

ROMANCE CLITICS AND FEATURE ASYMMETRY  
AN AUTOSEGMENTAL-BASED APPROACH

(A SLIGHTLY-REVISED VERSION OF MY DISSERTATION)

©

LOUIS-H. DESOUVREY

[louish8@gmail.com](mailto:louish8@gmail.com)

2005/2009

## Notice

This is a slightly-revised version of my doctoral dissertation. Mostly limited to the first chapter, the revision does not significantly modify the content. Instead of making a thorough revision, the points that turned out to be less satisfactory are (and will be) addressed in further research. Thus the analysis of *wh*-movement in French and English, as well as adverb placement (including negation), are modified somehow in subsequent research. Also, the main aspects of this dissertation are further discussed in a published paper, Desouvrey (2005).

Notice that clitics are referring elements, a property that is not thoroughly discussed in the dissertation. At that time, I did not have a satisfactory theory of coreference, and I implicitly assumed the formalism of Binding Theory, namely co-indexing, to indicate coreference relations. This has no significant consequences on any aspect of the proposed analysis of Romance clitics, although our current knowledge of coreference would have eased the discussion of certain facts (cf. Desouvrey 2002).

**UNIVERSITÉ DU QUÉBEC À MONTRÉAL**

**ROMANCE CLITICS AND FEATURE ASYMMETRY: AN  
AUTOSEGMENTAL- BASED APPROACH**

**A DISSERTATION**

**SUBMITTED**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE  
OF DOCTOR OF PHILOSOPHY IN LINGUISTICS**

**BY**

**LOUIS-HARRY DESOUVREY**

**FEBRUARY 2000**

## ACKNOWLEDGEMENTS

I would like to thank many people who helped me during the years of elaboration of this work. First of all, I thank professor Anna-Maria Di Sciullo, who accepted to be my advisor.

Having ideas and an advisor are not sufficient conditions to complete a dissertation. Three more readers were needed (only three, but at some point I felt as if it were like three hundred!). Mohammed Guerssel and Henrietta Cedergren allowed me some of their time and accepted to become members of my dissertation committee. I thank them for that, and for the comments they provided to me. A special thanks to Mohammed, who remained friendly with me while he was evaluating the thesis. Also, I am grateful to Angela Ralli, who readily accepted the task of being the external evaluator, and who brought to my attention interesting questions and facts.

I would like to thank other people whom I consulted from time to time about the data. My friends Sophie Rosset (France), Inma Jauregui, Pablo Mendez (Basque Country), and Chris Miller, who thanks to his competence in phonology and his great culture in linguistics, provided me useful comments, as he proofread the three first chapters of the thesis. A special thanks also to Sauveur Joseph, Grâce Masagbor, Stanca Somesfallea, and the other members of the Asymmetry Project for their friendship.

I am also grateful to my teachers, including professors and *chargés de cours*, at the Université du Québec à Montréal.

I am very grateful to Jeff Gruber, a great linguist, with whom I had many interesting discussions on languages, linguistics, and many more things. He introduced me to wonderful people, among others, Raymond Flournoy. At

Raymond's, I met a lot of kind people, including Shirin Finley. Thanks to her, I have a new friend, Joël Finley, to whom I am grateful for proofreading the final version of the dissertation, and above all for his friendship.

I am pleased to thank my mother Olivia, to whom I dedicated this thesis, for the sacrifices she assumed to provide me education in my childhood. I would like to thank also my sisters and brother Elsie, Bérénice, and Roro, as well as my cousin Sonny Exume and my friend Joël Maxis for encouragements. I owe in fact a special thanks to Sonny and our friend Laurier Bonhomme, who helped me to land in this country.

I am also grateful to my *compagne*, Stéphanie Kovacs, for her patience, encouragements, and more.

Despite considerable (non-linguistics) difficulties, I have been able to complete this work. Determination, abnegation, and enlightenment were needed above all. I am grateful to my Benefactor, the Lord, for these. O Lord, may You see this work as a token of my worship.

This work was partly supported by the Asymmetry Project. In addition, I got a year of support from the Fondation de l'UQAM (PAFARC). I also got from the CCIFQ a two-month funding for a trip in Jussieu (Paris). I thank professor Jean Lowenstamm who sponsored me there, my friend Sophie Rosset, with whom I have some useful discussion, and the *Atelier des doctorants en linguistique* that allowed me to present a paper. Finally I am grateful to the program *Prêts et bourses aux étudiants* of Quebec government, which warranted me a loan during hard times.

## TABLE OF CONTENTS\*

LIST OF FIGURES AND TABLES .....	ix
RÉSUMÉ.....	x
ABSTRACT .....	xi
INTRODUCTION .....	1
1. The problem .....	1
2. Objectives .....	2
3. Theoretical framework.....	3
4. Methodology.....	3
5. How to falsify the proposal .....	4
6. Organization of the thesis.....	4
CHAPTER I: ON CLITICS	
1.1 Introduction .....	6
1.2 The nature of clitics .....	7
1.3 Theories of clitics.....	10
1.3.1 The principles and parameters approach to grammar.....	10
1.3.2 Non-syntactic analyses.....	12
1.3.3 Syntactic analyses .....	15
1.4 Some remarks on the Principles and Parameters Theory .....	21
1.5 Assumptions .....	29
1.5.1 The model of grammar .....	29
1.5.2 Nonlinear phonology .....	30
1.5.3 Constraint interaction.....	35
1.5.4 Summary .....	37
1.6 The proposal .....	37
1.6.1 Case features.....	37

---

\* This TOC is that of the original version and does not necessarily reflect the actual number of each page.

1.6.2 Syntactic features in tiers .....	40
1.6.3 Feature tree structure .....	43
1.6.4 On the argument structure of verbs .....	55
1.6.5 Properties of the Computational Space.....	60
1.6.5.1 Merge.....	60
1.6.5.2 Constraints on Merge.....	63
1.7 Summary .....	71
CHAPTER II: AUTOSEGMENTAL CONSTRAINTS ON SYNTACTIC STRUCTURES	
2.1 Introduction .....	74
2.2 Feature checking .....	75
2.3 OCP effects.....	86
2.3.1 Movement .....	87
2.3.2 Fusion.....	102
2.3.2.1 X-slot theory .....	107
2.3.3 Discussion of the OCP .....	112
2.4 Transformation ordering and Clitic clusters.....	116
2.4.1 Clitic order .....	117
2.4.2 Clitic cooccurrence restrictions .....	123
2.4.2.1 Clitic licensing .....	125
2.4.2.2 Animacy effects.....	130
2.5 Summary .....	136
APPENDIX: AUXILIARIES.....	138
CHAPTER III: ADDITIONAL CLITICS IN FRENCH	
3.1 Introduction .....	143
3.2 The syntax of se .....	145
3.2.1 Middle se.....	151
3.3 The syntax of on .....	158
3.4 The syntax of en.....	162
3.4.1 Feature structure .....	162
3.4.2 OCP effects .....	167
3.4.3 Animacy effects.....	168
3.4.4 Subnominal en.....	172

3.5 The syntax of y .....	179
3.5.1 Feature structure .....	179
3.5.2 Subnominal y.....	181
3.5.3 Y as a subcategorized argument .....	183
3.6 On double clitics .....	187
3.6.1 On the incompatibility of two subnominal clitics .....	188
3.7. Latent Case .....	191
3.7.1 Inalienable possession.....	193
3.8 Summary .....	196
<b>CHAPTER IV: CROSS-LINGUISTIC EVIDENCE AND EXTENSIONS</b>	
4.1 Introduction .....	198
4.2 English personal pronouns .....	199
4.2.1 Object pronouns.....	200
4.2.2 Subject pronouns .....	202
4.3 Haitian Creole .....	204
4.4 Spanish .....	206
4.4.1 Morpheme inventory and features.....	206
4.4.1.1 Feature structure of the paradigm .....	207
4.4.1.2 Feature structure of Spanish verbs .....	208
4.4.2 Spurious se .....	210
4.4.3 On the I-II/me-lui cluster .....	214
4.4.4 Spanish enclitics .....	222
4.4.4.1 Imperative verbs.....	222
4.4.4.2 Infinitive verbs .....	224
4.4.4.3 Clitic climbing.....	227
4.4.5 Clitic doubling .....	231
4.4.5.1 The problem .....	231
4.4.5.2 Common Spanish .....	235
4.4.5.2.1 Indirect object doubling .....	235
4.4.5.2.2 Strong pronoun doubling.....	237
4.4.5.2.3 Possessor doubling.....	239
4.4.5.3 River Plate Spanish .....	241
4.4.6 Summary .....	245



4.5 Overview of Italian Clitics .....	246
4.5.1 Feature specification .....	246
4.5.2 Opaque clitics .....	248
4.5.3 Summary .....	251
4.6 Wh-operators as special clitics .....	252
4.6.1 Interrogative que.....	252
4.6.2 Interrogative qui .....	258
4.6.3 A null wh-operator .....	263
4.6.4 Superiority effects .....	266
4.6.5 Summary .....	271
4.7 French negation .....	271
4.8 Summary .....	279
CONCLUSION .....	282
REFERENCES .....	288

## LIST OF FIGURES AND TABLES

Figure 1.1 The Government and Binding Model. ....	11
Figure 1.2 Di Sciullo's 1996 model of grammar. ....	29
Table 1.1 Relations of grammatical roles to Case and $\Theta$ -features. ....	39
Table 1.2 Declension of personal pronouns in French. ....	44
Table 4.1 English personal pronouns. ....	200
Table 4.2 Personal pronouns in Haitian Creole. ....	204
Table 4.3 Spanish personal pronouns. ....	206
Table 4.4 Italian personal pronouns. ....	246

## ABSTRACT

This thesis deals with pronominal clitics in Romance. A feature-based analysis drawing heavily on nonlinear representations in phonology as well as the notions of constraint interaction is proposed. It is shown that pronominal elements are inherently specified for a series of features, including Case and thematic features, which form tiers independent of the morphemes bearing them. Case features are crucial in the analysis; it is shown that an element is a special clitic in the sense of Zwicky (1977) iff it is specified for Case.

Similarly to pronominal elements, it is suggested that verbs are inherently specified for Case and thematic features, which are organized hierarchically. Assuming the bare phrase structure theory, coupled with a strong lexicalist hypothesis (e.g., Di Sciullo, 1986), which excludes abstract functional heads in the phrase-marker, it is shown that verbs interact with Case-specified arguments in ways that fall under a well-known family of constraints in nonlinear phonology, known under the name of Obligatory Contour Principle (OCP; Goldsmith, 1976; McCarthy, 1986p; etc.). The OCP in syntax forces Case-specified arguments that are generated within the complement domain of the verb, i.e. the domain containing the internal arguments, to move outside this domain. In addition to movement, the OCP requires Cases of the verb and those of the arguments to fuse. Case fusion forces Case bearing elements to become adjacent, usually by movement of the element intervening between them. This process is much like gemination effects in nonlinear phonology.

The OCP and a few other constraints make it possible to account for various phenomena related to cliticization in French, Italian, and Spanish in a principled way. These are: clitic order, including the cause of the dative-accusative and accusative-dative variation observed only in French, clitic climbing, clitic doubling and the *me-lui*/I-II effects.

It is shown that opaque clitics in French imperative verbs, as well as Spanish and Italian double clitics, result from the violation of different constraints, while they are repaired by the same mechanism, namely a delinking process. In French imperative structures, a feature mismatch between the clitic and the verb triggers a delinking process that converts the clitic into a spurious strong form. In Spanish, the third person dative clitic prevents the accusative clitic from being licensed, which triggers a delinking process yielding the spurious *se*. Similarly, in all clitic clusters in Italian, the dative clitic must be manipulated in order to allow the licensing of the accusative clitic.

Other important results are that clitic doubling in standard and River Plate Spanish is shown to arise from a feature mismatch during the structure building stage in the computational space. Encliticization to infinitive verbs in Spanish and clitic climbing are shown to be due to a superiority effect, given the hypothesis that both clitics and infinitive verbs in this language are operators. The same constraints are shown to carry over to elements traditionally considered to be unrelated: *wh*-movement in French is shown to be triggered by an OCP effect; the only difference with pronominal clitics is that, for some reason, Case-specified *wh*-operators move to clause initial position in French and similar languages.

**Keywords:** Autosegmental phonology, OCP, constraint interaction, asymmetry, feature checking, movement, Romance clitics, clitic cluster, clitic order, opaque clitics, clitic placement, clitic doubling, clitic climbing, *wh*-movement, superiority effects, *me-lui* effects, negation.

## INTRODUCTION

### 1. The problem

The analysis of pronominal clitics in Romance has been the subject of many debates in generative grammar since the seminal work of Kayne (1975) and Zwicky (1977). Most researchers consider clitics as a category on their own, that is, a lexical entry which contains a set of grammatical properties, parallel to other categories such as nouns, adverbs, and adjectives. Because not all languages have pronominal clitics, theoretically oriented work, in particular that which emerges from the Government and Binding tradition, proposes that the presence or absence of clitics in particular languages is a matter of parametric variation.

In the literature on clitics, two tendencies show up. In one, clitics are analysed as syntactic elements and are subject for example to movement. In the other, roughly speaking, clitics are affixes generated in their surface position in the syntax or in the morphology. The latter type of approach was a reaction against the former, and was motivated mainly by the lack of complementary distribution between clitics and NPs in some languages (e.g., River Plate Spanish, and Rumanian). This is indeed a challenging problem for the GB framework: since each element is specified for a position in the phrase marker, and there is only one position per element, a clitic is not expected to cooccur with the argument it is intended to replace, given the theta criterion. However, in other languages such as French and Italian (the standard dialects at least), complementary distribution between clitics and NPs is strictly observed. Thus both approaches have some initial plausibility, when one considers a limited set of data. However, it is desirable to have a theory of clitics that can account for complementary distribution and cooccurrence with an NP. Besides clitic doubling, this theory must be able to explain in a principled way a number of other important questions, among them the following:

- a. *Clitic order*: what governs the ordering of clitics within a cluster.
- b. *Clitic cooccurrence*: why, universally, some clitic combinations are impossible.
- c. *Clitic variation*: why, in some languages, clitics take different forms depending on whether they are in clusters or in isolation.
- d. *Clitic doubling*: why, in some languages, clitic doubling of an NP is required.
- e. *Clitic placement*: why clitics are mostly proclitics in Romance and why they can also be enclitics.

