* surfaces as *le-moi* / *moi-le* (substandard) (cf. *dis-le-moi* / *dis-moi-le* 'tell it to me') but *me le* is correct in other moods, since the clitics are computed in the order DO>IO.

In simple gerund verbs, there is no reason for the clitic *o* to be compounded. It does not even have a reduced form to trigger a prosodic gain. Therefore, the *u*-epenthesis cannot apply, as seen in (48) (cf. 38d).

- (48) v  z  nd-o ...
 seeing her...

To sum up this section, it is shown that Romanian clitics, broadly speaking to include auxiliaries, interact to make up three types of compound. The compounds are sealed by a series of sandhi rules which obliterate any boundary, thus allowing the cluster to behave as a new word-like element. The compound rules and the sandhi rules operate in tandem so that any conflict are naturally resolved by the Elsewhere Condition, which prioritizes the more specific rule.

5. Accounting for special clusters

5.1 An apparent person restriction

In Romanian the Person Case Constraint (Bonet 1991) does not hold (S  vescu 2006). The following data show that a combination of a third person dative and a second person accusative, as well as the combination of a first and a second person clitic are acceptable.

- (49) a. **I** **te** - au recomandat ieri.
 3DAT 2ACC have recommended yesterday.
 ‘‘They have recommended you to him yesterday.’’
- b. **Mi** **te** - a prezentat Ion la petrecere.
 1DAT 2ACC has introduced John at party.
 ‘‘John has introduced you to me at the party.’’ (S  vescu 2006:11 and 13)

These facts ruin the Person Case Constraint (PCC) as a universal constraint. However in Desouvrey (2018) the PCC is restated in terms of thematic features, see (50). In my perspective, Nominative and Accusative bear no thematic feature, unlike the oblique case, which may depend on such thematic nodes like dative, locative, etc. In Romanian, DO and IO clitics are distinct morphemes in the singular, unlike French, for instance. Indeed, as seen above, they are specified for a single case, Accusative and Oblique respectively. If for instance, *te* is used as IO and *îti* as DO, an ill-formed sentence obtains. Therefore, I-II and III-II clusters are expected to be grammatical since they include only one thematic feature borne by the IO argument, the first element.

(50) **Person Case Constraint (revisited)**

A cluster of object clitics must contain one and only one thematic feature.

There are, however, certain restrictions on such clusters. Săvescu points out that a combination of a third person dative and a first person accusative (51a) is not grammatical for most speakers. Moreover, all speakers reject the combination of a second person dative and a first person accusative (51b). To account for these sentences, Săvescu invokes a person hierarchy ($1^{st} > 2^{nd} > 3^{rd}$) and a case hierarchy (Dative > Accusative), which summarizes the problem instead of accounting for it.

- (51) a. %I m - au recomandat ieri.
 3DAT 1ACC have recommended yesterday.
 ‘‘They have recommended me to him yesterday.’’ (Săvescu's (12))
- b. *Ți m - a prezentat Ion la petrecere.
 2DAT 1ACC has introduced John at party.
 ‘‘John has introduced me to you at the party.’’ (Săvescu's (13b))

I show that the ill-formedness of (51) follows from the analysis presented above. Each element of the cluster is generated as an argument in post-verbal position. Then both clitics are compounded under the T1C, and as a result the boundary is obliterated by the sandhi rule (22), as shown in (52).

- (52) a. îi # mă → îi + mă → imă (before movement)
- b. îți # mă → îți + mă → țimă (before movement)

These outbound compounds then move to the left edge of the clause, where they come to be adjacent to the auxiliary. Then naturally any speaker will proceed to apply T3C in order to create a new compound

with the auxiliary. However since *imă* and *țimă* are syntactic elements, not lexical entries, they must break into their parts (see (53)) for T3C to apply. Precisely at this point, the derivation fails under the natural assumption that the obliteration cannot be undone in the same derivation (perhaps an effect of the Strict Cyclicity, a well-known phonological principle). This analysis is supported by the fact that such compounds are somewhat more acceptable if they stay in situ, as is the case of gerundive and imperative verbs, since they are not laid in two different environments.⁹

