*Natural Language and Linguistic Theory* provides a forum for discussion of theoretical research that pays close attention to natural language data, so as to provide a channel of communication between researchers of diverse points of view. The journal actively seeks to bridge the gap between descriptive work and work of a highly theoretical, less empirically oriented nature.

**Photocopying.** *In the U.S.A.*: This journal is registered at the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923.
Authorisation to photocopy items for internal or personal use, or the internal or personal use of specific clients is granted by Kluwer Academic Publishers for users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the fee of USD 8.00 per copy is paid directly to CCC. For those organisations that have been granted a photocopy licence by CCC, a separate system of payment has been arranged. The fee code for users of the Transactional Reporting Service is 0167–806X/95 USD 8.00.
Authorisation does not extend to other kinds of copying, such as that for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale.
*In the rest of the world*: Permission to photocopy must be obtained from the copyright owner. Please apply to Kluwer Academic Publishers, P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

---

*Natural Language and Linguistic Theory* is published quarterly.
Subscription prices, per year: Institutions USD 252.00, Individuals USD 62.00.
Second-class postage paid at Rahway, N.J. USPS No.716–190.
U.S. Mailing Agent: Mercury Airfreight International Ltd., 2323 Randolph Ave., Avenel, NJ 07001.
Published by Kluwer Academic Publishers, Spuiboulevard 50, P.O. Box 17, 3300 AA Dordrecht, The Netherlands, and 101 Philip Drive, Norwell, MA 02061 U.S.A.
*Postmaster:* please send all address corrections to: *Natural Language and Linguistic Theory*, c/o Mercury Airfreight International Ltd., 2323 Randolph Ave., Avenel, NJ 07001, U.S.A.

*Printed on acid-free paper*

EULALIA BONET

# FEATURE STRUCTURE OF ROMANCE CLITICS*

The combination of pronominal clitics in Romance often triggers the appearence of unexpected (opaque) form, which always coincide with clitics that exist independently in the language. In this article, which analyzes opaque forms in Italian, Spanish, and, especially, Catalan, it is proposed that pronominal clitics constitute hierarchical structures of monovalent morphological features. This hierarchy reflects markedness relations. Most opaque forms are obtained through morphological rules that delink or insert morphological features, thus rendering the target structure identical to the structure of another clitic. Morphological rules take place within the Morphology Component, between S-structure and PF (cf. Halle 1990, 1991). Phonological information, not present in the syntax, is introduced by spell-out rules very late in the Morphology Component and provides the input to PF. Clitic order is determined through mapping to a template. Some other opaque forms arise at that point.

## 1. Introduction

Pronominal clitics have been the subject of much work in generative linguistics, especially with respect to their syntactic properties and, less often, their phonological properties. This article focuses on one aspect of pronominal clitics that has never received a systematic treatment. The phenomenon I am referring to is the appearence of opaque forms that often arise in clitic combinations. By "opaque forms" I mean outputs of clitic combinations that do not coincide with the output forms of those clitics in isolation. In Section 1.2 it will be shown that, although the existence of this phenomenon has been noted in the literature on clitics, only ad hoc solutions to it have been proposed. In this article, several opaque clitic combinations coming from varieties of Spanish, Italian, and, especially, Catalan will be discussed. It will be argued that the data presented provide a useful insight into the internal feature structure of pronominal clitics, and a proposal regarding this internal structure will be made, together with an analysis of the different types of opaque forms.

This article deals with a fairly limited set of data, which is meant to illustrate the phenomenon treated here. It should be taken into account, however, that there is a lot of dialectal variation among the Romance

* This paper is a revised version of some main aspects of my doctoral dissertation. I would like to thank the members in my thesis committee, Ken Hale, Morris Halle, Jim Harris, Mike Kenstowicz and Alec Marantz, as well as Joan Mascaró and Gemma Rigau, for all their comments. I would also like to thank the four anonymous *NLLT* reviewers for their helpful comments. Their contribution crucially improved the quality of the paper.

*Natural Language and Linguistic Theory* 13: 607–647, 1995.
© 1995 *Kluwer Academic Publishers. Printed in the Netherlands.*

languages with respect to the output forms of pronominal clitics, both in isolation and in combinations. None of the cases I have examined so far challenges the account proposed here.

### 1.1. *Opaque Forms*

One of the most well-known cases of opaque clitic combinations is the spurious *se* of Spanish, which appears instead of the usual third person dative clitic form *le(s)* when it combines with a third person accusative clitic. This is shown in (1c). (1a) and (1b) illustrate the use of the two clitics in isolation.[1] Pronominal clitics always appear in boldface in the examples. In the examples of combinations, I include in parentheses the non-existent transparent output form.

(1)a. El premio, **lo** dieron a Pedro ayer.
*the price 3rd-acc gave(3rd-pl) to Pedro yesterday*

b. A Pedro, **le** dieron el premio ayer.
*to Pedro 3rd-dat gave(3rd-pl) the price yesterday*

c. A Pedro, el premio, **se lo** dieron
*to Pedro the price, se 3rd-acc gave(3rd-pl)*
ayer. (**le lo/*lo le*)
*yesterday*
'they gave the price to Pedro yesterday'

As can be seen in (1b), the form of the clitic corresponding to a third person singular indirect object is *le* in isolation. However, in (1c) the surface form corresponding to this clitic is not *le* but *se*. We thus expect the output sequence *le lo* but find *se lo* instead.

A similar phenomenon can be found in Standard Italian with the combination of an impersonal clitic *si* with a third person reflexive clitic *si*. The examples in (2) are taken from Saccon (1988). (2a) and (2b) show that the form of the two clitics is *si* in isolation and also that the two instances

[1] In most of the examples that will appear in this article I use the Clitic Left Dislocation construction in order to provide an explicit antecedent for the clitics. Even though in the Clitic Left Dislocation construction the left dislocated constituent constitutes always new information, for convenience of exposition the English translation does not reflect this aspect.

Example (1a) sounds slightly forced in most dialects of Spanish because it is very common to double the indirect object with a clitic in this language.



of *si* occupy different positions, relative to the third person accusative clitic:[2]

(2)a. Lo **si** sveglia.
*3rd-acc impers. wake-up(3rd)*
'one wakes him/it up'

b. **Se** lo compra.
*refl 3rd-acc buys*
's/he buys it for herself/himself'

c. **Ci si** lava. (**si si*)
*ci si washes*
'one washes oneself'

Instead of the transparent sequence *si si* present in certain dialects of Italian, the output form is *ci si*. Later it will be concluded that the surfacing *si* form corresponds to the reflexive clitic.

Some dialects of Catalan provide a more spectacular case of opaque clitic forms, which, as far as I know, had not been noticed until very recently. In this case, the clitics involved are the neuter clitic *ho* /u/ and the ablative clitic *en* /n/.[3] (3c) shows the combination of these two clitics resulting in the opaque form /li/:[4]

(3)a. Això, **ho** trauré de l'armari després.
*this neut. will-take-out(1st) from the closet later*

b. De l'armari, **en** trauré això després.
*from the closet ablat. will-take-out(1st) this later*

[2] The difference in vowel quality present in (2b) (*se* instead of *si*) is irrelevant to the issues discussed in this article.

[3] The use of 'neuter' here has nothing to do with gender (nouns are always either masculine or feminine). *Ho* is related to sentential complements which function as direct objects and to certain types of predicates, among others. The term 'neuter' has always been used for the clitic *ho* in the traditional literature on Catalan.

[4] Some of the clitic forms discussed in this paper are not accepted in Standard (normative) Catalan, and therefore there is no common way to spell them. In these cases I use phonetic transcription.

(3)c. Això, de l'armari, [**li**] trauré
*this from the closet* [*li*] *will-take-out*(1*st*)
després. (**n'ho/*ho'n*)
*later*
'I will take this out of the closet later'

If the output form in the combination of a neuter clitic and an ablative clitic were transparent, the phonological sequence in (3c) would be [nu], a phonological sequence that can in fact be found in other (very restricted) cases. However, the resulting form [li] is not only opaque but is phonologically very different from the expected transparent form. Later we will examine other similar opaque output forms from Catalan.

Some dialects of Catalan provide an additional type of opaque output forms in clitic combinations that will be accounted for in a slightly different fashion, even though the two types of opaque forms sometimes interact and share crucial aspects. These opaque forms affect, in most cases, combinations of a third person accusative clitic and a third person dative clitic. (4c) illustrates these cases (in (4a) and (4b) the relevant clitics appear in isolation):

(4)a. Els llibres, **els** donaré a en Quim demà.
*the books* 3*rd-pl-acc will-give*(1*st*) *to the Quim tomorrow*

b. A en Quim, **li** donaré els llibres demà.
*to the Quim* 3*rd-dat will-give*(1*st*) *the books tomorrow*

c. Els llibres, a en Quim, [**ɔlzi**] donaré
*the books to the Quim* [*ɔlzi*] *will-give*(1*st*)
demà. (**els li/*li'ls*)
*tomorrow*
'I will give the books to Quim tomorrow'

One of the most puzzling aspects about (4c) is that, even though the output clitic form looks like the third person plural dative clitic in isolation, /lzi/ (the schwa in (4c) being epenthetic), the antecedent of the dative clitic in (4c) is singular, not plural (*a en Quim*). As will be shown later, the plural marker /z/ in the output form has to come from the accusative (plural) source (the antecedent being plural: *els libres* 'the books'). One of the expected possible transparent output forms, *li'ls*, appears in fact in varieties of the Valencian dialect of Catalan.



Finally, an additional fact from several dialects of Catalan that will have to be accounted for, and which also involves the third person dative clitic, is exemplified in (5c):

(5)a. De pomes, en donaré als nens demà.
*of apples gen will-give*(1*st*) *to-the children tomorrow*

b. Als nens, [**ɔlzi**] donaré pomes demà.
*to-the children* 3*rd-pl-dat will-give*(1*st*) *apples tomorrow*

c. De pomes, als nens, [**ɔlzɔni**] donaré
*of apples to-the children* [**ɔlzɔni**] *will-give*(1*st*)
demà. (*[**ɔlzin**]/*[***n*ɔlzi**])
*tomorrow*
'I will give apples to the children tomorrow'

As can be seen in (5b), the form of the third person plural dative clitic in isolation is /lzi/ (with schwa epenthetic). However, in (5c) this form is split by the genitive clitic /n/.