An analysis of clitic order and clitic cooccurrence restrictions was proposed in Perlmutter (1971). Assuming that the ordering of clitics is not a syntactic matter, he proposed a template where both order and cooccurrence restrictions are fixed. Although the template explains nothing in itself, and is at best a descriptive generalization, some researchers come to assume the templatic approach to clitic order, among others Bonet (1991, etc.). In fact Bonet (1991) describes the restrictions on clitic cooccurrence more systematically, and at the same time provides a type of representation which deals better with morphological variants, which she refers to as opaque forms. While Bonet proposes an approach that is not complete, in that she does not explain why the process she describes must take place in a given language, she gives compelling evidence that the traditional tools used to account for clitics are insufficient.

## 2. Objectives

This thesis deals with pronominal clitics in Romance, and French in particular. Regarding the questions listed above, the work will be exhaustive. All the peculiarities of cliticization of personal pronouns in French will be addressed, including procliticization, encliticization, variation in linear ordering, opaque forms, etc. A less detailed account of the main features of cliticization in Spanish, and in Italian to some extent, will also be offered: opaque forms, spurious *se*, clitic doubling, etc. It will also be shown why in languages such as English and Haitian Creole, personal pronouns cannot be (special) clitics. Furthermore, the syntax of wh-operators will be shown to follow from the proposal.

### 3. Theoretical framework

I agree with Bonet that current assumptions in syntactic theory may not permit an adequate account of cliticization phenomena. I will assume a nonlinear type of representation, modelled on those currently used in nonlinear phonology and morphology. This will enable us to derive most of the defining properties of clitics discussed in the literature from the Obligatory Contour Principle (OCP) (cf. Goldsmith, 1976).

### 4. Methodology

I consider clitics to be defined by a set of features, including phi-features and Case, which are among the syntactic features assumed in GB theory.<sup>1</sup> Indeed, I claim that clitics may not be taken to be an ontological category, nor a lexical entry (a similar claim is found in Everett (1996) and Zwicky (1994)). As mentioned above, the defining properties of clitics will be shown to follow from the OCP. Assuming that the tier-based representation I use is universal, clitics are expected to (crosslinguistically) vary only according to their feature specification.

Therefore, in this thesis features play a crucial role. In order to account for any phenomenon, all the elements involved must be known by their defining distinctive features. While features have come to have a greater importance in current grammatical analyses, they often are not accurately defined. In contrast, in this thesis I will attempt to justify all features which are attributed to a particular element. In doing so, the whole paradigm which it belongs to must be considered. In my view it is insufficient to assume that a given element is specified for some feature, for instance accusative. We must know whether there is another element which is specified or not for a Case feature, and how many features the paradigm contains overall. For instance, with respect to Case, an element which is specified for a single Case will behave differently in the syntax from an element which is specified for two Cases. As we will see, the problem at hand is somewhat complicated by the fact that

---

<sup>1</sup> In the technical sense of a grammatical feature, Case will be spelled throughout the thesis with a capital C.

there is no one-to-one correspondence in any paradigm between features and morphemes. In the proposed system, an element may be either unspecified, or specified for either one or two Case features, for instance. The element's syntactic behaviour will vary accordingly.

## **5. How to falsify the proposal**

As mentioned above, I will use a tier-based representation, drawing heavily on nonlinear phonology. This type of representation may not be familiar to all syntacticians, and therefore some may be tempted to reject it. However, beyond the nature of the representation, our approach is appealing to the extent that it provides a principled account of various phenomena which are not addressed in the literature. Because the constraints which are formulated in this thesis refer only to the features of lexical elements, and not for instance to any node in the syntactic tree, as is customary in current analyses, an alternative approach must perforce rely upon the notion of distinctive features. In addition, to succeed, it must account for the same facts with fewer claims and principles.

## **6. Organization of the thesis**

The discussion proceeds as follows. In chapter 1, I present an overview of the elements traditionally classified as clitics in Romance languages, highlighting their properties. This is followed by an overview of various analyses proposed in the GB framework. After showing that the GB framework (and its successor, Chomsky's Minimalist Program) is not able to handle all facts related to cliticization, I present an overview of nonlinear phonology, the specific assumptions that underlie the thesis, as well as my proposals, including an appropriate model of grammar.

In chapter 2, I discuss the syntax of clitics in French, showing that they can best be explained with the assumptions presented in chapter 1—and indeed, they must be. In particular, I will show that the OCP successfully explains all the cases of cliticization of personal pronouns, once their Case features are taken into account. In chapter 3, I pursue the analysis of the French pronoun system with pronouns which are traditionally considered non-argumental clitics. In chapter 4, I present some

cross-linguistic evidence for the analysis, focussing on the pronoun systems of various languages, including English, Haitian Creole, Spanish and Italian. While the analysis predicts that English and Haitian Creole cannot have special clitics in the sense of Zwicky (1977), it allows a principled account of opaque clitics in Italian and Spanish, as well as the phenomenon of clitic doubling which occurs in the latter. The analysis will be shown to account for *wh*-movement, as well as the syntax of negation in French.



# **CHAPTER I**

## **ON CLITICS**

### **1.1 Introduction**

In this chapter I will present an overview of the set of phenomena traditionally referred to as cliticization. I will highlight the fact that clitics, as usually understood, cannot be taken to be a class of grammatical categories in natural languages. The defining properties of elements classified as clitics, such as personal pronouns in Romance, can be found in varying degrees in many other grammatical categories, including nouns, adverbs, etc.

I present a survey of numerous analyses of clitics, focussing mostly on those embedded in the Principles and Parameters approach to syntax, also known as the Government and Binding (GB) framework (cf. among others, Chomsky, 1981, 1986b, etc.). As I will show, a common problem to these analyses is that they do not provide an answer to the following questions: (a) why do some elements in a given language come to show clitic-like behaviour? and (b) how can a grammar predict which elements show such behaviour? With these questions (and others) in mind (see the general introduction), I consider some basic claims of GB, arguing that, though restrictive, they are not suitable for the explanation of cliticization.

The theoretical problems that confront traditional approaches lead us to seek an alternative approach. Fortunately, nonlinear phonology appears to show the hallmarks of a good theory in that it is powerful, simple and restrictive. Drawing heavily on this theory of phonology, I claim that every syntactic element, that is every free morpheme, is defined by a set of features, including Case and phi-features (i.e. person, number, and gender). Each of these appears on its own tier, similarly to

phonological features. Such a representation makes it possible to account for the behaviour of pronouns with the same types of constraints and processes as those currently used in nonlinear phonology.

The chapter proceeds as follows. The next section contains a general overview of clitics, as discussed in the literature. Section 3 discusses some theories of clitics in the GB framework. In section 4, I attempt to show that many structural operations which are currently used in order to restrain the power of the grammar are not well-grounded on a purely conceptual basis. In section 5, I present the assumptions I build on, most notably the theory of nonlinear phonology (Goldsmith, 1976, etc.). I briefly introduce the main feature of this theory, showing that it contains a minimal set of stipulations, which make possible an account of both the similarity and the diversity of syntactic processes found in natural languages. I also present the model of grammar advocated in Di Sciullo (1996), which allows one to articulate the proposal. In section 6, I briefly present my proposal to the effect that syntactic features can (and must) be treated in a fashion similar to phonological features in nonlinear phonology, as well as a number of assumptions and specific claims respecting the lexicon and the argument structures of verbs. Section 7 summarizes the chapter.

## 1.2 The nature of clitics

Clitics get their name etymologically from the fact that they "lean", that is, they are dependent in some way, on an adjacent word. Traditionally, in generative grammar it is proposed that this dependency is two-fold: phonological and syntactic. This corresponds roughly to the classic distinction (by Zwicky, 1977) between simple clitics and special clitics.

Indeed, simple clitics are dependent on an adjacent host, in the sense that they must be syllabified with it; but syntactically, they appear in an expected position. Examples of simple clitics are English 's (*is/has*), 've (*have*) (see Zwicky, 1977; Zwicky and Pullum, 1983). On the other hand, special clitics are not necessarily prosodically dependent; their peculiarity comes from the fact that they appear in a rather special position for the function they fill. For instance, in French, direct objects follow the verb (cf. *Marie aime Paul* 'Mary likes Paul'), as in English, but a pronoun

corresponding to a direct object appears before the verb (cf. *Marie l'aime* ' Mary likes him' ). It is this word-order alternation that makes a clitic special. Special clitics can be found in Serbo-Croatian (see Halpern, 1995; Anderson, 1996; etc.), and in Romance pronoun systems (see references below). In this thesis, I will deal only with special clitics, especially those of French, Italian, and Spanish. But the theory will make it clear why similar categories in other languages (e.g. English) are not special clitics.

These two fundamental properties of clitics, namely prosodic deficiency and special syntax, make them rather difficult to analyse. Despite the considerable literature devoted to their study (see Nevis et al., 1994; Janse, 1996), their nature remains controversial. As Zwicky (1977: 1) points out, "[clitics are] morphemes that present analytic difficulties because they are neither clearly independent words nor clearly affixes." Indeed, clitics seem to belong to both of these categories, as suggested by this list of their most common properties in the literature (cf. Kayne, 1975; Zwicky, 1977; Anderson, 1992; etc.).

- (1)
  - a. Clitics must be adjacent to their host.
  - b. Only another clitic can intervene between a clitic and its host.
  - c. Clitics cannot be conjoined or modified.
  - d. Clitics cannot be stressed.
  - e. Clitics are likely to have morphologically distinguished Case.
  - f. Clitics may not occur at all in the absence of any verb.
  - g. Clitics may appear in second position of the sentence (or the clause).
  - h. Clitics do not allow further morphology on their hosts.

Elements with these descriptive features are common in many languages. Such features are not restricted to words with reduced forms or to some closed paradigm of functional elements, such as pronouns. Among morphemes displaying any of the above behaviour within a given language, some are undoubtedly free words. Indeed, many grammatical categories can have one or several properties of the above list (see Zwicky, 1994, for a list of some such categories). For instance, some syntactic constituents (NPs) present the same distribution as affixes with respect to a host, in

that no element can intervene between them; this is clearly the case of direct objects in English, as shown in (2) (examples from Zwicky, 1994).

- (2) a. \*We passed quickly the other racers.  
 b. We passed the other racers quickly.

In French, subject NPs do not allow an adverb to separate them from their verb, although they allow the negative clitic *ne* to stand in this position, as well as a relative clause, as shown in (3). Clearly, it appears that many words, whose nonclitic status is uncontroversial, possess some of the properties of clitics.

- (3) a. \*Jean souvent voit Marie.  
 b. Jean voit souvent Marie.  
     'Jean often sees Mary.'  
 c. Jean ne voit pas Marie.  
     'Jean does not see Marie.'  
 d. L'homme que tu as vu aime Marie.  
     'The man that you saw likes Mary.'

We may note that the most celebrated characteristic of clitics, namely the fact that in many languages they occur in the second position of the clause, either after the first phonological word, or after the first syntactic constituent, a phenomenon known traditionally as Wackernagel's Law, is in fact not a privilege of clitics or affix-like elements. In some well-known languages (e.g. German), the verb must occupy such a position, a phenomenon referred to as the verb-second effect.

These facts tend to show that the term clitic is not a well-defined one in that the properties it refers to can be found in a heterogeneous set of elements which are clearly not affixes. It is for this reason that Zwicky (1994) refers to clitics as an umbrella term, not a genuine category in natural languages. As he puts it, "umbrella terms are names for problems, for phenomena that present 'mixed' properties of some kind, not names of theoretical constructs." If then the clitic is not a genuine grammatical category, the goal of the grammatical theory should be to look for, and ultimately explain, the particular properties that make a word a clitic, whatever its

size (i.e. full or contracted morphemes, particles, etc.), or its grammatical nature (verb, pronoun, etc.). In other words, the set of descriptive properties in (1), as well as Wackernagel's Law, must not be taken to be axiomatic features, that is a part of the lexical entry of any element, from which the analysis should proceed. Thus, in this thesis the term clitic will be used descriptively, without implying any theoretical status. In the next section, I consider the main analyses of clitics that have been proposed in the Principles and Parameters theory (PPT), restricting the discussion to pronominal clitics in Romance languages, mainly French.

### **1.3 Theories of clitics**

Romance clitics are variously analysed in the literature. These analyses belong to two main types: lexical analyses, and syntactic analyses. Such analyses are constrained by the nature of the framework they are embedded in, the Principles and Parameters Theory (PPT). Thus, even though the various analyses are significantly different from each other, they share a common unsatisfactory feature: the descriptive properties listed in (1) are simply assumed, as noted above. As a result many striking facts related to the distribution of pronominal clitics in Romance are considered to be idiosyncratic, and are not accounted for. Below, I begin by presenting the main features of GB theory, and will discuss the two approaches to clitics found in this theory: syntactic analyses and non-syntactic analyses.<sup>1</sup>

#### **1.3.1 The principles and parameters approach to grammar**

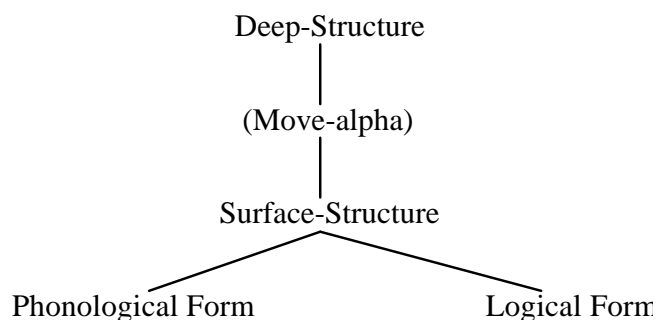
Consider the model of grammar known as Government and Binding theory (cf. Chomsky, 1981, 1986b, etc.; and many others). This model includes the following components: the lexicon, the syntactic component, the phonological component, and the logical form component. The syntax is the central part and contains the base rules (i.e. X-bar theory), D-structure and S-structure. D-Structure, which is a product of

---

<sup>1</sup> GB theory is abandoned by Chomsky (1995), who makes a number of new proposals about the nature of syntax, which are referred to as Minimalist Program (MP). MP contains some interesting insights, as we will see below, but it is not completely different from GB, as some unsatisfactory features of the latter are maintained. I will point out differences between GB and MP when relevant to the discussion.

the base rules, is mapped onto S-Structure by the general transformational rule *Move-alpha*. Finally S-Structure is mapped onto PF and LF. This is illustrated in Figure 1.1.