- (53) a. *imă* # *au* → * *îi* # *mă* + *au* / * *i* # *mă* + *au*
 b. *țimă* # *a* → * *îți* # *mă* + *a* / * *ți* # *mă* + *a*

That being said, certain speakers can be tempted to proceed anyway with the derivation. Given the partial similarity with (49a) (both clusters begin with a 3rd dative), an illusion of grammaticality obtains. We may note that the sentences in (49) yields no doubt; in effect, no speakers may be lured to make a further compound with the auxiliary, since there is no sandhi rule that deletes vowel *e* in the paradigm (except in the case of reflexive *se*, see below).

5.2 Clitics and imperative verbs

Consider now the contrast below. In (54) a cluster of a second person dative plural and a first person accusative clitic is ungrammatical as the complement of a plural imperative verb. However, the same type of cluster is perfect with a singular imperative (55).

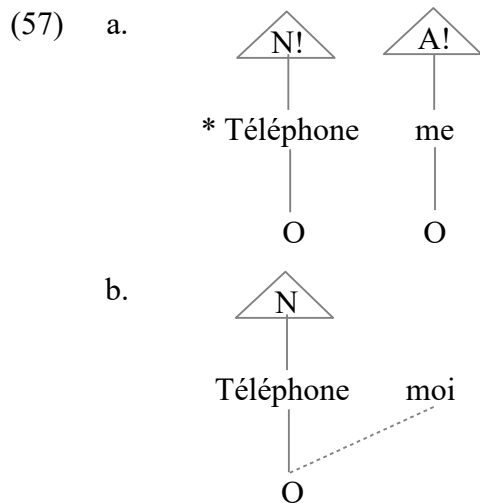
- (54) **Luai - vi - mă* drept martor, și vei câștiga procesul!
 take.2PL 2DAT 1ACC as witness and will.2nd win trial.THE!
 “Take me as a witness (for yourselves) and you will win the trial!” (Săvescu's (35))
- (55) *Ia - ți - mă* drept martor, și vei câștiga procesul!
 take 2DAT 1ACC as witness and will.2nd win trial.the!
 “Take me as a witness (for yourself) and you will win the trial!” (Săvescu's (36))

Analyzing those sentences under her Person Hierarchy, Săvescu builds on the assumption that (54) is a surrogate imperative, as opposed to a genuine imperative like (55). Referring to Zannutini (1997, not seen), she assumes that genuine imperatives lack the tense node, and as a result are not

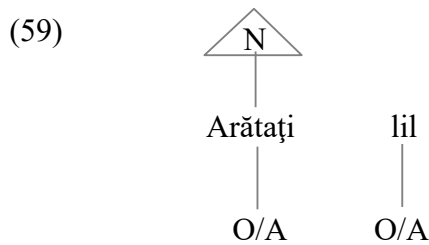
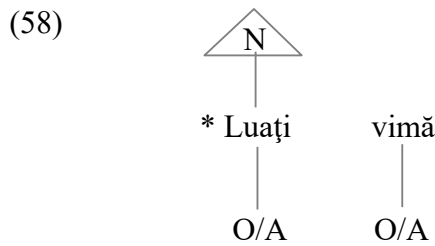
⁹ The judgement about these compounds in gerundives is reported in Rîpeanu Reinheimer et al. (2013).

sensitive to the Person Hierarchy. Under the present assumptions, it amounts to saying that imperatives normally need not have nominative case; a genuine imperative, not borrowed from another mood, may not have it. It is true that the extra-nominative case of verbs that lack an external argument is not without consequences on the grammar. In French, it bars first and second person clitics from appearing as an object of imperative verbs (cf. Desouvrey 2000). As illustrated below, the clitic *me* is the dative indirect object in (56a), but not in (56b), where it is replaced by its stressed equivalent *moi*. The problem is caused by a clash with Nominative and Accusative, which thus fail to pair up so as to make a tier, as shown in (57a). In (57b), on the other hand, the residual Nominative does not interfere with other case features, and therefore is ignored, and hence a well-formed sentence obtains.