### 1.2. *Problems with Previous Accounts*

As mentioned earlier, the phenomenon of opaque forms in clitic combinations has received little attention in previous accounts. In (6a), below, I repeat the formulation of the Spurious *se* Rule, as given in Perlmutter (1971). This rule accounts for the appearence of *se* in examples like (1c). In (6b) I reproduce the rule proposed in Wanner (1977) to account for cases like (2c), from Italian:

(6)a. Spanish Spurious *se* Rule (Perlmutter (1971), p. 22 (10))

$$\begin{bmatrix}\text{Pro}\\ \text{III}\\ \text{Dative}\end{bmatrix}\ \begin{bmatrix}\text{Pro}\\ \text{III}\\ \text{Acc.}\end{bmatrix}$$

1 2 → *se*, 2

b. Italian *si* to *ci* Conversion (Wanner (1977), p. 117 (26))

si si → ci si

Leaving aside the fact that Perlmutter's analysis mixes morphological information with phonological information, the main problem with the

rules in (6a) and (6b) is that they fail to capture an important generalization stated below:

(7) Generalization: opaque output forms in clitic combinations always result in another clitic form, indicating a closed system.

That is, it is never the case that opaque forms constitute a random phonological sequence. The output form of the Spurious *se* Rule coincides with the form of the reflexive or impersonal clitic *se*, while the output of the *si* to *ci* Conversion Rule coincides with the first person plural clitic (and also with one of the locative clitics). The same can be said about the Catalan form in (3c), which is more complex because the two input clitics change and appear with the surface form of two other clitics of the language, the third person accusative clitic and the oblique clitic.

The generalization just made is not captured in the rules in (6a, b) because anything can be expressed with the formalism used. It would be as easy to write random sequences like *pa*, 2, in the output of (6a), or *gu si*, in the output of (6b). The proposal to be presented in this article captures the needed generalization while blocking derivation of arbitrary sequences like *pa* or *gu*.

Notice that the generalization in (7) would equally be lost if one were to assume that opaque forms are simply listed in the lexicon forming a single item with the other clitic(s) they occur with. Under this view one should also expect to find random sequences in clitic combinations. But this is never the case.[5]

## 2. The Proposal

The main claim of this section is that pronominal clitics constitute hierarchical structures of monovalent morphological features and that most opaque forms are obtained after the syntax and before the insertion of phonological information, through delinking and insertion processes that modify the internal structure of the clitics involved. An additional and crucial claim is that the order among pronominal clitics is not determined syntactically but it is accomplished through mapping to a template. In this sense, the present proposal shares many aspects of Perlmutter (1971).[6]

[5] As pointed out by an anonymous reviewer, an additional problem with the rules in (6) is that they offer no clue as to why such a phenomenon should exist in natural language.

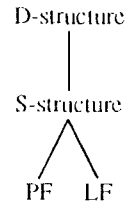
[6] See Simpson and Withgott (1986) for a somewhat different view of template morphology.



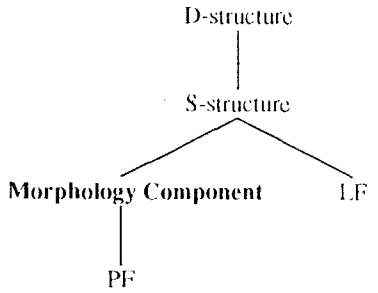
### 2.1. *Some Assumptions*

The model of the grammar in which I will frame my proposal is that of Principles and Parameters, which is the model shown in (8a).[7] However, I also assume a minor modification to this model proposed in Halle (1990, 1991), among others, which consists of the addition of a Morphology Component between S-structure and PF, shown in (8b):

(8)a.




b.




The Morphology Component takes as its input an S-structure that reflects hierarchical relations but not linear relations (cf. Marantz (1984, 1988)). Linear order, relevant only at PF, is established in the Morphology Component. Moreover, the syntax (D-structure, S-structure, LF) contains only morphosyntactic information. No phonological information is present at those levels, a view shared, for instance, by Otero (1976), Pranka (1983), and Anderson (1992). Phonological information is inserted late in the Morphology Component and provides the input to PF. The mismatches found between the morphosyntactic organization and the phonological organization (suppletion, multiple exponence, syncretism) are obtained in the Morphology Component.

I will follow Lumsden (1987) in assuming that S-structure contains fully specified syntactic feature matrices. On the surface, representations will be fairly impoverished with respect to these features, as will be discussed.

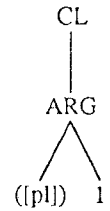
This paper, then, deals with the fate of pronominal clitics within the Morphology Component. At S-structure, pronominal clitics constitute nonlinearized sets of fully specified features; by PF they are impoverished, altered, linearized, and supplied with phonological information. A similar view of morphology can be found in Halle and Marantz (1993).

[7] By the term Principles and Parameters I refer, crucially, to a model of the grammar, not only a model of syntax. The proposal to be presented here might easily be adapted to other models, but such an enterprise is outside the scope of this paper.

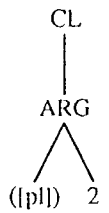
### 2.2. *The Internal Structure of Pronominal Clitics*

One of the main claims of this article is that pronominal clitics constitute hierarchical structures of unordered morphological features. For the most part they are abstract monovalent feature structures representing the morphological categories the clitics encode. These structures are mapped from a subset of the features, mainly $\phi$-features (in the sense of Chomsky 1981), that the clitics contain at S-structure. The projection of syntactic features onto morphological structures constitutes a first step toward the diverse morphological impoverishment (syncretism, multiple exponence, or the lack of overt morphology) that one finds across languages.[8] In (9) I give the structures corresponding to the pronominal clitics of Catalan, and afterwards I justify the use of these structures. Most Romance languages have basically the same (or fewer) structures.[9] For clarity, I also give the phonological form of the clitics (in proclitic position) that would correspond to the Barceloni dialect, discussed in Section 3.1.

(9)a. 1st person (sg, pl) (/m/, /nz/)

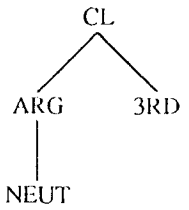



b. 2nd person (sg, pl) (/t/, /wz/)




c. Impersonal/reflexive . . . /s/

d. Neuter (/u/)

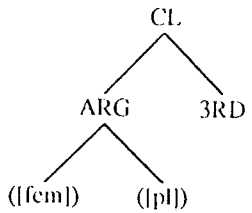



[8] Syntactic features might be contained in unorganized feature matrices or, alternatively, might be organized hierarchically. In either case, morphological impoverishment would involve the elimination of features in one of several ways.

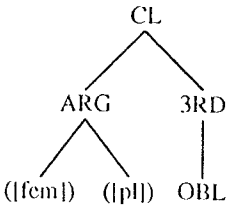
[9] The absence of some of these structures in certain Romance languages (e.g., (9g, h, i) in Spanish, or (9i) in certain dialects of Catalan) could be accounted for in one of two ways: either those forms are never generated in the syntax, or they are generated but they are lost in the mapping from S-structure to PF. I will not deal with this issue here.



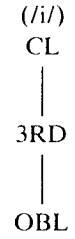
e. 3rd person acc. (sg, pl; masc-fem) (/l/-/lɔ/, /lz/-/lɔz/)



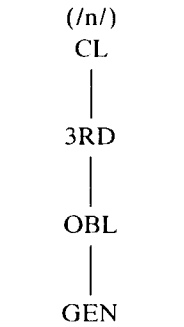

f. 3rd person dative (sg, pl) (/li/, /lzi/)



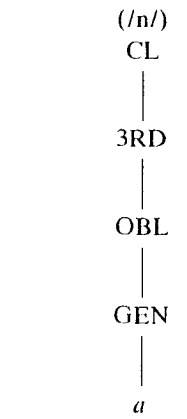

g. Locative, . . . (/i/)




h. Partitive, . . . (/n/)




i. Ablative (/n/)




The structures in (9) are similar, in many respects, to the internal structure of phonological segments, as proposed in Clements (1985), Sagey (1986), and McCarthy (1988), among others. Here, as in feature geometry, each structure has a hierarchical organization, and there is no linear ordering among elements at the same level. For instance, even though in (9f) I have written [ARG] before [3RD], there is no linear relation between them (at that point). They could have equally been written in the opposite order. The same could be said with respect to [feminine] and [plural] in

the same structure.[10] The relation among different nodes is in most cases one of dominance; [GEN], for instance, entails the presence of [OBL]. The opposite never holds. The node CL, which dominates each clitic structure, can be compared to the root node in phonological representations. The other labels in capital letters define natural classes. In fact, this is the main motivation for proposing the structures in (9): all the clitics sharing a given node pattern together in certain aspects, many of which will be discussed in later sections. There is, for instance, a major division between the clitics that have [ARG] and the clitics that have [OBL] in their structure. The latter have always been called 'adverbial clitics' in the traditional literature on Catalan, even though they often have little relation with adverbs.[11] The clitics that share the node [3RD], shown in (9d–i), also constitute a natural class. As will be shown later, [3RD] clitics are the only ones that undergo and trigger certain processes. The clitics that do not have this feature almost always appear in a transparent form; there is only one case in which a non-[3RD] clitic does not have a transparent form. Nodes like [plural] and [feminine] do not constitute defining properties of the clitics but are agreement features on them.

The feature hierarchy reflected in the structures in (9) constitutes a system of not binary, but privative or monovalent features. This is a basic difference between this system and the phonological feature geometry systems mentioned above. In those systems, all terminal features are binary. In the system argued for here, binary features have been rejected because there is no strong evidence for the existence of two values for any given feature. In addition, the hierarchical organization of morphological features also expresses markedness relations (the fewer number of nodes, the more unmarked). It has been observed, for instance (cf. Benveniste (1966) and Jakobson (1956/1971), among others), that third person is more unmarked than first or second person. (9c) corresponds to a bare reflexive or impersonal, unspecified for person (first or second). This clitic is morphologically defined as [ARG]. (9e), defined as [ARG 3RD] and corresponding to the third person accusative clitic, is also fairly unmarked. Agreement features are also privative or monovalent. The notion 'singular' is the lack of [plural], while 'masculine' is the lack of [feminine]. Arguments for this position, as opposed a position favoring binary features, can be found in Harris (1991a, b) and Bonet (1991).[12,13]

[10] A parallelism can be drawn between the structure in (9f) and the structure of complex segments in phonology. For instance, a labiovelar complex segment will have the articulator nodes [labial] and [dorsal] hanging from the place node. The fact that one label appears to the left of the other one does not imply that the labial articulation precedes the dorsal articulation or viceversa.