**Figure 1.1** The Government and Binding Model.



In this theory, the lexicon plays a secondary role; it is indeed considered as a simple list of items (words and affixes) and idiosyncratic properties (or exceptions) of particular grammars that cannot be derived by general principles in the syntax. It is also assumed that the lexicon contains lexical rules that serve to form derived words, i.e. it is the locus of what is traditionally called derivational morphology. Most PPT practitioners assume that inflectional morphology takes place in the syntactic component, where inflectional morphemes, such as tense, agreement, etc., head their own phrasal projection, along the lines of Pollock (1989) and Chomsky (1991).<sup>2</sup>

The grammar also contains a number of universal principles. As a given principle is not active in every single natural language, the theory allows for variation according to parameter setting. A parameter specifies open options within a principle. For instance, the position of a head with respect to its complement is accounted for by a parameter associated with X-bar theory, according to which a head takes a complement and the complex head-complement takes a specifier. Each particular

---

<sup>2</sup> However, different views on this matter exist. For instance, Di Sciullo and Williams (1987) take the position that inflectional morphology and derivational morphology occur in an independent morphology component. Iatridou (1990) and C.L. Baker (1991) present some serious objections to the Pollock-type approach (see also Bouchard, 1995). More recently, Ralli (1997) argues for an approach where inflection is viewed as a morphological process taking place within a morphology component.

language realizes this according to the head parameter, which has two values: either the head precedes the complement, or the complement precedes the head.<sup>3</sup>

The elements combined and manipulated in S-Structure are treated in the phonological component, which deals mainly with prosodic structures (e.g. syllable rules) and phonological rules. But it could as well be the locus of some syntactic phenomena that apply at the end of the derivation (e.g. stylistic rules). As for the LF component, it deals with semantic and logical information conveyed by the structures. Covert syntactic operations, including movement and deletion, also take place in the LF component.

Given PPT, special clitics in Romance languages can be accounted for in two fashions: (a) syntactically, and (b) non-syntactically. What I mean here is that in the first case clitics are generated in some syntactic position at deep structure and that they undergo movement to another position. In the second case, again according to my own survey of the literature, clitics are generated either in the syntax or in a morphological component, but what is common to the second type of approach is that clitics do not undergo any overt movement.

### **1.3.2 Non-syntactic analyses**

This type of analysis can be roughly characterized as follows: clitics, which are generally viewed as the spell-out of a set of (phi)-features, are base-generated in the position in which they appear, either under an  $X^0$  or a phrasal node (Borer, 1984; Bouchard, 1984, 1995; Everett, 1996; Jaeggli, 1982, 1986; Klavans, 1985; Labelle, 1985; Roberge and Vinet, 1989; Roberge, 1990; Rochette, 1988; Safir, 1985; etc.), or head their own phrasal projection (Sportiche, 1996). As just noted, even though clitics are dealt with in the syntactic component (D- and S-structures), these analyses may not be considered syntactic; indeed, some scholars who subscribe to this type of analysis explicitly consider clitics to be phrasal affixes (e.g. Everett, 1996, and Klavans, 1985). The fact is that in this model of grammar some morphology occurs

---

<sup>3</sup> Koopman (1984) and Travis (1984) propose to handle the head parameter with the directionality of thematic role and Case assignment.

in the syntactic component, and behaves in a similar fashion to syntactic categories. For example, the ending *-ed* in *liked* is at the D-structure level an autonomous syntactic element (see, amongst others, Chomsky, 1991; Pollock, 1989). Further, given X-bar theory, any element adjoined to an  $X^0$  element is considered in this model to be sublexical (see the introduction in Borer, 1986). From this perspective, clitics are clearly syntactic affixes, exactly like tense and agreement morphemes.

Other non-syntactic analyses bear the label non-syntactic more transparently, as the clitic is dealt with outside the syntactic component, more precisely in a separate morphological component. They are of two subtypes, depending upon the place of the morphological component in the model. In one type, the morphological component, which contains all inflectional affixes, including clitics, precedes the syntactic component. Clitics are inserted in the syntax at S-structure as inflectional affixes, a view advocated in Anderson (1992), Auger (1995), Bessler et al. (1992), Cummins and Roberge (1994), etc. In the last type, advocated in Bonet (1991, 1995a, 1995b), the syntax manipulates abstract features (cf. Lumsden, 1987) that will receive lexical contents in the morphological component placed after the S-structure in the PF portion of the grammar (see Halle and Marantz, 1993). One may note that some of these scholars (Bonet, as well as Cummins and Roberge) use a templatic approach to clitics (see also Watson, 1997), an idea originally proposed by Perlmutter (1971).

These two types of non-syntactic analyses face different problems and appear to be successful with respect to certain facts. The first type avoids the problem of clitic doubling, the main motivation for this analysis, by base-generating the clitic in a special position, this leaving the argument position as a place where a full NP can eventually appear. An undesirable consequence is that the clitic is on a par with inflectional morphemes. As Dufresne (1993) points out, (inflectional) bound morphemes are either prefixes or suffixes, but can by no means be both. Now, object clitics in French and other Romance languages can be enclitic or proclitic, that is, in a non-syntactic analysis, (phrasal) prefix or (phrasal) suffix; indeed, with all verbs, except imperatives, clitics appear in preverbal position, while with imperative verbs, they follow the verb. Negated imperatives, however, pattern like other verbs. Thus:



- (4) a. Marie **le** voit.  
       ' M. sees him.'  
     b. Regarde-**le**!  
       ' look at him!'  
     c. Ne **le** regarde pas!  
       ' Do not look at him!'

An account of this fact in these analyses would encounter many difficulties, as other bound morphemes of this language show no such behaviour. Thus, in the case of (4a), if one were to assume that object clitics are base-generated to the left of the verb, additional claims would be required to account for (4b). For the latter, one could say that imperative verbs must move leftward (without the clitic) to some functional head, as suggested for Greek and Balkan languages by Terzi (1996b), and Rivero (1994). But such a claim would not hold for negated imperatives, (4c). It would then be necessary to stipulate that negated imperatives pattern as other simple verbs. Clearly, an analysis which includes any of these ad hoc claims would be considerably weakened.

One may note that the problem is even more challenging for these analyses, since other Romance languages pattern differently from French. In Spanish and Italian for instance, gerundive and infinitive verbs allow enclitics only, which may become proclitic to a matrix verb (by clitic climbing). The latter possibility does not exist in French, except in causative constructions. Moreover, in French (pro)clitic order is subject, among other things, to the person of the clitics involved. For instance, one has the sequence *le-lui* (3.ACC.-3.DAT), but not *le-me* (3.ACC.-1.DAT); instead, one has *me-le* (1.DAT.-3.ACC.), as illustrated in (5).

- (5) a. Marie **le lui** donne.  
       ' Mary gives it to him.'  
     b. Marie **me le** donne.  
       \*Marie **le me** donne.  
       ' Mary gives it to me'

To the best of my knowledge, no attempt has been made to account for such a fact. Perhaps given the tools available within the GB theory, it may be that it is not possible to obtain a principled account. Indeed, as we will see in the next chapter, the distribution of such pronouns results from the interaction of their own Case feature and that of the verb, and some constraints on movement. Now, as the possibility of movement is ruled out, one has to stipulate for each cluster the order of the clitics.

In the second type of non-syntactic analysis, under which a distinct morphological component is active, the same questions arise as to why clitic position varies with respect to the type of verb (imperatives or not). Here also, the position of the clitic with respect to the verb must simply be stipulated. But the inadequacy of such analyses is more remarkable when one considers the facts in (5). These analyses rely on the existence of a template, which contains various slots to which clitics are anchored (one clitic per slot). The problem is that a template by its very nature does not have the flexibility to allow reordering of clitics. Thus in order to account for particular instances, it is necessary to change the template. In other words, one would have a template for each particular instance. These analyses are therefore inadequate to account for cliticization, at least in standard Romance languages.<sup>4</sup> I will return to the problems identified here, as they constitute serious challenges for the syntactic analyses also.

### 1.3.3 Syntactic analyses

As noted earlier, in what I refer to as syntactic analyses, the clitic is allowed to move from its D-structure position to another position. The movement analysis of clitics, usually referred to as the classical analysis because it was the first to be held in generative grammar, is constantly advocated in Kayne (1975, 1989, 1991, 1995), and is further developed with few modifications by Di Sciullo and Williams (1985, 1987), Di Sciullo (1990), Dufresne (1993), Laenzlinger (1993), Ouhalla (1989), Sportiche (1989), Uriagereka (1995), etc. Basically, according to this analysis the

---

<sup>4</sup> Following among other Rizzi (1986), Brandi and Cordin (1989), Auger (1994) argues convincingly that subject clitics in Picard are in fact agreement markers prefixed to finite verbs. One of the strong arguments she gives is the fact that subject doubling is obligatory in Picard. She seems to be correct also in her analysis of Quebec popular French (Auger, 1995). I will not discuss these facts here.

clitic is base-generated in the same phrase structure position as an ordinary NP, and is moved to a position under the V node, an affix position as discussed above, or to some other functional position. Thus the derivation of a simple sentence such as *Jean le connaît* 'John knows him' would proceed from the D-structure in (6a), yielding (6b) after the movement of the NP-clitic (irrelevant details set aside).

- (6) a. Jean [VP[V connaît [NP le]]].  
 b. Jean [VP [V le [V connaît] t]].

As in the case of non-syntactic analyses, the adjunction of the clitic to the verb ensures that it can never have identical properties with ordinary NPs, for the latter could never appear in this structural position (i.e. under the  $V^0$  node). For example, putting the clitic under a  $V^0$  node allows the fact to be captured that clitics cannot be conjoined, nor separated from their head by any lexical element or relative clause; compare (7) and (8).

- (7) a. \*Jean **la** et **le** connaît.  
       Jean her and him knows.  
 b. \*Marie **le**, qui est architecte, fréquente.  
       Marie him who is architect goes with.
- (8) a. Jean connaît Marie et Paul.  
       Jean knows Marie and Paul.  
 b. Marie fréquente Paul, qui est architecte.  
       Marie goes with Paul who is an architect.

Despite its initial plausibility, this type of analysis raises the same type of problems as the non-syntactic ones. Consider the ordering of clitics in Romance imperative verbs. As noted above, clitics never precede positive imperative verbs. At least three solutions are available within such a theory: (a) one may say that the clitic is right-adjoined to the positive imperative verb; (b) the clitic is left-adjoined to the verb as usual, but the latter then moves without the clitic; and (c) the clitic adjoins to an element preceding the verb (a functional head), and the verb moves past the clitic to another higher functional head. The first option cannot be maintained on pure

conceptual grounds, as minimally a theory which allows an element to be either left-adjoined or right-adjoined to the same type of head, according to the desired output, may not be restrictive enough. Note that Kayne' s (1994) Linear Correspondence Axiom, discussed below, is an attempt to avoid this type of problem.

The second option is not satisfactory either; indeed, since the clitic is in an affix position (under the  $V^0$  node), moving the verb without its adjunct would be exceptional, given GB assumptions. However, this option is the one used by Rooryck (1992) in his attempt to account for enclitic ordering in Romance imperatives (see also Roberts, 1991). Central to his argumentation is Rizzi' s (1990) Relativized Minimality, which includes the notions of trace, A and A' po sitions, government, governor and c-command, and the claim that  $C^0$  has a different nature in imperative verbs (cf. Rivero, 1994). All these claims are internal to GB theory, and appear to be successful only if one assumes certain premises, namely that inflectional morphology occurs in the syntax via movement of the verb to functional heads such as Agr, T, etc., and that the nature of the functional head can change from one language to another and even within a single language.

As for the last option, a variant has been developed, for instance, by Terzi (1996b), relying on Rivero (1994), for clitics in Greek, which patterns like standard Romance languages. In her analysis, clitics are left-adjoined to an empty functional head, and the verb is incorporated into the clitic (i.e. left-adjoined to the clitic); the complex V+clitic moves to a higher functional head ( $C^0$  or  $M^0$ ). These analyses are similar to the others with respect to their premises and their assumptions, which I will discuss in the next section. However, an analysis that can account for the phenomenon of clitic ordering in Romance without these complex apparatus would be preferable, as we will see.

Furthermore, one may note that with respect to clitic doubling, the movement analysis does not give rise to particular problems in (standard) French where clitics and NPs are in complementary distribution. However, as noted by its detractors within the GB framework, if the clitic cooccurs with a full NP occupying precisely the slot it is supposed to move from (as in some Spanish dialects), and with which it agrees in gender and number, it is reasonable to question the movement analysis. In

effect, in the base component an insoluble problem arises in that two elements, the clitic and the NP, are competing for the same position. Such a problem, of course, disappears under the copy theory of movement (Chomsky, 1995), where elements in a syntactic chain are deleted. But now the problem is to explain in a principled way why some languages allow deletion while others do not, hence allowing doubling (and resumptive pronouns). Note that the base-generation analysis of clitics allows the NP and the clitic doubling it to cooccur, but it does not say why a language comes to use such a construction.