- (56) a. Tu me téléphones!
 you me telephone!
 You telephone me!
- b. Téléphone-moi/*me!
 Telephone-me!



Returning to Romanian, a similar analysis may not hold, irrespective of the type of imperative (genuine or surrogate). In effect, in either type of imperatives, the cluster is compounded and obliterated, which eliminates the stranding feature of any clitic. The syntax comes to deal with the compound *vima* [A, O] in (58), instead of two separate clitics *vi* [A, O] + *mă* [A]. One can see that the configuration (58) is similar to that of a well-formed cluster (59).



From the present theory, it is likely that the compound fails in (54) because it has failed elsewhere with the same type of verb (recall the discussion around (51)). This is an effect of the harmonization process which aims to the elimination of exceptions in the grammar, whenever possible. This process arises from the assumption that there are no conditional constraints (if-then) in the grammar. That is, with any indicative verb, the cluster is rejected, independently of the actual use, namely imperative.¹⁰ On the other hand, (55) is acceptable because the verb has a special morphology (genuine imperative) which has not been used elsewhere with this type of cluster.

5.3 A coreference effect

Romanian presents a further interesting contrast involving clusters with a third person reflexive. As can be seen in (60), a third person dative reflexive is incompatible with a first or a second person clitic, but with a third person accusative instead, the cluster is grammatical. Furthermore, a cluster of a third person accusative reflexive and a first or second person dative pronoun yields a well-formed sentence, (60c).

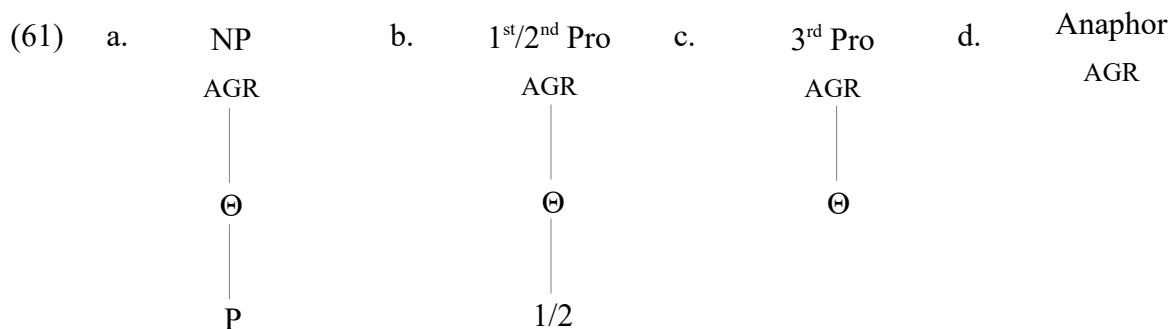
- (60) a. *Maria **și** **m-/ te** - a luat drept sclav.
 Mary REFL.DAT 1/2 ACC has taken as slave.
 “Mary has taken me/you to be her slave (for herself).”

¹⁰ On this view, it turns out that the grammar uses some kind of the so-called Murphy's law, which may be stated as follows: something that went wrong once in some place will always go wrong in any place.

- b. Maria **și I** - a luat drept martor.
 Mary REFL.DAT 3ACC has taken as witness.
 “Mary has taken him as a witness (for herself).” (Săvescu's (16))
- c. Maria **mi/ți s** - a prezentat la petrecere.
 Mary 1DAT/2DAT REFL.ACC has presented at party.
 “Mary has introduced herself to me/to you at the party”. (Săvescu's (17))

It is obvious that these sentences cannot be accounted for with the apparatus used so far. The dative reflexive and first / second dative are specified for the same case features, and therefore there would be no reason for the cluster **și-te* to be ungrammatical unlike *mi-te*. Similarly, the acceptability of the cluster *și-I* is not expected besides **și-m*, since both *îl* and *mă* are specified for the same case, namely Accusative. What we need to consider here is the intrinsic difference between the reflexive and third and second person clitics. The key point to be taken into account here is that the third person reflexive is an anaphor, as opposed to pronouns like first, second and third person clitics.