[11] The clitic corresponding to the third person dative belongs to both classes, since it has both [ARG] and [OBL]. The motivation for the presence of these two features in the structure of the third person dative clitic will become clear in Section 3.1.



The first person clitic constitutes an example of the morphological impoverishment reflected by the structures above. Even though a first person clitic can be syntactically dative or accusative in the syntax, reflexive or pronominal, masculine or feminine, these distinctions are neutralized in the Morphology Component for all Romance languages, where there is only one structure corresponding to first person (with the option of being plural or not).[14] Another, less clear, case of impoverishment of clitic structure is shown by the clitic *ne*. Some linguists (Elliott, 1986, Bartra, 1987) have suggested that, syntactically, there is only one clitic *ne*, which systematically receives Genitive Case regardless of the different positions it can occupy. The basic motivation for this proposal is to capture the fact that, at some level, there is a single form *ne*. Other linguists (Belletti 1988) have proposed that one of the instances of *ne* receives Partitive Case. If one were to take the latter approach, the "unification" of all instances of *ne* would arise in the Morphology Component in the mapping to the structures in (9).

A comment should be made about the terms chosen to define the natural classes introduced in this paper. These are *morphological* terms; although they suggest some syntactic category, they are defined within the theory proposed. [ARG], for instance, is an abbreviation for 'argument'. Although in many cases [ARG] clitics constitute syntactic arguments, this is not always the case (they can also be ethical datives, for instance). [ARG] is meant to suggest a certain relation with 'argument' but should not be identified with it. The same can be said of the terms [OBL]

[12] Of course, I could have advocated a system of binary features together with a theory of underspecification, where the non-active features would have been inserted late in the Morphology Component by default rules. The system of monovalent features here has been chosen because it is more restrictive and because it expresses markedness relations, as discussed in the text.

[13] Pronominal clitics are most probably not the only elements in the grammar with feature structures like the ones proposed in (9). Inflectional morphology presumably shares most of those structures, and the definite article might have the structure shown in (9e). However, such an extension of the present proposal is a much more ambitious project.

[14] As can be seen in (9a, b), first person has the specification [1], and second person has the specification [2]. In languages that can have inclusive first person plurals, both [1] and [2] would appear as sisters dominated by [ARG].

(cf. 'oblique'), [GEN] (cf. 'genitive') or [3RD] (cf. 'third person').[15] An objection that could be raised is that these natural classes are defined too vaguely (basically in terms of the behavior of the clitics that belong to them). However, this vagueness can also be found with other types of features. At the present time, the correlates of phonological features are more or less well-known, but this is not the case with syntactic features. The main motivation for proposing a feature system of any sort is that it allows us to account for the behavior of certain elements. In the present case, the goal is to account for certain morphological properties of pronominal clitics.

As will be shown in Section 3, most opaque forms are obtained through delinking or insertion rules that operate on the morphological structures represented in (9). Given that these structures are built from a closed set of features, it will follow that the forms resulting from delinking or insertion operations will always be identical to a possible clitic, and no random phonological sequences will ever be derived. The generalization expressed in (7), above, will thus be captured more naturally under this proposal.

### 2.3. *The Internal Organization of the Morphology Component*

In the previous section we have seen that the fully specified syntactic feature matrices of pronominal clitics are mapped onto the morphological structures represented in (9) in the Morphology Component.[16] Most opaque forms are obtained through morphological rules, i.e., delinking and insertion processes that modify the internal structure of the clitics (see Section 3).

After the application of morphological rules, morphological structures or their defining features are mapped onto a template that will determine

[15] I do not mean to imply that there is no relation between morphological features and syntactic features. On the contrary. It could very well be the case that the features relevant to the morphology are a subset of the features present in the syntax (cf. footnote 8). The labels used in this paper for some of the features have been chosen because of the similarities they have with syntactic features and in spite of the asymmetries between the two types. A more refined theory of syntactic and morphological features is needed to establish what the exact correlation is between them.

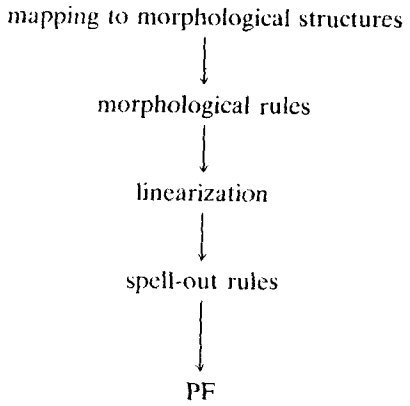
[16] In order to see exactly how this mapping works, more should be known about syntactic features. See also footnote 8.



the surface order of the clitics.[17] At that point, since two morphological structures or two features may compete for the same slot, other opaque forms (like (4c), above) will be obtained. Finally, a set of spell-out rules will provide morphological features with the relevant phonological information. This operation will provide the input to PF.

The internal structure of the Morphology Component is schematized below:

(10) **The Morphology Component:**




The different processes that take place within the Morphology Component are discussed in the following sections.

## 3. Morphological Rules

Morphological rules are insertion and delinking processes that modify morphological structures. Similar processes are found in phonology. In this section I will give examples of both types of operations, which will account for the examples in (1c), (2c), and (3c), among others. In the section on the spurious *se* of Spanish, we will also have to invoke a notion

[17] The surface order among pronominal clitics cannot be determined syntactically. This is especially clear when languages with a rich pronominal clitic system are considered. In Catalan, for instance, it cannot be said that accusative clitics precede dative clitics or viceversa, or that clitics related to argument positions precede or follow clitics related to adjuncts or to inherent clitics. The only generalizations that can be made for certain clitics are in terms of person: second person clitics always precede first person clitics. Moreover, one can find many differences in clitic order among dialects, but these differences do not seem to correlate with any syntactic differences. Arguments against a syntactic derivation of clitic order can be found in Perlmutter (1971) and Bonet (1991).

of structure preservation, which receives support from several American dialects of Spanish.

### 3.1. *Barceloní (Catalan)*

All the examples from Catalan given so far can be found in Barceloní, which is very rich in opaque forms. Some of the data I will discuss here can be found in only a subset of speakers of Barceloní.[18]

I repeat below (as (11)) the examples from Catalan that were given in Section 1.1 as (3):

(11)a. Això, **ho** trauré de l'armari després.
*this neut. will-take-out(1st) from the closet later*

b. De l'armari, **en** trauré això després.
*from the closet ablat. will-take-out(1st) this later*

c. Això, de l'armari, **[li]** trauré
*this from the closet [li] will-take-out(1st)*
després. (**n'ho/*ho'n*)
*later*
'I will take this out of the closet later'

As commented on earlier, it can be seen in (11a) that the phonological form of the neuter clitic in isolation is *ho* (/u/) and that the form of the ablative clitic in isolation is *en* (/n/). In (11c), which combines both clitics in the same sentence, we do not find the transparent form /n/ + /u/ (or /u/ + /n/) but a completely different output: [li]. The segment /l/ of this output is homophonous with the third person accusative clitic (in this system [ARG 3RD]), while /i/ is homophonous with the oblique clitic ([OBL]).

There are cases where these "changes" apply independently of one another. The example in (12c) shows the opaque output form of an ablative source, which becomes homophonous with the oblique /i/. In this case the ablative clitic *en* is combined with the genitive clitic *en*. This process can also be found in dialects of Italian. (12a) and (12b) show the clitics involved in (12c) as they appear in isolation.

(12)a. Del teatre, **en** sortiran tres homes.
*from-the theater ablat. will-exit(3rd-pl) three men*

b. D'homes, **en** sortiran tres del teatre.
*of men gen. will-exit(3rd-pl) three from-the theater*

c. D'homes, del teatre, **[ni]** sortiran tres. (**ne'n*)
*of men from-the theater [ni] will-exit(3rd-pl) three*
'three men will exit from the theater'

If we only had examples like (12), we would not be able to tell which one of the two instances of *en* surfaces as /i/ in (12c). In order to decide, we need examples like (11c) and, even better, examples like (13), below. It can be seen, from (13c), that when the ablative clitic *en* is combined with a third person accusative clitic, /l/ ([ARG 3RD]), the ablative clitic also shows up as /i/:

(13)a. El jersei, **el** trauré del calaix
*the sweater 3rd-acc will-take-out(1st) from-the drawer*
aquesta tarda.
*this afternoon*

b. Del calaix, **en** trauré el jersei
*from-the drawer ablat. will-take-out(1st) the sweater*
aquesta tarda.
*this afternoon*

c. El jersei, del calaix, **[li]** trauré
*the sweater from-the drawer [li] will-take-out(1st)*
aquesta tarda. (**l'en/*ne'l*)
*this afternoon*
'I will take the sweater out of the drawer this afternoon'

We can conclude, then, that in (12c) the *en* that surfaces as /i/ is the ablative, given that this is the clitic affected in the other cases (i.e., (11c) and (13c)).

(12c) and (13c) illustrate how one of the two "changes" that take place in (11c) operates in isolation (namely *en* → *hi*). (14c), below, illustrates

[18] Another important subset of speakers differ minimally from the subset described here. They require a slight modification of rule (17a), presented below in the text.

Although I use the term Barceloní for the variety discussed here, it must be said that the speakers of this variety come mainly from the northern part of the Barcelona area. The Catalan spoken in the Barcelona area is in general quite heterogeneous and often differs according to age and other factors.

the other change found in (11c), also operating in isolation (*ho* → *el*, in this case). Here, the clitics that enter the combination are the neuter clitic *ho* /u/ and the locative clitic *hi* /i/.

(14)a. Això, **ho** portaré a Sabadell demà.
*this neut. will-take*(1*st*) *to Sabadell tomorrow*

b. A Sabadell, **hi** portaré això demà.
*to Sabadell loc. will-take*(1*st*) *this tomorrow*

c. Això, a Sabadell, [li] portaré demà. (**hi ho/*ho hi*)
*this to Sabadell* [*li*] *will-take*(1*st*) *tomorrow*
'I will take this to Sabadell tomorrow'

As can be seen in (14c), the neuter clitic *ho* becomes homophonous with the third person accusative clitic *el*, that is [ARG 3RD].