Apart from their difficulty in explaining the problem of clitic ordering, both syntactic and nonsyntactic analyses appear to be successful (within the limits of the GB framework) with the facts they were designed for, namely complementary distribution in (standard) French and clitic doubling elsewhere. Thus, one may be tempted to think that a unified treatment of cliticization is not possible because of the very nature of this phenomenon, and that each particular grammar will in any case need some ad hoc mechanisms to account for various idiosyncratic features observed cross-linguistically, as suggested in Klavans (1985). However, on a purely theoretical basis, a unified analysis of clitics is preferable. Such an analysis would have to explain many things, among others, why there is no clitic doubling in standard French, in contrast with Spanish, and why clitic doubling in Spanish is mostly (but not always) restricted to NPs preceded by a preposition, a fact known as Kayne's Generalization. Sportiche (1996) includes a proposal which aims to unify cliticization theory by reconciling both types of analyses. On the one hand, he proposes that clitics are base-generated in their position as lexical elements heading their own projection called Clitic Voice (CLV), but individually realized as Nominative Voice, Accusative Voice, etc. On the other hand, he proposes that the clitic is always related to an XP present in the structure, and which is generated in argument position. This XP, which may be either a null element (PRO) or a full NP, must move to [Spec,CLV], yielding a familiar relation of Spec-Head agreement between the clitic head and the XP in its specifier. Thus, this analysis contains a movement component, while it is no longer the clitic that moves, but rather its related NP.

A range of parameters allows different types of clitic distribution to be accounted for with respect to the related XPs. The parameters concern the nature of the XP, the clitic head, and the movement: all three could be either covert or overt. It is, then, easy to see that in the case of clitic doubling, the XP must be overt, while its movement to [Spec,CL] is covert. On the other hand, in the complementary distribution case, the clitic head must be overt, while the XP may not be phonetically realized, i.e. it is PRO. The difference between complementary distribution (French) and clitic doubling (River Plate Spanish) is illustrated in (9a) and (9b) respectively.

- (9) a. Marie [AccV PRO<sub>i</sub> [Acc le ...[VP voit t<sub>i</sub>]].  
       Marie sees him.  
       b. pro [AccV — [Acc lo ...[VP veo [NP a Guille]]]]. (Jaeggli 1982)  
       I CL saw Guille.

In (9a) the clitic is an X<sup>0</sup> heading the AccV. The object of the verb is PRO, which moves to Spec-AccV either overtly or covertly. However, in (9b) the clitic also heads its own projection, but the NP object must move covertly to [Spec,AccV]. Thus, the setting of parameters for the desired output allows all possible cases to be derived.

Unfortunately, Sportiche' s theory is also unsatisfactory in that it exhibits the same features as the others. As noted above, along the lines of most work in the GB framework, the descriptive features of clitics (see (1)) are simply taken as defining properties which do not need to be accounted for. That is, given the properties in (1), the problem is reduced to the mechanical representation of the sentence, where parameters are adjusted according to the desired output. Now, as Adams (1987: 29) points out, "[...] when one attributes a property to a parameter one is saying, in effect, that that property cannot be explained, at least linguistically." In fact, it is clear that languages differ from each other, and perhaps GB theory is correct in attempting to handle cross-linguistic variation with parameters. However, one feels that in most cases it is preferable (and in fact it is possible) to explain why a language chooses a particular value for a given parameter.

The strongest argument against this type of analysis is of an empirical order. Indeed, like the other analyses discussed above, Sportiche' s theory fails to address

puzzling facts observed in cliticization in French. Let us consider some of them. Firstly, in any person but the third, the dative clitic precedes the accusative; in the third person, the accusative clitic precedes the dative, as shown in (10). Secondly, with positive imperative verbs, clitics are enclitics and the strong forms *moi* and *toi* are used instead of the corresponding weak forms, *me* and *te*, as illustrated in (11). Thirdly, a first person accusative object and a second person dative object are impossible at the same time in any structure, as seen in (12).

- (10) a. Max **me le** montre. ' M. shows it to me.'  
 b. Max **le lui** montre. ' M. shows it to him.'
- (11) a. Regardez-moi! vs. \*Regardez-**me**! ' look at me!'  
 b. Donne-**le**-moi! vs. \*Donnez-**le-me**! 'g ive it to me!' .
- (12) \*Marie **me te** présente.  
 Marie presents me to you.

Unfortunately, Sportiche does not address these issues. Indeed he explicitly claims that he has "no light to shed on these questions." I do not know how such a theory could account for these facts. I believe, in fact, that only a theory which includes distinctive features for personal pronouns, as discussed in Bonet (1991), may account for these facts; traditional analyses must invoke various ad hoc rules.

We may also note that a common feature of GB-like analyses is that they take for granted the fact that pronominal clitics in Romance are in most cases special clitics in the sense of Zwicky (1977). In GB theory, it is assumed that clitics undergo the rule *move-alpha*, at least by proponents of the movement analysis, but it is unclear why the NP that the clitic refers to does not move. In the Minimalist Program (MP) features play a central role, but they are not used distinctively. Indeed, in the MP every element, whether pronoun or noun, comes into the syntax fully inflected with features such as Case, gender, number, etc. Their behaviour depends on the assumption that features can be either weak or strong. For instance, one can say that clitics have strong features allowing them to move overtly, unlike NPs. Although this view of movement is more plausible, it is less clear how to know whether a feature is

weak or strong.<sup>5</sup> (Notice that in a recent version of MP (Chomsky, 1998), the notion of strong feature is discarded.)

The fact that the problems considered above have either not been addressed or have not received a satisfactory answer prompts us to look for an alternative representation for linguistic expressions. Before doing so, it is worth considering how a linguistic theory can be both restrictive and explanatory. The next section is devoted to this question.

#### **1.4 Some remarks on the Principles and Parameters Theory**

I will begin by reviewing the different mechanisms that have been put forward in order to constrain linguistic theory, showing that in most cases they are not as restrictive as claimed, because of their stipulatory character.

In generative grammar, linguistic analysis traditionally has a supreme goal, namely explanatory adequacy. Explanatory adequacy is the account of what the speaker knows about his/her language, and how he/she acquires this knowledge; thus, it refers to the psychological plausibility of the statements of linguistic theory, and ultimately makes it possible to choose among descriptively adequate grammars. A non-controversial fashion to attain this goal consists in discovering cross-linguistic principles and universal parameters, both being components of what is referred to as Universal Grammar.

The quest for explanatory adequacy leads naturally to the elaboration of theories that are as restrictive as possible. Among others, the Principles and Parameters approach to grammar, in both its versions (GB and MP) is a well-known attempt to reach this goal. As mentioned above, GB theory possesses a set of components which are organized in an optimal way so as to minimize the representations and to reduce redundancy in the grammar. An ongoing project is to reduce the transformational apparatus. For example, Kayne's work is a constant effort

---

<sup>5</sup> Ralli (1997), however, contains an attempt to deal with the notion of strong and weak features in morphology. As she puts her proposal "a feature is considered to be strong if at least one of its values is overtly realized. Otherwise the feature will be weak." (p.8)



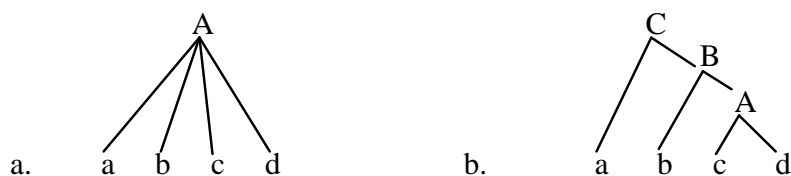
to restrict the syntactic representations available to learners. For instance, he proposes a strong condition on syntactic representation, namely that only binary branching is allowed (cf. Kayne, 1984), and recently, the Linear Correspondence Axiom (LCA), which is an attempt to drastically and in a principled way reduce syntactic representations and to derive the linear order of terminal elements (cf. Kayne, 1994).

How do these claims restrict the available syntactic representations for (young) learners? Consider the binary branching condition. A priori, it is preferable to any n-ary branching because it constitutes the simplest assumption, as I show immediately. In a grammar containing a rule of movement like *move-alpha*, where alpha may be a terminal element or not, the choice of a particular type of branching appears to have consequences for the derivation. Indeed if we have a set of terminal elements *a*, *b*, *c*, and *d*, immediately dominated by the same node, *A*, as illustrated in (13a), a movement rule could involve *A* (only if it is a component of a larger structure) or any single element dominated by *A*. But according to the internal logic of the representation, it must not simultaneously move two terminal elements, even if they are adjacent, because they do not form a constituent. On the other hand, in a binary branching structure, two adjacent elements may be moved together provided that they are dominated by the same node, that is, if they form a constituent. Thus in (13b), one can move either (a) *d* ; (b) *c* and *d* (=A); (c) *b*, *c* and *d* (=B); or (d) *a*, *b*, *c*, and *d* (=C). But the representation precludes moving *b* and *c* together without moving *d* also, because obviously only elements forming a constituent may move together. What I want to show with this example is the fact that, independently of any empirical content, the tree structure allows for some operations and excludes others. At the very least, for reasons of simplicity and economy, a grammar with a transformational component must favor (13b) over (13a).<sup>6</sup>

---

<sup>6</sup> It should be noted that Kayne (1984) reaches such a conclusion on a fairly different basis, namely the need for unambiguous paths in tree structures. Bouchard (1995) attempts to derive binarity in syntactic structure from the notion of one-element saturation. Thus while both Bouchard and Kayne rely on theoretical problems which arise within their respective frameworks, I base my argumentation on the nature of the representation, independently of empirical content.

(13)



Let us now consider the LCA, which is defined in (14). It is intended to heavily constrain syntactic representations. Indeed, it allows three conditions on linguistic expressions to be derived: (a) the basic word order in UG is Subject-Verb-Object, (b) only leftward movement is allowed in derivations, and (c) adjuncts are always to the left of the head. Furthermore, it allows the binary branching condition to be saved. While the latter is widely accepted by now, this is not the case for the three other conditions; however there seems to be some consensus about condition (b).<sup>7</sup>

(14) Linear Correspondence Axiom (LCA) (Kayne 1994: 33)

Let  $X$ ,  $Y$  be nonterminals and  $x$ ,  $y$  terminals such that  $X$  dominates  $x$  and  $Y$  dominates  $y$ . Then if  $X$  asymmetrically c-commands  $Y$ ,  $x$  precedes  $y$ .

I will not discuss the first condition, which requires cross-linguistic evidence. However, one cannot a priori exclude it, insofar as the movement of each category can be strongly motivated (but see Bouchard, 1994, for a critical evaluation of that approach). That is, it would be insufficient to invoke strength or weakness of features, as there are languages with no overt morphology which have the order SOV for example; e.g. Bambara, a language spoken in Mali (see Givón, 1984, and references therein).

Condition (b) appears to be weak, for it is not constrained by the nature of the representation. That is, given a syntactic structure, it is unnecessary to limit the movement in one direction only. What is needed, instead, is to explain why an

---

<sup>7</sup> The widespread consensus against rightward movement follows from the Empty Category Principle, which requires that a trace be governed by its antecedent. As government depends on the notion of c-command, the empty category must not be higher than the antecedent in the structure. However, as Bouchard (1995) points out, when the facts do not confirm this claim, the theory allows one to say either that the trace is deleted, that the element involved leaves no trace behind, or that the antecedent raises again at LF.

element moves, and why it does so in one direction rather than in another. The grammar would in fact be simpler if the direction of the movement were not arbitrarily limited to one option, thus allowing the movement of an element to take place according to its own position in a given structure.<sup>8</sup>

Now consider adjunction. The claim that an adjoined element must always be to the left of a head is also weak. When one adjoins an element B to a head A, a new segment of A is created, as opposed to a new category, and such a segment dominates both the other segment A and the element B (I will return to this in chapter 2). There are thereby two possible linear orders: either  $A > B$  or  $B > A$ . Consequently, admitting one linear order and ruling out the other is a decision which does not follow from the nature of the representation. Moreover, if only the other option were allowed, namely a ban on leftward adjunction, it would be difficult to see how this move should be incompatible with the representations.

Furthermore, we may note that the LCA relies on hierarchical structure in phrase markers, the existence of vacuous projections (i.e., non-branching bar-projections) of terminal elements, and the relation of asymmetric c-command between nonterminal projections. If one accepts the main axiom of GB theory, according to which the derivation proceeds from a rigid preexisting phrase marker, it is then clear that some hierarchical relationships are obvious (e.g., precedence and dominance), and even empirically grounded as far as basic constituency is concerned. However, the vacuous projection of terminal elements and the relation of asymmetric c-command raise some problems: the former is not empirically grounded (see Chomsky, 1994; Bouchard, 1995), and the latter is not logically constrained, as the definition is arbitrarily restricted to the first node up. Thus, it appears that the LCA exhibits the same problems as the left-adjunction-only and the leftward-movement-only claims.

So far I have considered the machinery of the theory, showing that when the representation offers two options for a given operation, one is arbitrarily eliminated

---

<sup>8</sup> Note that even in the GB literature, there is evidence for rightward movement; see for example Barbosa (1996).

in order to reduce the power of the grammar. By "arbitrary" I mean that one could have chosen the reverse, for instance rightward-movement-only, and right-adjunction-only without changing the theory or rendering it less plausible. Let us turn now to two basic tenets of GB theory: (cross-linguistic) principles and parameters. In the analysis of any phenomenon (linguistic or other), it is not necessary to have an explanation for all related phenomena. For instance, it was possible to produce some alloys before there was a satisfactory theory of atomic structure. Similarly, in linguistics it is not necessary to know why a language has a particular word order before studying some other aspects of its grammar. In fact, one can simply assume a particular word order as part of the evidence for an analysis. In this respect it is legitimate to consider the word order (irrelevant for the analysis) as a parameter subject to variation. The notion of parametric variation must then be treated as a heuristic procedure, and its interest is only typological. It should not, however, be taken as a final goal, as it explains nothing in itself. In other words, while one may legitimately assume that there is parametric variation in natural languages, since obviously languages are different from each other, by no means should a parameter be stated in order to account for a particular fact. From this point of view, we may note that Kayne's LCA, though unsatisfactory for reasons just discussed, is an attempt to explain what has been traditionally considered a parameter, namely variation in word order.