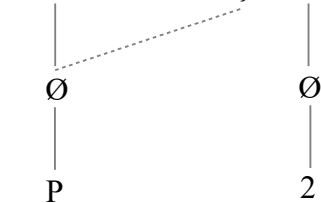
Referring elements, namely NPs, pronouns, and anaphors, have different representations, as discussed in Desouvrey (2003). The referential tree of an NP comprises a root node (Agr), a thematic node, and a terminal feature, represented with such capital letters as P, Q, etc., which are the virtualization of any real (or imaginary) world element. For first and second person pronouns, letters are replaced by numbers 1 and 2, respectively. Third person pronouns lack a terminal feature, while anaphors lack the thematic node, hence the terminal feature. This is shown in (61).



NPs and first and second person pronouns are referentially autonomous, since they are permanently linked to a real world element via a virtualization feature. A third-person pronoun must receive a terminal feature from an NP by spreading. In other terms, they do not have access to the real world. Similarly an anaphor gets a thematic node (and the associated terminal feature) from an NP or a

pronoun that can spread a thematic node to its bare root node. An important point is that the coreference relation is not derivable. It is set from the input and must be preserved throughout, whatever happens to the elements in the course of the derivation. Let us see how this accounts for the cluster in (60a). As shown in (62a), a coreference relation is correctly set between the anaphor and the subject. The cluster undergoes compounding, a process that preserves different features of each element, but reduces identical features or nodes to a single occurrence. The root nodes of both clitics are thus combined, and as a result the anaphor inherits the thematic node of the II-pronoun. Since the latter is referentially constant (which is indicated by the solid association line), the coreference relation between the anaphor and its normal antecedent is lost, (62b), hence the ill-formedness of the compound (which comes to refer to itself). For the sake of clarity, the full representation of the ill-compound is given in (63).

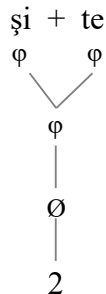
(62) a. Maria a luat își te drept sclav.



b. *Maria a luat își te drept sclav.

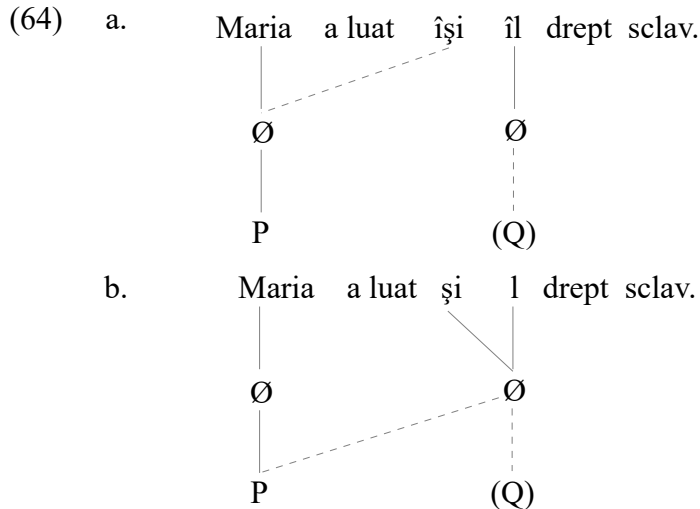


(63)



Let us see now what happens when the dative reflexive is in a cluster with a third-person pronoun, as in (60b). The same process takes place: the anaphor becomes a pronoun by acquiring a thematic node under compounding. However, its bond with the subject is not lost, since the third person pronoun

is referentially variable (remark the dotted line linking Q to the dummy thematic node vs. the solid line of the 2nd person pronoun in (62)). That is, having no direct link to the real world. It may catch any terminal feature from a nearby NP. Therefore the virtualization feature of the subject normally spreads to the compound, which can still get another referent (Q) from another NP in the discourse. This is shown in (64a, b).