Consider finally (15c), below, where the neuter clitic *ho* /u/ has disappeared. This clitic has been combined with a third person dative clitic:

(15)a. Això, **ho** donaré a l'Oleguer pel
*this neut. will-give*(1*st*) *to the Oleguer for-the*
seu aniversari.
*POSS.*(3*rd*) *birthday*

b. A l'Oleguer, **li** donaré això pel
*to the Oleguer 3rd-dat will-give*(1*st*) *this for-the*
seu aniversari.
*POSS.*(3*rd*) *birthday*

c. Això a l'Oleguer, [**li**] donaré pel
*this to the Oleguer* [*li*] *will-give*(1*st*) *for-the*
seu aniversari.
*POSS.*(3*rd*) *birthday*
'I will give this to Oleguer for his birthday' (**li ho/*ho li*)

For this case, we cannot say that the neuter clitic /u/ has undergone a phonological process of truncation due to the presence of an adjacent vowel (/i/, coming from the dative clitic in the sequence /li/ + /u/) since this process is not independently motivated in Catalan. However, given

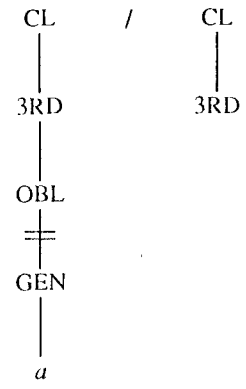


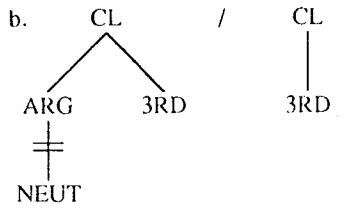
examples like (14c) and (11c), where the neuter clitic surfaces as /l/, we can extend the account we need for those cases to (15c). This would yield /l/ + /li/ or /li/ + /l/. As will be shown, the final [li] in (15c) is the product of a morphological rule plus the process that affects combinations of two third person clitics, exemplified in (4c), which will be addressed in Section 4.

Below I summarize the opaque output forms that have to be accounted for so far. To the left of the arrow I give the input clitic combinations, using the minimal number of nodes that identify them. (16) gives the cases in the reverse order used for the examples. As can be seen in (16c–d), /n/ "changes" to /i/ only when it is an ablative, that is, when it has the structure {[GEN], *a*}, which is slightly (but crucially) different from the structure of the other uses of /n/ (bare [GEN]):

(16)a. {[ARG], [OBL]}/l(z)i/ + [NEUT]/u/ → /l(z)i/
b. [OBL]/i/ + [NEUT]/u/ → /li/
c. [ARG 3RD]/l(ə)(z)/ + {[GEN], *a*}/n/ → /li/
d. [GEN]/n/ + {[GEN], *a*}/n/ → /ni/
e. {[GEN], *a*}/n/ + [NEUT]/u/ → /li/

The five opaque forms given in (16) are accounted for by the following two morphological (delinking) rules:

(17)a. 


b. 


Through the application of rule (17a), the structure of the ablative clitic becomes identical to the structure of the locative clitic *hi* (cf. (9g)). In the case of (17b), the structure of the neuter clitic becomes identical to the

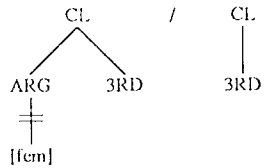
structure of the third person pronominal clitic *el* (cf. (9e)). Notice that, in these two rules, both the trigger and the target have the feature [3RD]. All the clitics involved in these rules belong to the natural class [3RD]. First person, second person, and reflexive or impersonal clitics, which lack this feature, never serve as triggers or targets for these rules.[19]

In (18) I give the derivations of the cases in (16), following the same order. I omit the agreement features [feminine] and [plural] because they are not relevant here.[20] Even though I use the phonological content, for convenience, to refer to the clitics involved and to the result obtained, it has to be kept in mind that at that point in the Morphology Component there is no phonological information. This information is introduced by spell-out rules very late within this component. The morphological structures corresponding to all Catalan clitics, together with their phonological form, can be found in (9), above.

[19] An anonymous reviewer suggests that all morphological rules could be conceived as delinking operations (in Section 3.2 a morphological rule is proposed which consists of the insertion of a feature). Given the data shown in this paper, this could certainly be a possibility worth considering. However, one can find cases where a feature that is deleted in one language or dialect is added in another one. For instance, as shown in (14c), (15c) and the rule in (17b), in Barceloní the feature [NEUT] is delinked in certain cases. On the other hand, in varieties of Majorcan and Valencian, the feature [NEUT] is added to the morphological structure of the [ARG 3RD] clitic in combinations of two third person clitics. (One could alternatively see the Majorcan and Valencian case as a delinking operation, but then the Barceloní case would have to be an insertion operation.) None of these cases can be accounted for through phonological processes.

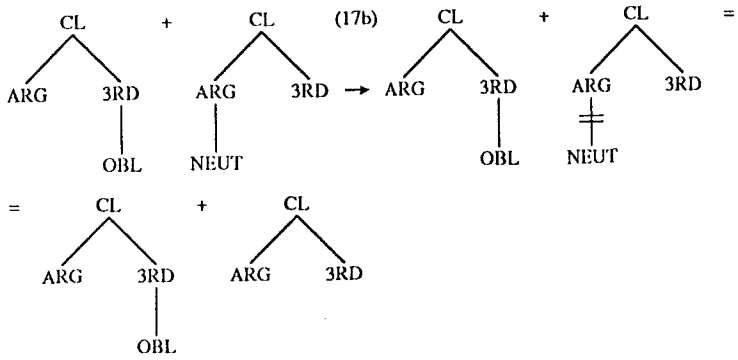
[20] In Barceloní, the feminine marker disappears whenever the accusative third person clitic is combined with any other [3RD] clitic, but not with non-[3RD] clitics. This deletion can be accomplished by an additional morphological delinking rule that would be formulated as shown in (i):

(i)

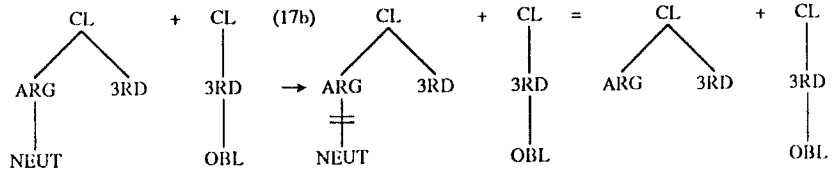




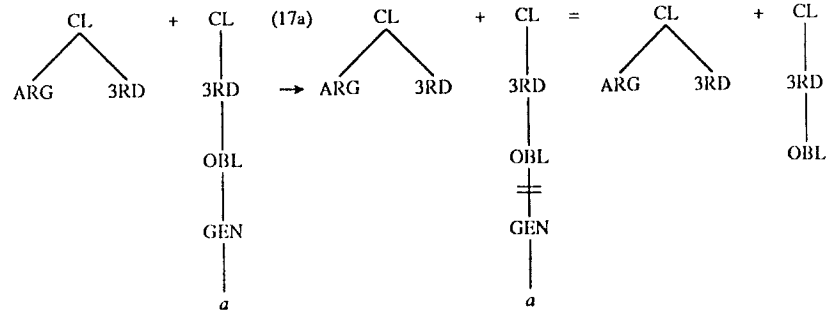

(18)a. /li/ + /u/ → /li/ + /l/(→/li/)



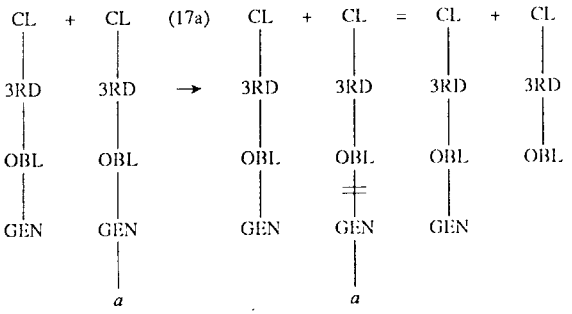

b. /u/ + /i/ → /li/



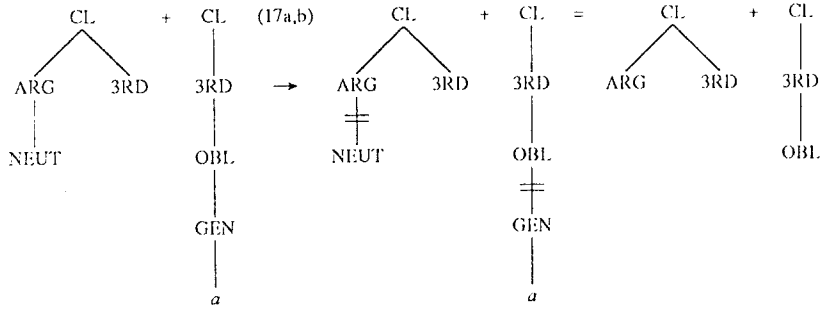

c. /l/ + /n/ → /li/



(18)d. /n/ + /n/ → /ni/




e. /u/ + /n/ → /li/




Notice that in (18a) the delinking of the node [NEUT], through rule (17b), renders the morphological structure of this clitic identical to the structure of a third person accusative clitic (cf. (9e)). This combination becomes thus identical to the combination of two third person clitics (accusative and dative), a type of combination discussed in Section 4. Therefore, the form obtained in (18a), after the application of rule (17b), does not constitute the end of the derivation of the combination of a neuter clitic with a third person dative clitic.

As can be seen in (18e), the complete lack of transparency in the output form is due to the fact that the two rules in (17) are applicable (in any order), thus altering the internal structure of the two clitics involved. The neuter clitic (/u/ in isolation) becomes identical to the third person accusative clitic (realized phonologically as /l/), and the ablative clitic (/n/ in isolation) becomes identical to the oblique clitic /i/. These rules apply independently of one another in (18a, d).[21]

[21] The derivations given in (18) show that the five opaque output forms are not arbitrary. They are all obtained through the application of two different rules. If the output forms were in fact arbitrary one could expect random, not related, sequences in each case. However, it is not a coincidence, for instance, that the combination of a neuter clitic and an ablative clitic results in the opaque sequence /li/; this sequence is the result of two independently motivated rules (cf. (18c)).



### 3.1.1. *A Note on Filters Barring Phonologically Identical Sequences*

In the literature on clitics, one can often find cooccurrence restrictions that are supposed to rule out sequences of phonologically identical clitics. This is, for instance, the option taken by Wanner (1977) for the non-occurring Standard Italian sequences **ne ne*, **ci ci*, **si si* and **vi vi*. According to him, this absence "shows that clitics are dependent on phonological information even for what concerns the syntactic problem of their surface ordering and cooccurrence restrictions" (Wanner (1977), p. 123).