Consider now the principles in PPT. As a general rule a grammatical theory may contain some principles which are stated over the grammatical concepts, or the representation it uses (X-bar in the case of GB). So the principles used in GB are in themselves no better or worse than those of another theory; what counts in comparing and evaluating competing theories is, among other things, the simplicity and the number of principles, as well as the predictions they allow for. However, the way principles are used may weaken a given theory. Within GB theory principles are stated as absolute, that is, inviolable. But, as Speas (1997) shows, they always include some hedges, that is "a clause which extends the principle in order to cover problematic cases." (p.180) For instance, there is a principle in PPT known as Full Interpretation, which is given in (15).

## (15) Full Interpretation

There can be no superfluous symbols in an output representation.

By "superfluous symbols" is meant extraneous elements which are uninterpretable semantically and phonetically. Speas points out that sentences actually may contain such elements; the expletive pronoun in English, for instance, is such a word, see (16) (Speas' examples (6.13)). The theory deals with this problem by allowing the hedge in (17), as clarified by Speas (p. 181 and 184). Note that she further cites (p.184) twelve important principles of PPT (and MP) which contain hedges, including Case Theory and Binding Theory.

- (16) a. It is clear that constraints are inviolable.  
 b. There are three cats on the porch.  
 c. It rained all night.

## (17) Hedge of Full Interpretation

Expletives delete right before the point at which they are semantically interpreted, whereas they survive phonetically.

Clearly, it would be simpler to formulate the principles without any hedge, and one could thus assume that they are violable, somewhat in the manner of Optimality Theory (see below), as Speas argues.

Consider now Optimality theory (OT). In spite of its initial plausibility, it contains certain problems that cast doubt on its results. The fact is that on the one hand, a number of OT's constraints seem to be descriptive observations of the phenomena to be accounted for. On the other, a given constraint may be posited to outrank other constraints without independent evidence. To take an example, consider how OT uses a well-known principle of nonlinear phonology, namely the Obligatory Contour Principle. Spanish and Italian have a construction in which the reflexive clitics *se* (Sp.) and *si* (It.) appear as an impersonal subject, similar to French *on* and English *one*, as illustrated in (18).

- (18) a. Si sveglia Paolo. (Italian)  
 b. Se despierta a Pablo. (Spanish)

' One wakes up Paul.'

Now if the object is reflexive, the expected forms with double occurrences of *se/si* are ungrammatical (19). Instead, both languages use a distinct strategy, as shown in (20).

- (19) a. \*Si si sveglia. (It.)  
 b. \*Se se despierta. (Sp.)  
 ' One wakes up oneself.'
- (20) a. Ci si desperta.  
 b. (Uno) se despierta.  
 ' One wakes up oneself.'

Attempting to account for this fact in OT, Grimshaw (1997b) invokes the OCP, which, according to her, acts there as a dissimilation constraint, that is, the sequence of identical morphemes is prohibited: \*XX, as she formulates it. In her view, the ill-formedness of *si si* or *se se* is crucially due to the OCP's outranking any other relevant constraint. However, this use of the OCP is highly suspect here; in effect, it involves morphemes and not distinctive features, as in nonlinear phonology (see below). Besides, this interpretation of the OCP does not account for the fact that it is the first occurrence of the morpheme, and not the second, which dissimilates in Italian.

Furthermore, Grimshaw goes on to argue that the OCP is responsible for the spurious Spanish *se* construction. What is referred to as spurious *se* is the *se* which appears instead of the expected *le* (3.DAT.) in the cluster \**le lo* --> *se lo*. Again, a robust problem rises, which she recognizes herself: since both clitics (*le* and *lo*) are not identical, it is not obvious at all that the OCP, as defined for the impersonal *se*, is involved here (see also Heap, 1996, for a similar account).<sup>9</sup>

---

<sup>9</sup> In fact, Grimshaw speculates on three possible ways to see the nature of the \*XX constraint with respect to spurious *se*: (a) a phonological constraint against consecutive morphemes with identical onsets, given that each morpheme begins with *l*; (b) a morphological constraint which prohibits two [-Reflexive] clitics; (c) a template which contains only one slot for third person clitic, such as that proposed in Perlmutter (1971) and adopted in Bonet (1991). As I will show in chapter 4,

Consider another OT analysis in syntax. With respect to the problem of clitic ordering in finite and non-finite structures in Romance, Anderson (1996) suggests a solution along the following lines. In finite structures, a constraint of the form  $\text{EDGEMOST}(\text{cl}, L)$  aligns the clitic with the left edge of the relevant syntactic domain, ensuring that it will occur to the left of the verb. In non-finite contexts, where the clitic appears to the right of the verb, another constraint will require that the imperative in French and both the imperative and infinitive in other Romance languages be at the left edge of the same domain, yielding a constraint conflict. This conflict may be resolved by ranking the latter constraint higher (infinitive left), which entails a violation of the former. Although this analysis has the merit of being simple and clear, it does not explain why both clitics and non-finite verbs should be leftmost in some languages.

To conclude, I have attempted to show that the principles of PPT are somewhat weak because they always include some hedges (see Speas, 1997), and that parameters, though useful, should not be taken as definitive explanations. Even within the PPT, one observes cases where progress has been made by accounting for what has been treated as a parameter. Indeed, we have seen that the LCA is an attempt to account for the word order parameter in a principled way. Furthermore, it should be noted that some principles and constraints seem quite arbitrary to the extent that their object can be changed without affecting the plausibility of the theory. For instance, the OT constraint  $\text{EdgeMost}(\text{cl}, L)$ , mentioned above, could be replaced by another constraint of the kind  $\text{INNERMOST}(\text{cl})$ , without affecting the plausibility of that theory. This type of question is not new; for instance Chomsky and Halle (1968) offer a serious objection to their own phonological theory, as they open the ninth chapter of SPE as follows:

"The entire discussion of phonology in this book suffers from a fundamental theoretical inadequacy. Although we do not know how to remedy it fully, we feel that the outlines of a solution may be sketched, at least in part. The problem is that our approach to features, to rules, and to evaluation has been overly formal. *Suppose, for example, that we were systematically to interchange features or to replace  $[\alpha F]$  by  $[-\alpha F]$  (where  $\alpha = +$ , and  $F$  is a feature) throughout our description of English*

---

spurious *se* is not triggered by the OCP, although its placement in preverbal position is due to the OCP, like any other clitic pronoun.

*structure. There is nothing in our account of linguistic theory to indicate that the result would be the description of a system that violates certain principles governing human languages.* To the extent that this is true, we have failed to formulate the principles of linguistic theory, of universal grammar, in a satisfactory manner. In particular, we have not made any use of the fact that the features have intrinsic content." (p.400, emphasis mine)

Although this critique addresses phonological theory, its scope extends to many contemporary syntactic theories. If we were to replace the feature F in the italicized portion of the quote by the principle P of the theories just discussed, for instance, and proceed to change the value of alpha in the same manner, it would be impossible to say why the resulting system should be less plausible or natural.

To summarize, I have attempted to highlight some current assumptions which are intended to restrict the power of the grammar. It appears that they are not so restrictive in that they do not follow from the representation. In the next section, I present the assumptions I build on.

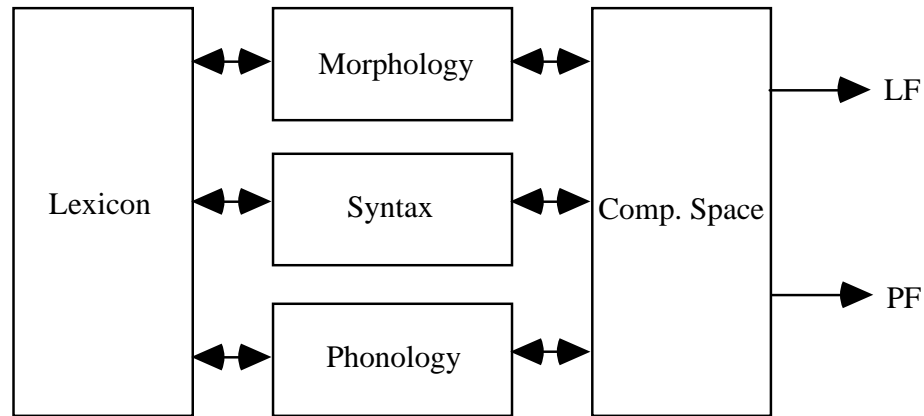
## **1.5 Assumptions**

### **1.5.1 The model of grammar**

In this thesis, I will assume the model of grammar proposed in Di Sciullo (1996). This model includes the following components: the lexicon, the syntax, the morphology and the phonology. The grammar also includes a computational space, which branches to the levels PF and LF. This highly modular model of grammar is diagrammed in Figure 1.2. (The size of the boxes is not significant, nor is the order of the components.)



**Figure 1.2** Di Sciullo's 1996 model of grammar.



Each component includes a set of principles which deal with different parts of linguistic expressions. The lexicon is connected to the other components, but not to the computational space, as can be seen in the diagram. The computational space is the locus where all operations occur; it receives instructions from the morphology, the syntax and the phonology. All types of instruction are processed simultaneously, not linearly, even though they interact with each other, consistent with the Modularity of Computational Space Hypothesis (see below).

In this model of grammar, local asymmetrical relations play an important role. Di Sciullo (1998/to appear) defines asymmetrical relations as in (21) and the notion local domain on which it relies in (22). As discussed below, the verb is central in the structure in that it is the element which projects after each merger. I take the verb to be in a local asymmetrical relation with each argument. The verb precedes the internal argument, and both define the complement domain of the verb (see below). The VP follows the subject and therefore both are in a local asymmetrical relation. This relation has some consequence for the interaction of features, as we will see below.

## (21) Local asymmetrical relations

x and y are in a local asymmetrical relation, iff there is a formal relation r that is true for the pair (x,y) and false for the pair (y,x) in a local domain.(p.2)<sup>10</sup>

## (22) Local domain

A local domain is a domain defined by immediate precedence and immediate dominance.(p.2)

In this thesis, I will be concerned mainly with the phonological and the syntactic components, as well as the computational space, and to some extent the lexicon. The morphological component and the interfaces PF and LF, which deal respectively with the articulatory system and the interpretation of structures, will not be discussed in detail. In the next subsection I turn to consideration of the phonological component, presenting the basic tenets of nonlinear phonology, which I build on.

### 1.5.2 Nonlinear phonology

Designed initially for tone phonology in African languages (Goldsmith, 1976), nonlinear representation has been readily extended to nonconcatenative morphology (McCarthy, 1979, 1981), and to every aspect of phonology (Sagey, 1986; Halle, 1992; etc.). The basic axioms of nonlinear phonology are listed below. For the sake of simplicity, I illustrate these axioms with tonal phonology.

- (23) a. Tones and tone-bearing units (or positions in a string) are independent of each other, indeed they each appear on different tiers.
- b. Tones and tone-bearing units are linked by association lines.

In many African languages, each morpheme is accompanied by a tonal melody which contributes to its meaning. For instance, a language may have three types of tones: high tone (H), low tone (L) and a contour tone. The latter is a combination of

---

<sup>10</sup> The page number refers to the 1998 version of her paper. The definitions are stated in terms of implications. It is not in the scope of this dissertation to discuss the difference between the two definitions of asymmetry.

the former: either HL (falling tone) or LH (rising tone). For instance, consider the following facts from Margi (cf. Kenstowicz, 1994), where the acute accent indicates a high tone, and the grave accent a low tone.<sup>11</sup>

- (24)
- |    |        |             |    |       |            |
|----|--------|-------------|----|-------|------------|
| a. | tsá    | 'b eat'     | b. | dlà   | ' fall'    |
| c. | ndábyá | ' touch'    | d. | gèrhù | ' fear     |
| e. | pèzú   | ' lay eggs' | f. | hɸ    | 'g row up' |

In contrast to linear phonology, where the tone is treated as a feature of the vowel, in nonlinear phonology, the monosyllabic examples (24a) and (24b) can be underlyingly represented as (25a), then as (25b), consistent with the basic claims (23a) and (23b), respectively.

- (25)
- |    |     |     |    |     |     |
|----|-----|-----|----|-----|-----|
|    | H   | L   |    | H   | L   |
|    |     |     |    |     | \   |
| a. | tsa | dla | b. | tsa | dla |

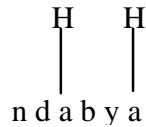
Given the representations in (25), where tones are independent of the syllables bearing them, the bisyllabic example (24c) and (24d) can be logically conceived of as having the structure in (26) underlyingly, where a single tone feature is associated with two positions within the morpheme. To the extent that a general principle of economy holds in the derivation, this multiple linking in (26b) is preferable to the one in (27), since it contains fewer occurrences of the feature. Indeed, in the latter there are two occurrences of the High tone, while in the former there is a single occurrence.

- (26)
- |    |        |    |        |
|----|--------|----|--------|
|    | H      |    | H      |
|    |        |    | \      |
| a. | ndabya | b. | ndabya |

---

<sup>11</sup> In the transcription of Margi, the letter *ɸ* stands for *ɸ* for convenience.

(27)



The representation in (27) is ruled out by what Goldsmith, relying on an idea by Leben (1973), refers to as the Obligatory Contour Principle (OCP), a formulation of which is given in (28). This structural constraint, which is motivated by reasons of economy, has considerable syntactic interest, as I will show in this thesis.

(28) Obligatory Contour Principle (OCP)

Adjacent identical features are banned in the autosegmental representation of a morpheme.