Consider the case in which the reflexive clitic is the direct object (Accusative), as in (60c) above. Unlike its dative counterpart, the direct object reflexive is not π -specified, and in addition it is the only clitic to normally undergo the elision of its vowel *e* (cf. *le*/**l*, *te*/**t*, *ne*/**n*). The *e*-elision rule, which normally applies to a compound with vowel-initial auxiliaries, is stated as follows:¹¹

$$(65) \quad se \rightarrow s / [\text{ ____ } + V]$$

The derivation proceeds from input (66), in which the intended coreference with the reflexive anaphor and the subject does exist, since the dative association line is not in the same plane with the \emptyset -thematic of the subject.¹² Now since reflexive *se* is assumedly not π -specified, it need (and may) not make a compound with the 1st person clitic. Therefore, the latter moves to the left edge of the structure, as seen in (66b). From this derivational step, two further operations occur at the same time: the subject regains

11 The orthography of clitic *se* may be due to a decision by the language regulation body. According to Avram (1986), referring to Lombard (1972 not seen), *se* is usually spelled *să/s* in dialectal Romanian. In the standard language, *să* is not used but its reduced form is retained. The rejection of *să* may have been motivated by the necessity to formally distinguish the reflexive clitic from the subjunctive marker.

12 If they were in the same plane, the coreference bond would be lost, since line crossing is banned in the representation, just as in nonlinear phonology.

its initial position under OCP-2, while the reflexive adjoins to the first person clitic, which comes to be the closest host outside the V-domain, as seen in the last derivational step (66c).

- (66) a.
- $\delta(1)$
 |
 Maria a prezentat *îmi* se la petrecere.
 |
 \emptyset
 |
 P
- b. *îmi* [Maria a prezentat se la petrecere]
- c. Maria *îmi=se* a prezentat la petrecere → Maria *mi=s* a prezentat la petrecere

The latter undergoes two further processes: T3C, which triggers *î*-deletion, and then the *e*-elision of *se*. These occur as follows. The two-tiered cluster (67a) is linearized as in (67b). Then the elision rules apply, deleting the vowels, as seen in (67c). It should be clear that the obliteration of *îmi+a* may not take place before the linearization, since *i* is linked to another morpheme tier. If the obliteration were to take place in this context, it would yield a structure like (68), which is not linearizable under the natural assumption that linearization may not undo any word.¹³

- (67) a. *îmi* + a
- |
 se
- b. *îmi=se* + a
- c. *mi=s* + a

- (68)
- *m i a
 x x x
 |
 x x
 s e

The difference between (62), (64), and (66) is a bit subtle. However, as can be seen, the input

¹³ In Portuguese, clitics lie within infinitive verbs between the root and the infinitive affix. This phenomenon precisely occurs because the boundary between the affix and the root is not obliterated. Indeed, infinitive formation exhibit a high level of regularity in this language (see Desouvrey 2008, Rouveret 1993).

configuration is intrinsically different for each structure in a principled way. Therefore the variation is expected. In a nutshell, an ill-formed structure obtains if the coreference bond is lost as a result of compounding.

6. Conclusion

I have presented an account of Romanian clitic doubling and clitic clusters. Their peculiarity is due to the features π and ψ , which has not yet been isolated in other Romance languages. Their effects are clear: π -specified clitics (all of them but *o* and *se*) undergo an outbound movement while the feature $[\psi]$ imposes a strict ordering of the clitics within a cluster. It is shown that to overcome the effects of these features, two clitics are combined into a compound under three types of morphosyntactic rules. All compounds are sealed by a sandhi rule, which obliterates their boundary. As a result, the cluster comes to behave as a new lexical element. The theory I build on allows a fine-grained analysis of such a complex pronominal paradigm which turns out to go like a clockwork.