Even though it is true that one tends not to find sequences of phonologically identical clitics, it is not impossible to find them. For example, in the Italian dialect spoken in Conegliano, sentence (2c), which I repeat below as (19a), is actually realized as (19b):

(19)a. **Ci si** lava.
*ci si washes*
'one washes oneself'

b. **Si si** lava.
*si si washes*
'one washes oneself'

In (19b) there is a sequence of two phonologically (and phonetically) identical clitics, and the sentence is perfectly fine. Some varieties of Valencian provide a further example, involving third person clitics. In Valencian, the third person plural dative clitic becomes homophonous with the third person plural (masculine) accusative clitic. In (20a) there is no phonological identity (even though the clitics are fairly similar in this respect) because the accusative clitic is feminine. In (20b), with no feminine markers, the two clitics involved become identical phonologically:[22]

[22] In this particular example, (20b), the clitics are identical phonologically, not phonetically ([elzels]).

(20)a. Als xiquets, les coses, **els** **les**
*to-the boys* *the things* *3rd-pl-dat* *3rd-pl-fem-acc*
portaré després.
*will-take*(1*st*) *later*
'I will take the things to the boys later'

b. Als xiquets, els llibres, **els** **els**
*to-the boys* *the books* *3rd pl-dat* *3rd-pl-acc*
portaré després.
*will-take*(1*st*) *later*
'I will take the books to the boys later'

In this variety of Valencian, the combination of two third person clitics, accusative and dative, always gives transparent results. This is not the case in Spanish (which uses the spurious *se* in precisely this context) or in other dialects of Catalan.[23]

We could have tried to account for the non-occurrence of /n/ + /n/ in Barceloní, shown in (12), with the use of a filter against phonologically identical sequences.[24] The filter could have been expressed as in (21):

(21) */n/ /n/

Notice, however, that we need rule (17a) (which forces the ablative to "look like" the oblique) to account for the opaque forms described in (16c) (a third person accusative with an ablative) and (16e) (a neuter with an ablative). Rule (17a) can perfectly account for the non-occurrence of */n/ + /n/ because it forces the sequence of two instances of [GEN] (one of them also with the feature *a*, the ablative) to ultimately become [ni]. It is important to note that if rule (17a) had not applied in (16c) and (16e), no phonologically identical sequences would have emerged. The output of the combination in (16c) would have been [lɔn] (found in certain dialects) and the output in (16e), [nu] (or [un]).

Given that rule (17a), required independently for (16c) and (16e), can account for the non-occurence of */n/ + /n/ (as well as for the actual output), the filter in (21) is unnecessary. The analysis proposed with rules (17a, b) is superior to a phonology-based analysis because it accounts for

[23] As an anonymous reiewer points out, sequences of phonologically identical clitics can also be found in French and in German, for instance.

[24] The sequence /n/ + /n/ would be realized phonetically as [nɔn], which can actually be found in certain varieties of Catalan.



all the opaque cases in a uniform manner, and not with arbitrary restrictions on phonological form. If there were no examples of sequences of phonologically identical clitics one could perhaps posit some universal constraint against such sequences (this is what Wanner seems to be proposing, at least uniformly for Italian). However this is not the case, as shown by the examples in (19) and (20).

The type of filters discussed in this section face another problem: it is very common in phonology not to find adjacent phonologically identical segments. These cases are usually accounted for by the Obligatory Contour Principle (see, for instance, McCarthy 1986). However, as far as I know, one cannot find this type of restriction among *sequences* of segments, where the phonologically identical segments are not adjacent. The clitic cases discussed here consisting of sequences of segments would be very difficult to account for in phonological terms. These filters are easy to describe in prose, but they are extremely difficult to formalize.[25] Perlmutter (1971) gives an additional argument against the use of filters blocking sequences of phonologically identical clitics. He shows that for Spanish these filters should also rule out sequences with only partial phonological identity, because sequences like *le la*, *le los*, *les la*, and many others never occur. This idea would be very difficult to capture.[26]

### 3.2. *Standard Italian*

I will now turn to (2c), which is repeated below as (22):

(22)a. Lo **si** sveglia.
*3rd-acc impers.* *wakes-up*
'one wakes him up'

[25] An anonymous reviewer suggests that the filters that rule out sequences of phonologically identical clitics are pertinent, for instance, to the barring of *-ing . . . -ing* sequences in English (cf. Ross (1972), for example). Under the view defended in this paper, the English restriction could also be considered a morphological restriction barring identical morphological features, instead of a problem related to phonologically identical sequences (if the English gerund presented allomorphy, the choice between these two possibilities would be straightforward). A phonological account of the *-ing . . . -ing* restriction would face the formalization problem mentioned in the text.

[26] Rules barring sequences of identical phonological elements and morphological rules make different predictions, leaving aside the formalization problem, pointed out earlier in the text: the former should rule out any sequences of identical elements (for instance, any sequence *si si* in Italian), while the latter look only at the morphological (not phonological) makeup of the elements involved. Given that it is not uncommon to find sequences of identical syllables across languages, the idea of a phonological filter seems rather inadequate.

(22)b. Se lo compra.
*refl. 3rd-acc buys*
's/he buys it for herself/himself'

c. Ci si lava. (**si si*)
*ci si washes*
'one washes oneself'

Even though the phonological form of the impersonal clitic and the third person reflexive clitic in isolation is in both cases *si* (as shown in (22a, b)), when these two clitics appear in the same sentence, one of them surfaces as *ci*, a phonological form that coincides with the form of the first person plural clitic and the locative clitic.

In order to account for this opaque form, let us note two peculiarities concerning the impersonal clitic, peculiarities which are not found in other Romance languages, as far as I know. The first of these peculiarities can be observed comparing (22a) and (22b). Both sentences contain a third person accusative clitic *lo*. What is interesting about these examples is that the two instances of *si* occupy two different positions with respect to *lo*: the impersonal clitic *si* follows the third person accusative clitic (cf. (22a)), while the third person reflexive *si* precedes it (cf. (22b)).

The other peculiarity that seems to distinguish Italian from other Romance languages is that impersonal *si* triggers plural agreement. The following example has been reproduced from Cinque (1988) (it corresponds to his (25b)). The plural elements are written in boldface:

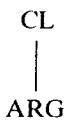
(23) Si è stati abbandonati a se stessi.
'one has been(pl, masc) abandoned(pl, masc) to oneself(pl, masc)'

This type of evidence led Cinque (1988) to conclude that the impersonal clitic is morphologically specified as plural.[27] This idea can be incorporated into the present proposal by saying that the morphological structure of the impersonal clitic is as shown in (24a). The structure of the impersonal clitic can be compared to the structure of the third person reflexive clitic (in (24b)), which lacks agreement features:

[27] As Rizzi (1986) points out, all *arb* elements in Italian have plural number.

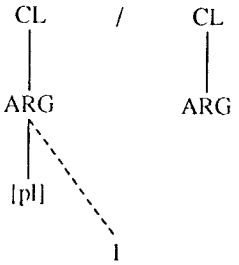


(24)a.


b. 


Often third person reflexives and impersonals share the structure in (24b). This is the case in Catalan and Spanish, for instance. As shown by the example in (23), there is evidence for the presence of a feature [plural] in the morphological structure of the impersonal clitic in Italian. The structures in (24a) and (24b) occupy different slots in the template that takes care of their linearization. This different ordering is reflected in (22a) and (22b).

Once the structure in (24a) is assumed for the impersonal clitic, the change from *si* to *ci* is fairly easy to account for. The rule responsible for the change, a two-step process, inserts the feature [1] to the structure in (24a) and links it to the impersonal clitic when this clitic is combined with the bare [ARG] clitic. This is shown in (25):[28]

(25) 


Once the rule has applied, the structure corresponding to the impersonal clitic becomes identical to the structure of the first person plural clitic. This structure is later spelled out as *ci*. We can then say that, in the *ci si*

[28] The formulation in (25) is defective because, as it is stated, the insertion rule should operate in the context of any [ARG] clitic (then the rule would apply in the context of a second person clitic or a first person clitic, for instance, which is not the case). Rule (25), like some other morphological rules, probably has a deeper motivation than is suggested in the text: a restriction parallel to the Obligatory Contour Principle (OCP) in phonology (cf. McCarthy 1986, among others), which would bar sequences of identical morphological structures or morphological features, as was suggested in passing in footnote 25. Then, the insertion of [1] in (25) would be some sort of morphological dissimilation rule triggered by the morphological quasi-identity of the two clitics. I leave this issue for further research.

sequence, *ci* corresponds to the impersonal clitic, not to the third person reflexive clitic.

It was said earlier that the form *ci* corresponds to the first person plural clitic but also to the locative clitic. One might wonder why *si* does not become identical to the locative clitic instead of becoming identical to the first person plural clitic. Under the present proposal, the change from the structure of *si* to the structure of the locative would involve more operations than the change presented above. Given that the structure of the locative is defined as bare [OBL] (cf. (9g)), the operations would have to be (1) delinking of [ARG] and (2) insertion and linking of the node [OBL] (together with the node it dominates, [3RD]). This derivation would evidently be more costly than one unique operation, namely the insertion of [1], postulated above. Moreover, the locative clitic in Italian has two possible spell-outs: *ci* and *vi*. If the change of *si* to *ci* were a change to [OBL] (the locative), one might expect to also find the sequence *si vi* in combinations of two *se* clitics. This is never the case.

### 3.3. *Spanish*

This section is devoted to the spurious *se* in Spanish and some of its consequences. We will see that a crucial difference between Spanish and Italian with respect to the bare [ARG] clitic triggers an additional process in certain American dialects of Spanish.