Now, in the previous example, a single feature is linked to two positions. Another logical possibility is the linking of two features to a single position. It is indeed such a possibility that gives rise to the contour tone (24f), where a low tone and a high tone are associated with a single position, as shown in (29).

(29)



Now consider the bisyllabic example in (24) which exhibits two different tones. This case can be initially represented as in (30a). Associating each tone with a vowel yields the surface representation (30b).

(30)



Notice that (30b) is the desired output, but it is not the only possible output, given the representation. That is, from (30a) one could also have either (31a) or (31b) if one were not looking for a particular output. Thus the representation must be

constrained so as to rule out these alternatives. Indeed, in order to prevent structures such as (31), Goldsmith proposes two further axioms, namely a well-formedness condition on the representations (32), and a manner of association, referred to as the Universal Association Convention, (33).

(31)



(32) Well-Formedness Condition (WFC)

Association lines must not be crossed in autosegmental representations.

(33) Universal Association Convention (UAC)

- a. Associate tones with segments one-to-one and from left to right.
- b. If there are more segments than tone, spread the rightmost tone to the toneless segments.

The WFC appears to be logically suitable, and is highly plausible to the extent that it would go against common sense if crossing of association lines were allowed. One may note further that it holds in every aspect of autosegmental representation, and it seems to be true in all languages. As for the Association Convention, it rules out a structure such as (31b) even if the language allows contour tones elsewhere. The usefulness of these two conventions becomes more obvious in cases where there are more positions than tones. Consider the hypothetical example in (34a), where a tonal melody LH is to be associated with a three-syllable morpheme.

(34)



By associating the tonal melody consistently with (33a), the representation (34b) obtains where the rightmost tone spreads onto the extra position. By virtue of the WFC the low tone cannot spread onto the rightmost syllable because it would cross

the high tone association line. Indeed, in tone languages, a melody such as LHH or HLL is current, but one never finds a melody such as HLH or LHL, which would entail the spreading of the leftmost tone.

However, (33a) appears to be too strong, as it is not constrained by the representation. One may ask why one could not have a one-to-one association proceeding from right to left, such that the L tone is associated to the remaining syllables, as shown in (35). There is nothing, in fact, in the representation which prohibits a leftward spreading as in (35). Indeed, further investigation into autosegmental phonology has shown that the initial claim in (33a) is too strong; given the representation, it is similar to leftward-only-movement in GB syntax.

(35)

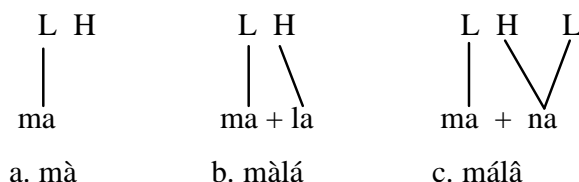


Here, the point is that one can stipulate some structural prohibitions that obviously constrain the representation or the operations one can perform on it, but such constraints must be plausible and grounded in the representation. That is, even though one does not find languages with left-to-right association, it is necessary to take into account the other alternative which follows from the architecture of the representation.

To sum up, autosegmental representations allow two types of linking of tones with segments: one-to-one, and one-to-many. The latter is composed of two cases: either one tone is linked to many segments, or two tones are linked to a single segment. All these results follow logically from the initial axiom. Yet there is another logical possibility which appears to be empirically supported: a tone may remain floating, in the sense that it is not linked to any segment. Let us illustrate this with the hypothetical example in (36). Assuming that a one-syllable word has the special tonal melody LH, and that the association convention associates it left-to-right, the representation in (36a) obtains, where H is not associated (i.e. it remains floating). Evidence for its presence arises when it is followed by another morpheme. By virtue

of the spreading process, if the latter is toneless it will surface with a high tone (36b), if it bears a low tone it will surface with a falling tone (36c).

(36)



In a later section, I will explore the hypothesis that syntactic features can be represented in a nonlinear fashion and in addition are subject to similar constraints, for instance the OCP, to the spreading process, and so one.

### 1.5.3 Constraint interaction

Another basic concept that I assume is the notion of constraint interaction and the violability of constraints. One finds this claim in many modern phonological theories, including Paradis' Theory of Constraints and Repair Strategies, and Prince and Smolensky' s Optimality Theory, which I have already discussed above (see Paradis and Lacharité, 1993, for a comparison between these phonological theories, and references).

As discussed earlier, Optimality Theory, which has been adopted by many syntacticians, assumes a set of universal constraints which are variably ranked in individual languages. I would like to explore another view of constraint violation. I will distinguish representational-type (or structural-type) constraints from semantic-type constraints (or interpretive-type). The former include the OCP, the Structure Preservation Constraint, etc., and are not ranked with respect to each other. As I proceed, I will provide evidence to the effect that operations triggered by representational-type constraints apply simultaneously, but that such operations are blocked if they yield a constraint violation which cannot be put right.

Unlike structural-type constraints, semantic-type constraints deal with the semantic interpretation of the structure. For instance, the Animacy Hierarchy (see chapter 2), as well as the interpretation of anaphoric elements with respect to their

referent (see Desouvrey, 1996b, 1997), are such constraints. In a given derivation these constraints must be satisfied in order for the structure to be properly interpreted. Furthermore, some features are more semantic-like and will take precedence over others. For instance, it will be shown that a scopal element (having a feature referred to as  $\pi$ ) may override any principle requiring it to move (see the discussion of enclitics in Spanish and wh-words in French in chapter 4). In other words, semantic-type constraints outrank structural-type constraints. As a result, a structural-type constraint will fail to apply where it compromises the interpretation of the structure.

The interaction of both types of constraints follows from the interaction of the various modules of the grammar. In Di Sciullo's (1996) model, the interaction is due to a central property of the architecture of the grammar, namely the Modularity of Computational Space (MCS), which is defined below.

(37) Modularity of Computational Space Hypothesis (DiSciullo, 1996: 2)

The computational space includes interacting types of derivations leading to optimal target types of configurations.

To conclude, semantic-type constraints must be satisfied before phonology-type constraints. Beyond this, there is neither constraint ranking nor transformation ordering. All transformations apply simultaneously and a maximal number of constraints must be satisfied in the derivation. Some special cases which seemingly challenge this claim will be pointed out as we proceed.

#### 1.5.4 Summary

Although my assumptions are different from those in most work in the field, they are not new. Nonlinear phonology is a well-established theory and some syntacticians have attempted to adapt its insights to syntactic phenomena, as we will see in the next section (cf. Yip et al. 1987). The notion of constraints is rather familiar, as it is well-discussed in the literature. Moreover, my proposal is part of a larger research program in which asymmetrical relations and modularity are crucial. In the next section, I will develop my proposals, and present a number of other assumptions.



## 1.6 The proposal

### 1.6.1 Case features

Syntactic features such as Case and thematic roles play a crucial role in GB theory.<sup>12</sup> Indeed, an important module of this framework is labeled Case theory, and it is intended to account for the positions in which NPs may appear or not, i.e. the positions where they may be assigned a Case and a thematic role (argument (A)-position). There is a distinction between structural and inherent Case. Nominative and accusative are (abstract) structural Cases (at least in English-like languages, including French), that is, they are assigned abstractly by a Case-assigner, namely the functional head Tense (nominative) and the verb (accusative), under the general condition of government. Unlike structural Case, inherent Case is linked to a thematic role. Nouns and adjectives are assumed to be inherent Case assigners. An inherent Case assigner Case-marks an NP if it governs and theta-marks that NP (cf. Chomsky, 1980, 1986a; etc.)

In order to attain my objectives in this thesis, I am led to explore a slightly different view. Assuming a radical underspecification theory in which there is an opposition between a feature and its absence (cf. Trubetzkoy, 1967; Kiparsky, 1982; Archangeli, 1988; Pulleyblank, 1988; etc.), I assume that lexical elements may either bear an inherent Case or bear no Case at all.<sup>13</sup> That is, a Case is necessarily inherent

---

<sup>12</sup> The terms thematic roles / features, semantic roles / features are synonymous. They will be sometimes replaced by the shorthand terms  $\Theta$ -role and  $\Theta$ -feature, following the literature.

<sup>13</sup> This view of underspecification, according to which there is an opposition between a feature and its absence, for instance, [F] vs. [ ], is the simplest one can have. In contrast, an equipollent theory of underspecification would allow for a given feature [F], a positive and a negative value, for instance [-F] and [+F], in all of its manifestations. And it is not logically excluded that many elements bear a negative value, that is, elements that are not specified [+F] are necessarily [-F]. On this view, in a given paradigm of four elements A, B, C, and D., if A is [+F], then the other elements must be [-F], obviously an unnatural result. This problem disappears under a radical theory of underspecification, as assumed in this thesis, for one can simply allow one element to be specified [+F], while the others are not specified for F. However, Rooryck (1994) argues for an even more complicated theory of underspecification: a three-valued system. In his system, elements are either specified negatively and/or positively for a feature, or are underspecified. In the latter case, underspecification may be either variable or nonvariable: a nonvariable feature is equivalent to 0-Feature, while a variable feature may be either positive or negative. Rooryck suggests that variable features 'are inherently flexible and open for the features surrounding them', as opposed to 0-Features

to the element bearing it; a lexical element enters the syntax with its inherent Case, if any. In this sense, an element with inherent Case will be said to be specified for Case, as opposed to a Caseless element, which is not specified for Case (i.e., it has no inherent Case).<sup>14</sup>

An argument which does not bear Case inherently may be assigned Case by a Case-assigner, namely a verb or a preposition.<sup>15</sup> I claim that some Cases are related to a thematic feature, while some others are not. Assume that in French (and similar languages) the grammatical relations of subject and direct object (DO) are related respectively to nominative Case and accusative Case, while the indirect object (IO) is related either to accusative Case or oblique Case (see Marantz, 1984, for a theory of grammatical relations). I take accusative Case corresponding to the DO to be unspecified for thematic feature. Oblique Case and accusative Case corresponding to an indirect object are related to some thematic role, and are marked by certain prepositions. In this sense, those prepositions are overt Case/theta-markers.

I will show that a constraint on merger requires that the Caseless complement of a verb be introduced by a relevant preposition, if any, in order to agree with the head. When a verb bears a Case not related to a thematic feature (call this a bare Case), the argument will not be marked by a preposition, presumably because there is no preposition which bears a bare accusative Case. The argument which is assigned a

---

which are neutral. The evidence he attempts to adduce in favor of this theory is mainly related to the *that*-trace effect in English and *que/qui* alternation in French (he assumes the GB theory). Since I have presented elsewhere an account of these phenomena, where a radical underspecification theory has successfully been used, Rooryck's theory is highly questionable. (On the so-called *that*-trace effect in four languages, including English, French, Italian and West-Flemish, see Desouvrey 1996b.)

<sup>14</sup> To set the terminology, an element bearing a Case (or other features) is said to be specified for Case (or for the feature it bears). An element which does not bear Case is unspecified for Case. In this sense, I will use also the term underspecified, which in my view is equivalent to nonspecified. A Case is necessarily inherent, hence lexical. Case assignment, which involves the saturation of the Case-assigner, will be discussed in chapter 2. Also, as mentioned above, the term clitic will be used as a shorthand for Case-specified pronoun, but recall that in my view it is not a grammatical category such as verb, adjective, etc. The term NP, as opposed to clitic in the sense assumed here, will be used conveniently to refer to a noun or a noun phrase, for instance *Mary* and *the book*, which does not bear a Case as defined above. Thus the terms clitic and NP do not entail an opposition between NP and DP or head and nonhead, as in GB theory.

<sup>15</sup> This is not obligatory, as an *argument* may remain without Case, as I will show. Thus the Case filter is not needed.

bare accusative Case will be interpreted as the theme by default. All types of verbs in French assign a bare nominative Case to the subject, although they may vary according to the thematic feature they assign to the subject. Thematic features which are assigned to the subject are never marked via prepositions, parallel to the accusative Case, as no preposition is specified for nominative Case (see below). The following Table summarizes the proposal.<sup>16</sup>

**Table 1.1** Relations of grammatical roles to Case and  $\Theta$ -features.

	Case	$\Theta$ -feature	overt marker
Subj	Nom	yes	no
DO	Acc	no	no
IO	Obl/Acc	yes	yes

Notice that the relation between grammatical relations and Case is a semantic-type constraint, which may not be altered in the course of the derivation. As we will see, a structure is rejected whenever a mismatch arises between Cases and grammatical relations.

To conclude, I suggest that there is no distinction between structural Case and inherent Case. An element may be specified for Case or not. In French, an indirect object verb bears a Case which is related to a thematic feature, hence the use of a preposition as an agreement marker.

### 1.6.2 Syntactic features in tiers

Traditionally, it is assumed that syntax and phonology deal with different principles. This assumption is motivated by the fact that they deal with different objects: syntax is concerned with strings of words (organized in tree structure), whereas phonology deals with strings of phonemes or larger units including syllables,

---

<sup>16</sup> To repeat, Table 1.1 is valid for most verbs in French-like languages, and perhaps more generally nominative/accusative languages as opposed to ergative/absolutive languages. Thus, it should not be taken as universal. Indeed, the relations of grammatical roles to Case and thematic features vary crosslinguistically, as is well known. In Icelandic for instance, the subject argument may bear a non-nominative Case (quirky Case). I leave this matter for future research.

feet, etc.<sup>17</sup> However, it is not a logical necessity that this should be the case; nothing indeed prevents different objects from being treated with the same apparatus. In fact, if a theory treats apparently different objects alike, it should be simpler than another theory which does not, on purely conceptual grounds. Let us assume that syntactic objects and phonological objects are regulated by similar representations and may have some principles in common.<sup>18</sup>

If syntax is like phonology, one may think of the syntactic features borne by different elements as interacting with each other much more like phonological features. In order to give some substance to this idea, one can proceed in two ways. First, one can assume that a given syntactic structure is provided with an autonomous Case melody, similar to the tonal melody of most African languages, as discussed above. Yip, Maling, and Jackendoff (1987) argue for a similar theory. In their system, surface Cases are assumed to appear on a single tier in the order N[ominative], A[ccusative]. This is illustrated below with their examples (1) and (2). A sentence with a simple transitive verb has the configuration in (38a), once it is assumed that the Universal Association Convention associates Cases with NPs one-to-one and from left to right (at least in English). If the verb is intransitive, there are more Cases than arguments, and therefore the extra Case is ignored, consistent with standard assumptions in nonlinear phonology, (38b).