7. References

- Avram, Andrei (1986). Sandhi Phenomena in Romanian. In *Sandhi Phenomena in the Languages of Europe*, Henning Andersen (ed.), Walter de Gruyter, pp 551-574.
- Avram, Larisa and Martine Coene (2009). Null Objects and Accusative Clitics in Romanian. <https://www.researchgate.net/publication/2600650174>
- Bențea, Anamaria (2010). On Restrictive Relative Clauses in Romanian: Towards a Head-Raising Analysis. *Generative Grammar in Genova* 6: 165-190.
- Bonet, Eulalia (1991). Morphology After Syntax: Pronominal Clitics in Romance. Doctoral dissertation, MIT.
- Chiriacescu, Sofiana and Klaus von Heusinger (2010). Discourse Prominence and *Pe*-marking in Romanian. *International Review of Pragmatics* 2: 298–332.
- Cojocaru, Dana (2003). Romanian Grammar. SEELRC.
- Desouvrey, Louis-Harry (2000). Romance Clitics and Feature Asymmetry: An Autosegmental Based-Approach. Doctoral dissertation. <http://ling.auf.net/lingBuzz>.
- Desouvrey, Louis-Harry (2002). Adverbs, Negation and OCP Effects. www.semanticsarchive.net.
- Desouvrey, Louis-Harry (2003). The Proper Treatment of Coreference Relations. www.semanticsarchive.net.
- Desouvrey, Louis-Harry (2005). Romance Clitic Clusters: The Case Connection. In Heggie, L. and F. Ordóñez (eds), *Clitic and Affix Combinations*. Amsterdam: John Benjamins.
- Desouvrey, Louis-Harry (2006). Underspecification and Long-Distance Antecedent: The Case of Chinese Ziji. www.semanticsarchive.net

- Desouvrey, Louis-Harry (2007). Wh-Interrogatives: The OCP Cycle. www.semanticsarchive.net.
- Desouvrey, Louis-Harry (2008a). Vector Effects on Wh-Interrogatives. <http://ling.auf.net/lingBuzz/000755/> and www.semanticsarchive.net.
- Desouvrey, Louis-Harry (2008b). Superiority Effect and Clitic Placement in European Portuguese. <http://ling.auf.net/lingBuzz/000764>.
- Desouvrey, Louis-Harry (2009). Nominal Predication, Relative Clauses and Wh-Extraction: The Amazing Syntax of Haitian Creole Se. <http://ling.auf.net/lingBuzz>.
- Desouvrey, Louis-Harry (2010). The Syntax of the So-Called Internally-Headed Relative Clauses: A Study of Korean Clause Structures. <http://ling.auf.net/lingbuzz/001170>.
- Desouvrey, Louis-Harry (2013). Periodic Linking: The How, Where and Why of Ergativity and Analogous Phenomena. [Http://ling.auf.net/lingbuzz/001655](http://ling.auf.net/lingbuzz/001655)
- Desouvrey, Louis-Harry (2018). The Syntax of Italian Clitics. <http://ling.auf.net/lingbuzz/>.
- Hill, Virginia (????). The Direct Object Marker in Romanian: A Historical Perspective. Ms., University of New-Brunswick, Saint John.
- Jaeggli, Osvaldo (1982). *Topics in Romance Syntax*. Dordrecht: Foris.
- Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Cambridge, MA: Blackwell.
- Luis, Ana (????). Mixed Clitic Placement: Evidence from Romanian. University of Coimbra, Portugal.
- Popescu, Alexandra (2000). The Morphophonology of the Romanian Clitic Sequences. *Lingua*, 110, 10: 773-799.
- Rîpeanu Reinheimer, Sandra, Liliane Tamowski, and Andra Vasilescu (2013). Pronouns (Chapter 5, sections 1.1-1.7). In Carmen Dobrovie-Sorin, and Ion Giurgea (eds.), *A Reference Grammar of Romanian*, Volume 1, Amsterdam: John Benjamins Publishing.
- Rouveret, Alain (1993). Cliticisation et Temps en Portugais Européen. *Revue des langues romanes*, 93, 2: 337-371.
- Săvescu Ciucivara, Oana (2006). Challenging the Person Case Constraint : Evidence from Romanian. In *Romance Linguistics 2006. Selected Papers from the 36th Linguistic Symposium on Romance Languages (LSRL), New Brunswick, March-April 2006*, J. Camacho, N. Flores-Ferrán, L. Sánchez, V. Déprez and M. José Cabrera (eds), John Benjamins, pp 255-269.