#### 3.3.1. *The Spurious se Rule*

The Spurious *se* Rule is responsible for the opaque output in examples like (1c), repeated here as (26c). (26a, b) show the use of the clitics involved in (26c) when they appear in isolation:

(26)a. El premio, **lo** dieron a Pedro ayer.
*the price 3rd-acc gave(3rd-pl) to Pedro yesterday*

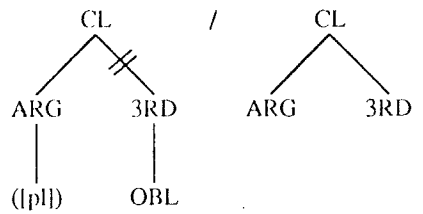
b. A Pedro, **le** dieron el premio ayer.
*to Pedro 3rd-dat gave(3rd-pl) the price yesterday*

c. A Pedro, el premio, **se lo** dieron
*to Pedro the price se 3rd-acc gave(3rd-pl)*
ayer. (**le lo/*lo le*)
*yesterday*
'they gave the price to Pedro yesterday'



By the Spurious *se* Rule the third person dative clitic (*le(s)*) becomes identical to the reflexive or impersonal clitic (*se*). The Spurious *se* Rule delinks the node [3RD] with its dependent [OBL] from the structure of the third person dative clitic. This is shown in (27), below:

(27)




With the application of the rule in (27) an illegitimate object is created because bare [ARG] clitics in Spanish are incompatible with agreement features. I give below the structure of the bare [ARG] clitic:

(28)

Therefore, because of structure preservation, once the structure in (27) is created [plural], an agreement feature incompatible with bare [ARG], is delinked and later deleted by Stray Erasure.[29] Stray Erasure, a convention that is commonly used in phonology, states that all the features that are not linked by the end of the derivation are deleted because they are uninterpretable (see, for instance, Kiparsky (1985)). Then the output of rule (27) will effectively be identical to (28) and therefore will be spelled out as *se*.[30]

In the section below I present data from certain dialects of Spanish showing that some of the features left floating because of the conflict created by the Spurious *se* Rule can be reassociated and show up on the accusative clitic instead.

[29] The intuitive idea behind the statement that bare [ARG] is incompatible with agreement features is that the mapping to the Morphology Component creates an inventory of impoverished morphological structures which cannot be altered to create new (less impoverished) structures out of this inventory. A convention could be invoked stating that morphological structures cannot be enriched within the Morphology Component (enriched in the sense of recovering types of features lost in the mapping to that Component).

[30] The "change" from *le* to *se* by the Spurious *se* Rule, which could be viewed as a dissimilation rule related to the OCP (mentioned in footnote 28), is the minimal possible change in Spanish, given the inventory of clitics (as morphological structures) in this language (five basic structures). All other changes (such as from *le* to *me*) would have to involve the insertion of additional features and the delinking of some.

### 3.3.2. *The Spurious se Rule and Agreement Features*

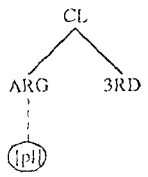
In (29b) I give an example from several American dialects of Spanish that can be compared to the corresponding version in Iberian Spanish in (29a). In (29b) I give in boldface the crucial segment:[31]

(29)a. El libro, a ellos, ¿quién se lo prestó?
*the book to them who se 3rd-acc lent(3rd)*
'who lent the book to them?'

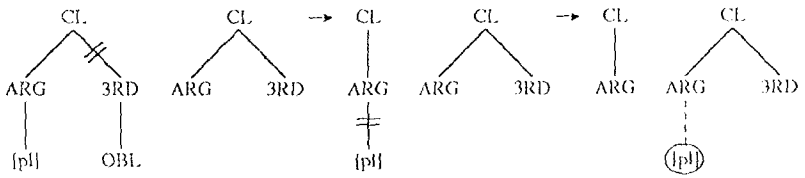
b. El libro, a ellos, ¿quién se los presto?

In (29) the antecedent of the indirect object is plural (*ellos* 'them'), while the antecedent of the direct object clitic is singular (*el libro* 'the book'). In (29a), from Iberian Spanish, the accusative clitic is correspondingly singular (*lo*). In many American dialects, however, the accusative clitic surfaces with a plural marker (*los*), in spite of the source being singular. This plural marker can only come from one place: the third person dative clitic, which has become *se* by the Spurious *se* Rule. This phenomenon can be captured by the rule in (30a). The whole process – application of the Spurious *se* Rule plus the relinking of floating features – is illustrated in (30b). The circle around [pl] indicates its floating status after being delinked:

(30)a.




b.




As shown in (30b), the floating agreement feature [plural] is linked to the

[31] Most of the data in this section come from the handout of a talk given by C. Company at the Universitat Autònoma de Barcelona and given to me by Gemma Rigau. These data have been checked with Hamila Cuna-Stainton. Similar data are given in Kany (1951).



[ARG] node of the structure corresponding to the third person accusative clitic. The resulting structure will be then spelled out as *los* (like a third person plural accusative clitic). When the structure corresponding to the accusative clitic is already plural the association process applies vacuously.

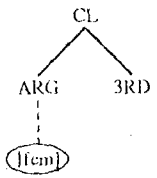
In dialects like colloquial Mexican and Uruguayan, this "transfer" of agreement features also takes place if the dative source is feminine. This is shown in (31b). (31a), again, contains the Iberian Spanish version:

(31)a. Si ella me quiere comprar el caballo, yo se
*if she 1st-dat wants buy the horse I se*
lo venderé.
*3rd-acc will-sell(1st)*
'if she wants to buy my horse, I will sell it to her'

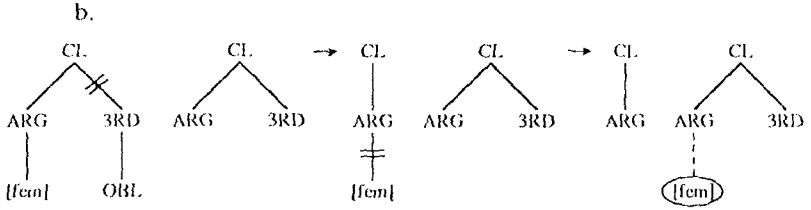
b. Si elle me quiere comprar el caballo, yo se
*if she 1st-dat wants buy the horse I se*
la venderé.
*3rd-acc-fem will-sell(1st)*
'if she wants to buy my horse, I will sell it to her'

In (31b), the source for the accusative clitic is masculine singular (*el caballo* 'the horse') but the clitic surfaces with the feminine marker *a*. Similarly to (29b), this marker has to come from the dative clitic, whose source is feminine (*ella* 'she'). The rule necessary to account for (31b), very similar to (30a), is given in (32a). I also show the whole process in (32b):

(32)a.




b.



Notice that, for at least the dialects under discussion, we have to postulate that the dative clitic can have the morphological feature [fem], even though no feminine marker ever appears on the dative clitic itself.

The rules in (30a) and (32a) have both applied in the example in (33b). This example can be compared with (33a), from Iberian Spanish:

(33)a. Si ellas me quieren comprar el caballo, yo
*if they(fem) 1st-dat want(3rd-pl) buy the horse I*
se lo venderé.
*se 3rd-acc will-sell(1st)*
'if they want to buy my horse, I will sell it to them'

b. Si ellas me quieren comprar el caballo, yo
*if they(fem) 1st-dat want(3rd-pl) buy the horse I*
se las venderé.
*se 3rd-acc-fem-pl will-sell(1st)*
'if they want to buy my horse, I will sell it to them'

The source of the dative clitic ("changed" to *se*) is feminine plural (*ellas* 'they'). The features [feminine] and [plural] appear, by the application of the rules in (30) and (32), on the accusative clitic in (33b), whose source is singular masculine (*el caballo* 'the horse').

The data discussed in this section provide strong support for the analysis presented in this paper. The appearence on the accusative clitic of the agreement features belonging to the dative clitic follows naturally from the analysis of the spurious *se*: the Spurious *se* Rule facilitates the relinking of the floating agreement feature to an appropriate target. The "transfer" of features is thus a consequence of the application of the Spurious *se* Rule. This "transfer" of features would be very difficult to account for if one were to try to give, for instance, a syntactic analysis of the spurious *se*. No available syntactic mechanisms would derive this "transfer" of features. Under approaches stating that clitic sequences are unanalyzable lexical items, the connection would be missed between the presence of the spurious *se* and the "transfer" of features, for example.

## 4. Mapping to a Template: More Opaque Forms

The mapping to a template can work in one of two ways: either each morphological structure is assigned to a slot (probably the most common case), or only the most specific defining features of a structure are assigned



to a slot. Barceloní, the dialect of Catalan that will be discussed in this section, illustrates both possibilities: the mapping of features affects [3RD] clitics, while the mapping of structures affects the other clitics. As will be shown, then, a third person dative clitic ([ARG OBL]), which is a [3RD] clitic, will occupy two slots of the template: the slot for [ARG] (also for third person accusative clitics) and the slot for [OBL] (also for locative clitics).

Below I give the template for Barceloní:

(34)

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| CL | CL | CL | [ARG] | [GEN] | { [OBL] / [NEUT] } |
| ARG | ARG | ARG | | | |
| | 2 | 1 | | | |

At the end of this section I will give an example in which all the slots in the template have been filled.[32]

In (34), the node [ARG] appears in two kinds of slots: the slots that require a whole clitic (slots 1–3) and the slot that allows only for the mapping of a terminal feature (slot 4). When the mapping takes place, the Elsewhere Condition, originally due to Panini (see Kiparsky 1982, for instance), will ensure that slots 1–3 have precedence over slot 4. Then, a bare [ARG] clitic (*es*) will not be mapped onto slot 4 because there is another slot, with more specific information, that matches exactly the structure of the bare [ARG] clitic, namely slot 1. Moreover, a third person accusative clitic ([ARG 3RD]) will not be mapped onto slot 1 because it does not match exactly the structure required to fill that slot (it also has the feature [3RD], not present in slot 1).

The two mapping systems impose different restrictions when there are two possible targets for one and the same slot. On the one hand, the mapping of whole clitics (slots 1–3) allows for one and only one clitic in each slot. When there are two possible targets for a given slot the derivation crashes: no grammatical output is possible. On the other hand, the mapping of terminal features does not impose any restrictions on the number of features that can be assigned to a given slot. The output will always be possible. Let us examine these possibilities in more detail.

[32] Even though in Barceloní all [3RD] clitics follow all non-[3RD] clitics, this is not the case in other dialects of Catalan (e.g., Balearic).

The form *es* can correspond to an impersonal clitic or to a third person reflexive clitic, among others. As shown by the examples in (35), two instances of *es* (one reflexive, the other one impersonal) are impossible (cf. (35a)), and a single *es* cannot represent both an impersonal clitic and a reflexive clitic (cf. (35b)). If one wants to say something like 'one washes oneself', either the reflexive must be expressed by a strong pronoun (cf. (35c)), or the impersonal must appear in a strong form (cf. (35d)):

(35)a. *Es es renta.
*es es washes*

b. *Es renta.
*es washes*

c. **Es** renta **a un mateix**.
*es washes to one self*

d. **Hom es** renta
*imp. refl. washes*
'one washes oneself'

(35a) would never be generated because there is only one slot for the bare [ARG] clitic (slot 1). This slot is filled in (35c) and (35d) (there is only one target for slot 1). In (35b) there is only one clitic *es*. As mentioned earlier, this clitic cannot be understood as both a reflexive and an impersonal. Under the present proposal, (35b) is ruled out because of the restriction on the mapping of whole clitics: even if one of the two clitics is mapped, the other one is not, and the derivation crashes.