---

<sup>17</sup> As Chris Miller points out (p.c), "l' objet de la syntaxe est, d' un point de vue plus vaste, l' organisation temporelle des éléments lexicaux dans la phrase, ce qui se traduit par des principes et des contraintes qui régissent: (a) l' ordre séquentiel des éléments dans les langues et (b) l' encodage simultané, ou inversé, d' éléments dans des canaux primaire et secondaire, et donnant lieu aux encodages parallèles et aux inversions de dominance." He further notes that "le phénomène d' encodage parallèle est restreint aux langues signées, où chaque bras/main est potentiellement un canal d' encodage autonome." In my perspective, even in oral languages, syntactic structures are not linear; but their apparent linearity is forced by the nature of the output device, essentially the mouth (see chapter 2).

<sup>18</sup> Bromberger and Halle (1991) convincingly argue that the Principles and Parameters approach, which is prevalent in syntax, cannot apply successfully to phonology. They say nothing, however, as to whether or not syntax can be like phonology.

- (38) a.  $\begin{array}{ccc} & \text{N} & \text{A} \\ & | & | \\ \text{John} & \text{hit} & \text{Mary} \end{array}$       b.  $\begin{array}{cc} & \text{N} \quad \text{A} \\ & | \\ \text{John} & \text{paused} \end{array}$

This theory allows various phenomena involving Case in Icelandic, including the problem of quirky Case, to be handled successfully. Furthermore, under this theory, the difference between ergative and accusative languages follows from the direction of the association: left-to-right in accusative systems and right-to-left in ergative systems. I will assume that the Case-in-tiers hypothesis can be implemented in various ways across languages. In Icelandic and other similar languages, Case clearly forms an autonomous melody, as Yip et al. argue; but this may not necessarily be so in other types of languages, including Romance, which have a less developed case morphology.

Given the data I will deal with, I would like to explore another alternative. Under a strong lexicalist hypothesis, according to which inflectional morphology, as well as derivational morphology, occur in the lexicon or in a separate morphological component (see among others Di Sciullo and Williams, 1987), one can assume that a tensed transitive verb is specified for both nominative and accusative Cases. From this perspective, a simple transitive verb in French can be represented as in (39), where the N[ominative] Case and the A[ccusative] Case are each linked to the verb via an association line.

- (39)  $\begin{array}{c} \text{N} \\ | \\ \text{Paul aime Marie} \\ | \\ \text{A} \end{array}$

In (39), assuming that the NPs *Paul* and *Marie* are unspecified for Case, they must be assigned Case by the verb. This can be realized by assuming the Principles of

Spreading in (40) (adapted from Piggot, 1992)<sup>19</sup>, and, for the time being, the reasonable stipulation in (41). Thus the application of (40) and (41) to (39) yields the configuration in (42), where the verb assigns bare accusative and nominative Cases to the arguments. (I will return to the (b)-clause of (40).)

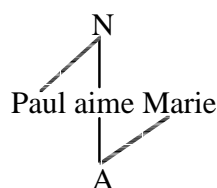
(40) Principles of Spreading (Piggot, 1992: 35)

- a. A feature (x) may spread only to a position not specified for (x).
- b. The spreading of a feature (x) may be arrested only by a position specified for (x).

(41) Direction of Case spreading

In French the nominative Case spreads leftward and the accusative Case rightward.

(42)



Notice that in the GB framework nominative Case is assigned by a tensed INFL head. This ensures that an infinitive verb, which contains a tenseless INFL, will not have an overt subject by virtue of the Case filter (43). Within the present proposal, the same result, if not a superior one, can be achieved by taking infinitive verbs to be specified for accusative Case only. Having no nominative Case, infinitive verbs are not allowed to select an overt subject in unmarked contexts (see below).<sup>20</sup>

(43) Case filter (cf. Chomsky 1980, 1981)

\*NP, if NP has a phonetic matrix and no Case.

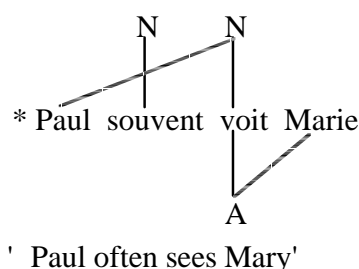
---

<sup>19</sup> As can be seen, this principle does not introduce any new thesis into the phonological theory; it is an instance of the basic axioms, namely the Universal Association Convention (UAC) and the Well-formedness Condition (WFC).

<sup>20</sup> Thus, since in some contexts an infinitive verb may have an overt subject, the Case-filter is not accurate, and indeed lacks generality (see section 1.6.5.2).

I will assume further that adverbs may be specified for Case (see chapter 4). Suppose that most French adverbs are specified for nominative Case. Then the ill-formedness of the structure in (44) is explained readily by (40b) (the No-Crossing Lines Constraint).<sup>21</sup>

(44)



Having briefly illustrated how Case features can be treated in the same manner as phonological features, I turn now to the organization of features.

### 1.6.3 Feature tree structure

Suppose that the lexicon is composed of various paradigms (see Di Sciullo, to appear). Under this view, personal pronouns, relative pronouns, and nominals form distinct paradigms, different from each other with respect to their feature matrices. In many languages, pronouns have richer morphology than nouns in that they bear various features including Case, person, number, etc. Most of the personal pronouns in (standard) French are given in Table 1.2. At first sight, these pronouns decline according to their grammatical functions (subject, direct object, and indirect object), as well as gender, person, and number. Traditionally, those pronouns which decline for grammatical relations are referred to as weak forms or clitics (see for instance Kayne, 1975), two terms we will use occasionally for convenience.

---

<sup>21</sup> As we will see below the initial position of an element in a structure depends upon its Case specification. An element with nominative Case will be generated to the left of the verb, and an element with accusative Case to the right of the verb. Thus if adverbs are specified for nominative Case in French, they must be initially to the left of the verb. If in a given language adverbs are specified for Accusative Case, they must appear to the right of verbs; we believe that this is indeed the case in English. Some aspects of the syntax of adverbs will be accounted for briefly in chapter 4, focussing on the adverb of negation in French.

**Table 1.2** Declension of personal pronouns in French.

	Subj	DO	IO	Strong forms
1.	je	me	me	MOI
2.	tu	te	te	TOI
3a.	il	le	lui	LUI (masc.)
3b.	elle	la	lui	ELLE (fem.)
4.	nous	nous	nous	NOUS
5.	vous	vous	vous	VOUS
6a.	ils	les	leur	EUX (masc.)
6b.	elles	les	leur	ELLES (fem.)

Let us take pronouns to be defined by a set of features, including person, number, gender, Case features, thematic features, etc. The first three will be referred to as  $\Phi$ -features, following common practice in the field. Case features include nominative [N], accusative [A], and oblique [O]. Thematic features comprise benefactive, goal and locative (as far as pronouns are concerned); goal and locative pronouns do not appear in Table 1.2.

I wish to suggest that pronoun features are grouped according to the tree structure in (45), which is analogous to feature tree structure in nonlinear phonology. The first element of the tree, namely pronoun, may be seen as the root node, and is a convenient label for a set of categorial features, which I set aside here (see Di Sciullo, 1999). The non-terminal nodes are much more similar to class nodes in nonlinear phonology.<sup>22</sup> As Goldsmith (1990: 280) points out, "[a class node] serves only as a geometrical way-station for the passage of information up and down the feature tree." The node AGR (for agreement) is intended to express the fact that pronouns, and generally most grammatical elements, may be combined with each other to form larger structures (see below). It has two dependencies, the nodes K and

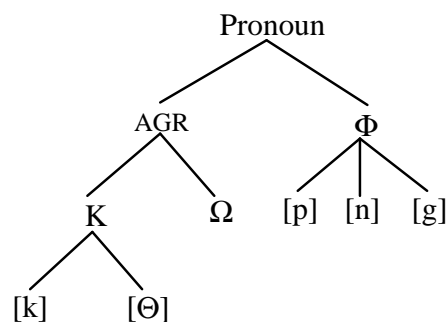
---

<sup>22</sup> This notion of class node may be related to the notion attribute (type of feature), as opposed to value (a particular feature) (see Steele, 1995). Such a view is defended in Ralli (1997), although she does not use a nonlinear-type of representation. See also Di Sciullo (to appear) for a configurational Type/Restrictor representation of feature asymmetry for inflectional, categorial and aspectual paradigms.



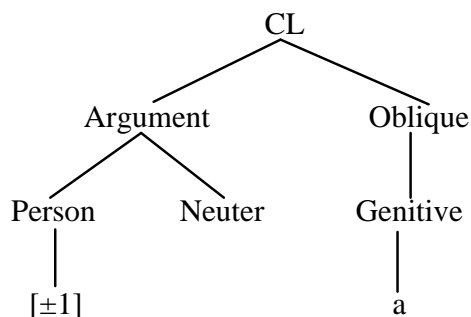
$\Omega$ . The former groups Case features and thematic features, represented by [k] and  $[\Theta]$  respectively, while the latter may host various types of features, including animacy, etc. Notice that each node appears on its own tier, and therefore, there is no linear order in the representation. That is, it would not matter if the node AGR appeared in the right branch of the root node and the node  $\Phi$  in the left branch.

(45)



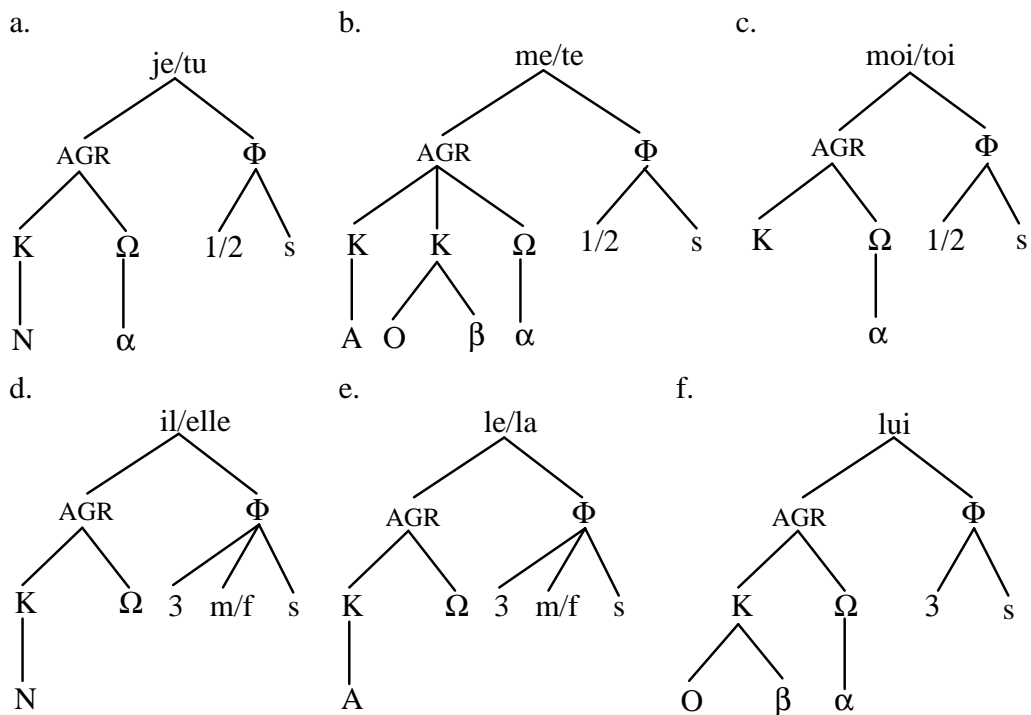
This structure is significantly different from that proposed in Bonet (1991), which is given in (46). According to her, each clitic (in Catalan) is a subset of (46). As she briefly describes it, "the nodes written in uppercase are defining properties of the clitics, while the information in lowercase refers to the agreement features (or other types of features) that the clitics can have in addition." She writes further that "the main purpose behind this structure is to reflect hierarchical relations as well as markedness." (p.16) We may note that the name clitic (CL) is explicitly mentioned in her representation, suggesting that she takes clitics to be a category on their own, like for instance adjectives and verbs. In our representation, in contrast, there are features only, and the term clitic is not mentioned. Indeed, as I will show, a pronoun will be considered a clitic only if it contains Case features. Furthermore, our tree structure is not intended to express markedness. I find no reason to assume that the more features a pronoun has the more marked it is. Rather, it appears that use of a Caseless pronoun yields a marked structure, as we will see below.