Let us now consider some cases in which terminal features coming from different [3RD] clitics compete for one and the same slot. The most common case where this issue comes up is in combinations of two third person clitics, accusative and dative. These two clitics contain the node [ARG], which is mapped to slot 4 in the template (cf. (34)).

I repeat below, as (36), the examples I gave in (4), with a combination of an accusative third person clitic and a dative third person clitic. As usual, I first give the clitics involved in isolation (ignore the schwas):

(36)a. Els llibres, **els** donaré a en Quim demà.
*the books 3rd-pl-acc will-give*(1*st*) *to the Quim tomorrow*

b. A en Quim, **li** donaré els llibres demà.
*to the Quim 3rd-dat will-give*(1*st*) *the books tomorrow*



c. Els llibres, a en Quim, [ɔlzi] donaré
*the books to the Quim* [ɔlzi] *will-give*(1*st*)
demà. (**li'ls*/**els li*)
*tomorrow*
'I will give the books to Quim tommorrow'

A transparent output form for this combination is found in varieties of Valencian: *li'ls*. In Barceloní (in fact, in Central Catalan), the opaque form /lzi/ is phonologically identical to the form of the third person plural dative clitic in isolation. However, we cannot say that only one clitic surfaces in (36c), and that this clitic is the dative clitic, because the output form contains the plural marker /z/, while the source for the dative clitic is singular (*a en Quim* 'to Quim'). The plural marker can only come from the accusative clitic, which is coindexed with the plural argument *els llibres* 'the books'.[33] Below I give a table showing all the output forms in combinations of two third person clitics in Barceloní. Notice that the plural marker /z/ surfaces whenever one of the input clitics (or both) is plural (d = dative, a = accusative, s = singular, p = plural, m = masculine, f = feminine):[34]

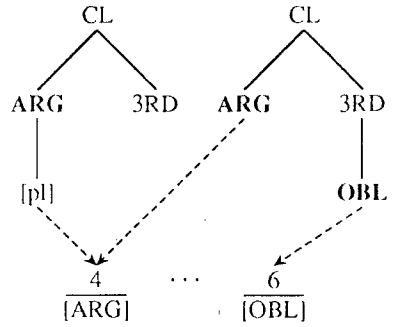
(37)
| | |
|---|---|
| ds + ams: | /li/ |
| ds + afs: | /li/ |
| ds + amp: | /lzi/ |
| ds + afp: | /lzi/ |
| dp + ams: | /lzi/ |
| dp + afs: | /lzi/ |
| dp + amp: | /lzi/ |
| dp + afp: | /lzi/ |

As I said earlier, there are no restrictions on the number of instances of a given feature that can be assigned to a particular slot: all of them are mapped. In the combinations in (37), the two competing nodes are instances of [ARG], which is mapped onto slot number 4 in the template in (34). The plural marker will be mapped whenever at least one of the competing [ARG] features dominates [plural]. This mapping is illustrated in (38). In (38a) the input clitics are third person plural accusative and third person singular dative. In (38b) the accusative third person clitic is

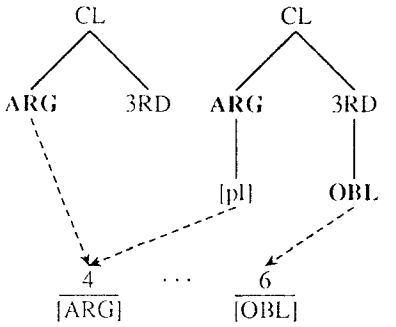
[33] In certain dialects of French, and very sporadically in Catalan, it is the case that in the same type of combination (3rd acc + 3rd dat) only the dative surfaces.
[34] The feminine marker never shows up in combinations of two third person clitics in Barceloní. See footnote 20.

singular, while the dative clitic is plural. I only give the relevant part of the template:

(38)a.




b.

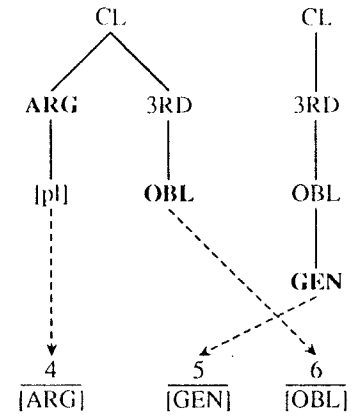



The dative clitic in isolation, singular or plural, will have the same surface form as the combinations of two third person clitics because the same type of information is mapped in both cases.

Notice that [ARG] and [OBL] do not occupy adjacent slots in the template. The intervening slot (number 5) is occupied by [GEN], later spelled out as /n/. The mapping of the combination of a third person dative clitic and a [GEN] clitic is shown in (39):



(39)




The example in (5c), repeated below as (40c), illustrates the final output resulting from the mapping in (39). I also repeat (5a, b), as (40a, b), respectively, to show the clitics involved as they appear in isolation:[35]

(40)a. De pomes **en** donaré als nens demà.
*of apples part, will-give*(1*st*) *to-the children tomorrow*

b. Als nens, [**ɔlzi**] donaré pomes demà.
*to-the children 3rd-pl-dat will-give*(1*st*) *apples tomorrow*

c. De pomes als nens, [**ɔlzɔni**] donaré demà.
*of apples to-the children* [*ɔlzɔni*] *will-give*(1*st*) *tomorrow*
'I will give apples to the children tomorrow'

(40c) shows that the dative clitic /lzi/ has been "split" by the partitive clitic /n/. This splitting is accounted for by the mapping to the template, illustrated in (39), together with the claim that the mapping of [3RD] clitics in Barceloní is based on morphological features, not on morphological structures.

A comment has to be made about the ablative clitic *en*. As can be seen in (34), there is no specific slot for [*a*]. Instead, the node that dominates this feature, [GEN], is mapped. The feature [*a*], which is not mapped onto the template and which does not receive a specific spell-out, is

[35] In these combinations, when the dative clitic is singular, the surface form is /ni/, instead of the predicted /lni/. This case can be accounted for through a morphonological rule that deletes /l/ only when adjacent to /n/. This rule cannot apply when the dative clitic is plural because of the /z/ intervening between /l/ and /n/ (see (40c)).

nevertheless needed because the clitic that has this particular Case, and not any other instances of *en*, is the target of morphological rule (17a).

The example in (41) illustrates the use of the six slots in the template in (34) (again, ignore the schwas):

(41) [sə **tə** **mə** **lzə** **n** **i**]
[CL, ARG] [CL, 2] [CL, 1] [ARG, pl] [GEN] [OBL]
vas quedar tres
*kept*(2*nd*) *three*
'you took three of them from mine (e.g., children)'

*quedar-se* 'to keep for oneself' is a pronominal verb which in the present case appears in the second person singular (this marking also appears on the auxiliary *vas*). The second person singular reflexive syntactic clitic is "represented" by two phonological clitics: /s/ (bare [ARG]) for the reflexive and /t/ ([2]) for second person.[36] /m/ ([1]) represents an ethical or affected dative. In the gloss to the example I have written 'mine' in order to reflect the involvement of the speaker. /lzi/ ({[ARG], [OBL]}, which appears as a discontinuous clitic because of the presence of /n/ [GEN])) is a dative of inalienable possession (reflected more or less in the gloss by 'children', in parentheses). Finally, the penultimate clitic in the clitic cluster in (41), /n/, is coindexed with an empty category N sister of the quantifier *tres* 'three'. The situation in which the sentence in (41) would be uttered is like the following: I have children, and there are a few apples that belong to them; now you have taken three of them from my children. Notice that in this sentence there is a mismatch between the number of syntactic clitics and the number of phonological clitics: there are four syntactic clitics and six phonological clitics (occupying six different slots).

## 5. Spell-out

The last operation that takes place within the Morphology Component, and which provides the input to phonology, is spell-out rules, which provide morphological features with phonological content.

Below, as an example, I give the relevant spell-out rules needed for Barceloní proclitics, which make direct reference to the material mapped onto the template. As mentioned above, the specification *a*, present in

[36] This "slitting" of reflexive clitics into a reflexive (bare [ARG]) and a first or second person clitic ([1] or [2]) occurs only under very specific conditions. See Bonet (1991) for a discussion of these cases.



the structure of the ablative clitic, has no spell-out. In this case, the node that immediately dominates *a*, [GEN], which was also mapped onto the template, is spelled out instead.

(42)a. [CL, 1] → /n/ / [plural]
b. [CL, 1] → /m/
c. [CL, 2] → /w/ / [plural]
d. [CL, 2] → /t/
e. [CL, ARG] → /s/
f. [ARG] → /l/
g. [NEUT] → /u/
h. [GEN] → /n/
i. [OBL] → /i/

There are two spell-out rules for first and second person. This is necessary because of the allomorphy found in these clitics. The examples in (43a–d) illustrate the spell-out rules in (42a–d). The relevant segments appear in boldface:[37]

(43)a. En Miquel **ens** vol veure.
*the Miquel* 1*st-pl-acc wants see*
'Miquel wants to see us'

b. En Miquel **em** vol veure.
*the Miquel* 1*st-acc wants see*
'Miquel wants to see me'

c. En Miquel **us** vol veure.
*the Miquel* 2*nd-pl-acc wants see*
'Miquel wants to see you(pl.)'

d. En Miquel **et** vol veure.
*the Miquel* 2*nd-acc wants see*
'Miquel wants to see you (sg.)'

The application of (42a) vs. (42b), (42c) vs. (42d), and (42e) vs. (42f) is determined by the Elsewhere Condition. (42a) and (42c) will take always precedence over (42b) and (42d), respectively. Therefore, sequences like

[37] The phonological form /w/ introduced by the spell-out rule in (42c) is realized as a vowel in certain contexts, like (43c). Some of the arguments for proposing an underlying glide can be found in Bonet (1991).

*/mz/ or */tz/, in which (42b, d) have preceded (42a, c), respectively, will never be derived. Spell-out rule (42e) applies only if [ARG] is dominated by CL. This will be the case with reflexives and impersonals. If [ARG] is not dominated by CL, which is the case of [3RD] clitics after the mapping to the template, rule (42f) will apply instead.

The spell-out rules for the agreement features [plural] and [feminine] are given in (44), below. They are not specific to pronominal clitics, but apply elsewhere in the language. I follow Harris (1991a, b) in claiming that the final vowel found in nouns and adjectives (and pronominal clitics) is not the direct spell-out of 'feminine' or 'masculine' but is the exponent of a declensional class. There is only one gender feature: [feminine] (in his terms *f*). Adapting his account for Spanish to Catalan, we can say that the schwa that appears in feminine clitics in Catalan is the spell-out that corresponds to the default declensional class for the feature [feminine], ]*a*:

(44)a. [plural] → /z/
b. [feminine] → ]*a*
c. ]*a* → /ɔ/

Presumably clitics are affected by these rules at the same point they apply to other items in the language. The relative order between the spell-out of [plural] and the spell-out of declensional classes is determined in Harris' approach, by a certain class of linearization rules. I do not have anything to say in this regard.

## 6. Conclusions

In this article several opaque output forms in Romance pronominal clitic combinations have been analyzed. It was observed that opaque forms always coincide with clitics that exist independently in the language, a generalization not captured (or even noted) in previous discussions of the phenomenon (cf. Perlmutter (1971) and Wanner (1977)). In this article it has been proposed that pronominal clitics constitute hierarchical structures of morphological features and that most opaque forms are the result of the application of morphological rules – delinking and insertion operations that modify the internal structure of clitics. With this proposal, the generalization mentioned above is easily captured because clitics are drawn from a closed set of morphological features, and delinking operations can only simplify the existing structure, while insertion operations can only add features from a closed set. Random outputs cannot be obtained. Morphological operations, together with the linearization of clitics, through the mapping to a template, take place within the Morphology Component, between S-structure and PF. A crucial assumption for the proposal made here is the idea that pronominal clitics lack phonological content throughout the syntax and until very late in the Morphology Component. The introduction of phonological information through spell-out rules provides the input to PF, which deals exclusively with phonological processes. The changes that often affect clitics, of the sort explored here, are then the product of the manipulation of morphological material, not phonological material. If the changes were solely due to phonological factors, it would be an accident that the resulting opaque form always coincides with an independently existing clitic. In the present proposal, there is no place for filters against sequences of phonologically identical clitics, for instance.



One of the ideas that might be pursued in the future, as mentioned in footnote 8, is that, similarly to the Morphology Component, hierarchical feature structures are also present in the syntax and that the mapping to the Morphology Component consists of the pruning of syntactic feature structures in different ways depending on the properties of each individual language. One of the features that surely would be present in this more complex syntactic feature structure would be [±animate], a feature that seems often to have an active syntactic and/or semantic role in many languages, but that does not have, so often, a morphological correlate.

Clitic order is determined in the Morphology Component through the mapping to a template. At that point more opaque forms might be created when two morphological features compete for one and the same slot since both of them are mapped onto that same slot. Recall the cases from Barceloní discussed in Section 4, where the appearence of an opaque form is caused by the combination of two third person clitics, accusative and dative. If, on the contrary, whole clitics compete for the same slot, the derivation will crash; no grammatical output is possible. Linearization through the mapping to a template constitutes a challenge to the idea that morpheme order is determined syntactically. This idea has become popular, especially after the work by Baker (1985, 1988) and Pollock (1989). In the present proposal, a syntactic approach to the ordering of clitics has been rejected, and an alternative has been proposed. This does not necessarily imply that the Mirror Principle "effect" (Baker (1985)) is nonexistent, given that this principle has been related to both linear ordering and hierarchical structure. The former but not the latter concept has been rejected in the present paper. The alternative proposed here allows not only for a description of clitic order but also for an account of the numerous opaque forms that often arise in clitic combinations. Opaque forms would be very difficult to account for within a syntactic approach.

It is not the case that opaque forms are only found with clitics. Similar cases exist in verbal agreement morphology. Basque, for example, has a phenomenon called Ergative Displacement (analyzed in Laka 1993), which very basically consists of the appearence, in specific circumstances, of an absolutive marker with the person features of the (instead) expected ergative marker. In other words, the ergative marker "acquires" the form of the absolutive marker. This phenomenon, then, is not very different from the spurious *se* of Spanish or the *si* to *ci* of Italian, to mention some of the cases discussed here. A natural step, in the future, would be to extend the proposal made here to verbal agreement morphology in general. After all, verbal agreement morphology, as well as pronominal clitics consists of a set of $\phi$-features, mainly Case, person, number, and gender.

## References


Anderson, Stephen R.: 1992, *A-Morphous Morphology*, Cambridge University Press (Cambridge Studies in Linguistics, 62), Cambridge.

Baker, Mark: 1985, 'The Mirror Principle and Morphosyntactic Explanation', *Linguistic Inquiry* **16**, 373–416.

Baker, Mark: 1988, *Incorporation: A Theory of Grammatical Function Changing*, University of Chicago Press, Chicago.

Bartra, Anna: 1987, 'Encara n'hi ha més (entorn de *en* i alguns SNs genitius', *Llengua & Literatura* **2**, 377–427.

Belletti, Adriana: 1988, 'The Case of Unaccusatives' *Linguistic Inquiry* **19**, 1–34.

Benveniste, Émile: 1966, *Problèmes de linguistique générale*, Éditions Gallimard (Bibliothèque des Sciences Humaines), Paris.

Bonet i Alsina, M. Eulàlia: 1991, *Morphology after Syntax: Pronominal Clitics in Romance*, Ph.D. dissertation. MIT, distributed by MIT Working Papers in Linguistics.

Borer, Hagit (ed.): 1986, *Syntax and Semantics 19: The Syntax of Pronominal Clitics*, Academic Press, New York.

Chomsky, Noam: 1981, *Lectures on Government and Binding*, Foris (Studies in Generative Grammar, 9), Dordrecht.

Cinque, Guglielmo: 1988, 'On *si* Constructions and the Theory of *Arb*', *Linguistic Inquiry* **19**, 521–582.

Clements, George Nick: 1985, 'The Geometry of Phonological Features', *Phonology Yearbook* **2**, 223–250.

Elliott, W. Neil: 1986, 'On the Derivation of *en*-Clitics, in H. Borer (ed.), *Syntax and Semantics 19: The Syntax of Pronominal Clitics*, Academic Press, pp. 97–122.

Halle, Morris: 1990, 'An Approach to Morphology', *Proceedings of the Northeastern Linguistics Society* **20**, 150–184.

Halle, Morris: 1991, 'The Latvian Declension', in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1991*, Kluwer, Dordrecht, pp. 33–47.

Halle, Morris and Alec Marantz: 1993, 'Distributed Morphology and the Pieces of Inflection', in K. Hale and S. J. Keyser (eds.), *The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*, (Current Studies in Linguistics, 24) MIT Press, Cambridge, pp. 111–176.

Harris, James W.: 1991a, 'The Exponence of Gender in Spanish', *Linguistic Inquiry* **22**, 27–62.





Harris, James W.: 1991b, 'The Form Classes of Spanish Substantives', in G. Booij and J. Van Marle (eds.) *Yearbook of Morphology 1991*, Kluwer, Dordrecht, pp. 65–88.

Jakobson, Roman: 1956/1971, 'Shifters, Verbal Categories, and the Russian Verb', in *Selected Writings*, vol. II, Mouton, The Hague, pp. 130–147.

Kany, Charles E.: 1951, *American-Spanish Syntax*, University of Chicago Press, Chicago.

Kiparsky, Paul: 1982, 'Lexical Morphology and Phonology', in I.-S. Yang (ed.), *Linguistics in the Morning Calm*, Hanshin, Seoul, pp. 3–91.

Kiparsky, Paul: 1985, 'Some Consequences of Lexical Phonology', *Phonology Yearbook 2*, Cambridge University Press, Cambridge, pp. 85–138.

Laka, Itziar: 1993, 'The Structure of Inflection: A Case Study on $X^0$ Syntax', in J.-I. Hualde and J. Ortiz de Urbina (eds.), *Generative Studies in Basque Linguistics*, John Benjamins, Amsterdam, pp. 21–70.

Lumsden, John Stuart: 1987, *Syntactic Features: Parametric Variation in the History of English*, Ph.D. dissertation, MIT.

Marantz, Alec: 1984, *On the Nature of Grammatical Relations*, (Linguistic Inquiry Monographs, 10) MIT Press, Cambridge.

Marantz, Alec: 1988, 'Clitics, Merger, and the Mapping to Phonological Structure', in M. Hammond and M. Noonan (eds.), *Theoretical Morphology. Approaches in Modern Linguistics*, Academic Press, New York, pp. 253–270.

McCarthy, John: 1986, 'OCP Effects: Gemination and Antigemination', *Linguistic Inquiry* **17**, 207–263.

McCarthy, John: 1988, 'Feature Geometry and Dependency: A Review', *Phonetica* **43**, 71–99.

Otero, Carlos: 1976, 'The Dictionary in a Generative Grammar', paper presented at the 91st Annual Convention of the MLA, New York.

Perlmutter, David M.: 1971, *Deep and Surface Structure Constraints in Syntax*, Holt, Rinehart and Winston, New York.

Pollock, Jean-Yves: 1989, 'Verb Movement, Universal Grammar, and the Structure of IP', *Linguistic Inquiry* **20**, 365–424.

Pranka, Paula: 1983, *Syntax and Word Formation*, Ph.D. dissertation, MIT.

Rizzi. Luigi: 1986, 'Null Objects in Italian and the Theory of *pro*', *Linguistic Inquiry* **17**, 501–558.

Ross, John Robert: 1972, 'Doubl-ing', *Linguistic Inquiry* **3**, 61–86.

Saccon, Grazziella: 1988. 'On the Order of Clitic Clusters in Italian', unpublished manuscript, Harvard University.

Sagey, Elizabeth: 1986, *The Representation of Features and Relations in Nonlinear Phonology*, Ph.D. dissertation, MIT.

Simpson, J. and M. Withgott: 1986, 'Pronominal Clitic Clusters and Templates', in Hagit Borer (ed.), *Syntax and Semantics 19: The Syntax of Pronominal Clitics*, Academic Press, New York, pp. 149–174.

Wanner, Dieter: 1977, 'On the Order of Clitics in Italian', *Lingua* **43**, 101–128.


Received 27 February 1992
Revised 19 February 1994

Unitat de Filologia Catalana
Facultat de Ciències de la Comunicació, Edifici I
Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona), Spain
ilftb@cc.uab.es