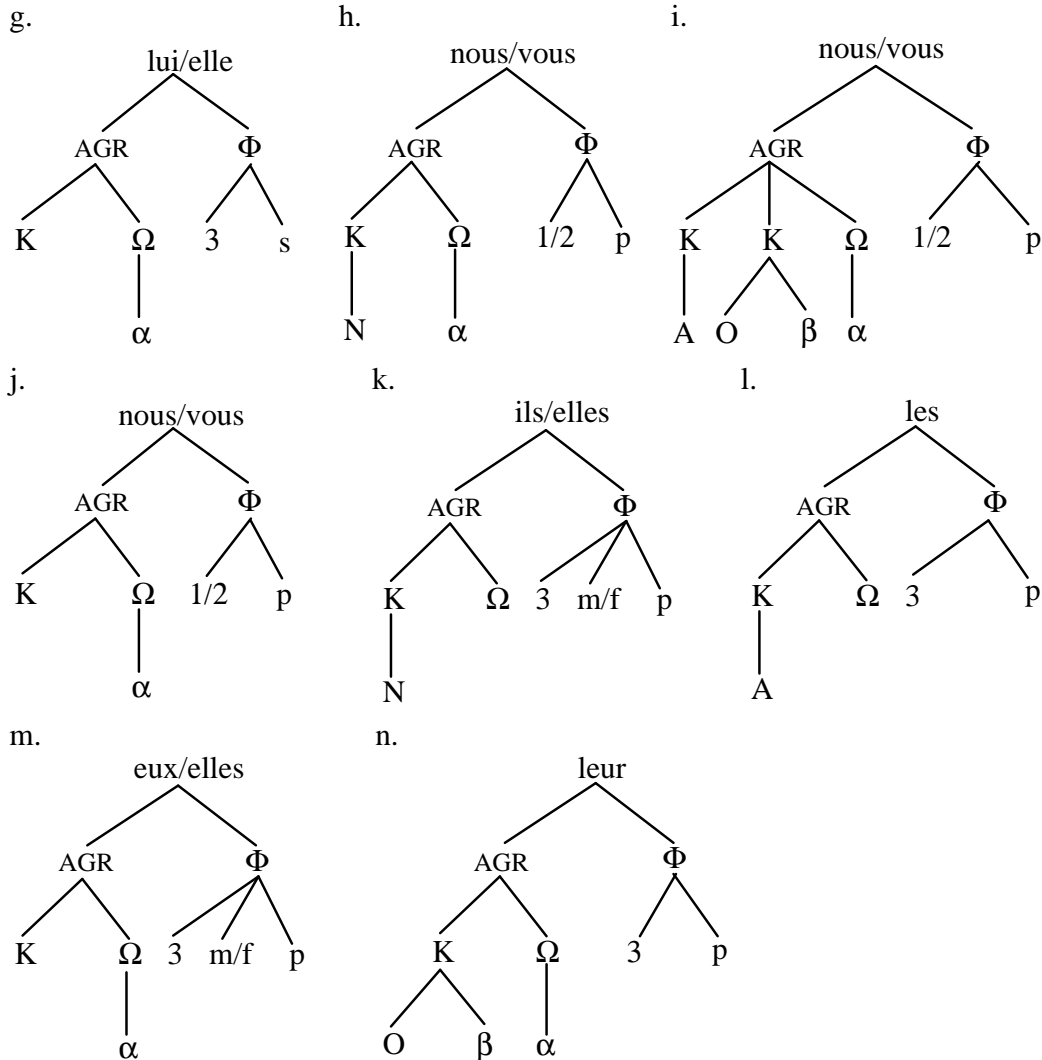
(46)



The feature structure of the pronouns appearing in Table 1.2 is given below, consistent with the model just outlined (see (45)). This feature structure will be motivated shortly. (The integers refer to person features. Conveniently, a single tree structure may be used for two different pronouns separated by a slash, as appeared in some root nodes (e.g., *moi/toi*), and similarly, the distinctive features of both pronouns in the root node are shown under the relevant node separated by a slash (1/2).)

(47)





Consider the third person singular pronouns, which are given in (47d-g). Traditionally, the pronouns in (47d, e, f) are considered clitic pronouns (or weak forms), while the pronouns in (47g) are referred to as strong forms. The subject pronoun (see (47d)) is specified for nominative Case [N] but not for thematic features; it is also unspecified for animacy, hence the lack of animacy features under the  $\Omega$  node. Under the  $\Phi$  node, the features [s] and [m] (*il*) / [f] (*elle*) stand for singular, masculine and feminine respectively. (47e) is identical to (47d), except that the accusative feature appears under the K node. In (47f) the K node dominates two features: oblique Case and the thematic feature [ $\beta$ ] (benefactive or malefactive), and

the  $\Omega$  node dominates the animacy feature  $[\alpha]$ ; but in contrast to the previous pronouns the  $\Phi$  node has no gender feature. The strong pronouns (47g) are unspecified for Case; the masculine form, *lui*, is syncretic with the oblique benefactive pronoun (47f), while the feminine form is syncretic with the nominative feminine pronoun (47d). As we will see, in each case there are actually two different pronouns, each having a set of divergent properties. In the rest of this section, I will provide evidence to the effect that the feature structure assumed for each pronoun in (47) is plausible.<sup>23</sup>

Given my proposal that some elements are specified for Case while some others are not, I will construct an argument on the natural assumption that an element which does not bear Case (i.e. which is unspecified for Case) is likely to have a larger distribution than one which bears Case. This must be mediated by the notion of grammatical relations, as assumed above. Thus, an NP such as *Paul*, which does not bear Case, may appear in the following contexts: left dislocation (48a), answer to a yes-no question (48b), subject of a verb (48c), and object of a verb (48d) or a preposition (48e).

- (48) a. **Paul**, il vient d' arriver.  
       ' Paul, he just came.'  
       b. Qui as-tu vu? **Paul**!  
       ' who do you see? Paul!  
       c. **Paul** est venu.  
       ' Only Paul came.'  
       d. J' ai vu**Paul**!  
       'I saw Paul!'

---

<sup>23</sup> Notice that many analyses of clitics rely on features, which are only stipulated, since they are not empirically motivated, as they are here. For instance, Laenzlinger (1993) assumes the following features for French weak pronouns. On such an assumption, the term clitic must be used as a diacritic to discriminate between strong pronouns and weak pronouns.

(i)	me:	[-gender, +number, +person],	[-Case]
	te:	[-gender, +number, +person],	[-Case]
	nous:	[-gender, +number, +person],	[-Case]
	vous:	[-gender, +number, +person],	[-Case]
	se:	[-gender, -number, +person],	[-Case]
	le/la/les:	[+gender, +number, +person],	[+Case (ACC)]
	lui/leur:	[-gender, +number, +person],	[+Case (DAT)]

- e. Selon **Paul**, ...  
 'According to Paul...'

Now, strong pronouns mostly pattern like ordinary NPs. If indeed in the sentences in (48) one replaces the NP *Paul* by the strong pronoun *LUI*, a grammatical structure obtains in each case, as shown in (49). But if, instead, the NP is replaced by the weak pronoun *le*, an ill-formed sentence results in each case, as illustrated in (50). Since the distribution of strong pronouns is quite parallel to that of NPs, one may conclude that they do not bear inherent Case.<sup>24</sup>

- (49) a. LUI, il vient d' arriver.  
           'h im, he just came.'  
       b. Qui as-tu vu? LUI!  
           ' who did you see? him!  
       c. LUI est venu.  
           'him came.'  
       d. J' ai vuLUI! (strongly stressed)  
           ' I saw him!'  
       e. Selon LUI, ...  
           ' According to him'

- (50) a. \***Le**, il vient d' arriver.  
           'h im, he just came'  
       b. \*Qui as-tu vu? **Le**!  
           ' who did you see? him!'  
       c. \***LE** est venu  
           'him came'  
       d. \*J' ai vu**le**!  
           ' I saw him!'

---

<sup>24</sup> An evaluator objects that this discussion is circular and does not demonstrate that strong pronouns are unspecified for Case, because I use a 'tenuous natural assumption', namely that an element which does not bear Case is likely to have a larger distribution than one that bears Case, but not a demonstrated fact. I accept her objection as it is, since at this point I do not know how to demonstrate independently of this assumption that an element bears Case. (On Case, see Blake, 1994.)

- e. \*Selon **le**, ...  
'According to him'

It is clear that the strong pronouns are not specified for Case, unlike the weak pronouns. I will provide empirical evidence for the alleged Case features of the weak pronouns, i.e. I will show why it is more plausible to take the pronouns *me* and *te*, for instance, to be specified for both [A] and [O,β], and the pronoun *je* to be specified for [N] and not for [A]. Consider the weak pronouns of the third person. If it is true that the pronouns *il* and *elle* are specified for nominative Case (see (47d,k)), they must always be used as the subject of their clause, since nominative Case is related to the subject (at least in French-like languages). Indeed *il* and *elle* correspond to the subject of well-formed sentences, as shown in (51). Therefore, one can conclude that they are specified for nominative Case. (Furthermore, since these pronouns may refer either to a person or a thing, it must be the case that they are unspecified for animacy.)

- (51) a. **Il** a nettoyé la maison.  
He cleaned the house.  
b. **Elle** a embrassé sa mère.  
She kissed her mother.  
c. **Ils** ont bombardé l'hôpital.  
They bombed the hospital.

Similarly, given the assumption that accusative Case and oblique Case correspond respectively to direct and indirect object, if the pronoun *le* is specified [A] and *lui* [O,β], they must be used only as direct and indirect object respectively. This is borne out, as can be seen in (52).<sup>25</sup> Therefore, I conclude that *le* is specified accusative and *lui* oblique-benefactive. While the accusative pronoun can refer either to animate beings or things, the oblique pronoun may refer only to animate beings. Therefore I assume that the latter, but not the former, is specified for animacy.

- (52) a. Paul **la** connaît. / \*Paul **lui** connaît. (DO verb)

---

<sup>25</sup> The idiomatic English translation, which is usually noted in simple quotes '...', is sometimes omitted when the meaning is obvious from the gloss.

- Paul knows her-A. / Paul knows her-O.
- b. Paul **lui** nuit. / \*Paul **la** nuit. (IO verb)  
Paul disturbs her-O. / Paul disturbs her-A.
- c. Marie **le** / \***lui** montre à Paul.  
Marie shows it to Paul.
- d. Marie **lui** / \***le** montre le livre.  
Marie shows the book to him.

Notice that the only difference between the strong forms LUI/ELLE and the weak forms *il*, *le/la*, *lui* is the lack of Case features in the former. I take this to be the main distinction between clitics and non-clitics. More generally, as I will show, cliticization in Romance pronoun systems obeys (53). As will be discussed later, clitics are partly licensed by virtue of (53b).

- (53) Condition for cliticization
- A pronominal element is a clitic iff it is specified for Case, that is, if it bears an inherent Case.
  - A pronominal element may cliticize onto a verb iff the latter bears the relevant Case.

The argument that I have developed above for the third person singular pronouns holds also for the third person plural pronouns (47k-n): their distribution is indeed identical, except that there is no gender distinction in the plural for the accusative pronouns.

Consider now first and second person singular pronouns whose feature structures are given in (47a-c). I take the pronouns *je* and *tu* to be specified for nominative Case since their distribution is strictly restricted to the subject position, exactly as the third person *il/elle*. Thus, when *je* and *tu* are used as subjects, well-formed sentences obtain, as illustrated in (54). But as expected, if these pronouns are used in any other position or in isolation, ill-formed structures obtain, as illustrated in (55).

- (54) a. Connais-**tu** Marie?

Do you know Marie?

- b. Hier, **je** suis allé au cinéma avec Marie.  
Yesterday, I went to the movies with Marie.

- (55) a. Marie est allée *te*/**\*tu** voir ce matin.  
Marie went to see you this morning.  
b. Toi/**\*tu**, as-**tu** vu Marie?  
Have you seen Marie?

Let us turn now to the pronouns (*me* [1] and *te* [2]). I assume that these pronouns are specified for two Cases, namely accusative and oblique. This is represented with two K nodes in (47b): one dominates [A], and the other dominates [O,β]. This is supported by the fact that these pronouns can be used either as arguments of a direct object verb or an indirect object verb. For instance, the verb *voir* 'to see' is a DO verb and *obéir* 'to obey', an IO verb. As seen in (56), the pronouns *me* and *te* correctly appear as complement of either of these verbs.

- (56) a. Paul est venu **me** voir.  
Paul came to see me.  
b. Paul ne veut pas **m'**obéir.  
Paul does not want to obey me.

With verbs taking both a direct object and an indirect object (double object verb), the pronouns *me* and *te* can be either the direct object or the indirect object, as expected. This is illustrated in (57): in (57a) the pronoun *me* is the indirect object, while in (57b) it is the direct object of the ditransitive verb.

- (57) a. Marie **m'**a donné un livre.  
Marie gave me a book.  
b. Marie **m'**a présenté à Paul.  
Marie presented me to Paul.

Given underspecification theory, one could argue alternatively that the pronouns *me* and *te* are unspecified for Case. However, this cannot be correct. Indeed, as shown above, the fact that strong pronouns lack Case features allows them



to be used in many contexts.<sup>26</sup> Thus, if the pronouns *me* and *te* were unspecified for Case, one would expect them to appear in the same environment as the strong pronouns *moi* and *toi*. Since this is not the case (see (58)), one must therefore conclude that these pronouns are specified for Case.<sup>27</sup>

- (58) a. Quant à moi/\***me**, je vais au restaurant.  
As for me, I' m going to the restaurant.
- b. C' est moi/~~me~~ qui ai vendu l'or dinateur.  
It' s me who sold the computer.
- c. Toi/\***te**, tu restes ici.  
Yóu stay here.
- d. Toi/\***te** voyager en première classe!  
Yóu, travel in first class!
- e. Toi et moi/\***te** et **me** sommes de bons amis.  
You and me are good friends.

Consider now first and second person plural pronouns, which show full syncretism in all grammatical functions. I claim that there are in fact three distinct morphemes for each grammatical function, as given in (47h,i,j). As can be seen, the pronouns in (47d) are exactly parallel to their singular counterparts: in both, indeed, there are a nominative form, a doubly specified form, and an unspecified form. This is the simplest claim, as I show in what follows.

In order to provide evidence that (47h-j) is correct, one can show that the following alternatives are wrong: (a) there is a single unspecified morpheme; (b)

---

<sup>26</sup> According to the same evaluator, " [this sentence] is misleading for the property of lacking Case features to which he refers has been proposed by him as a descriptive label to account for the distributional properties of strong pronouns. Therefore this cannot be adduced as *allowing* strong pronouns (sic) to be used in many contexts. There appears to be a confusion in ordering relations." Unfortunately, this is unfair, because I did not propose that. Rather, I attempted to show that strong pronouns are not specified for Case (see fn. 24).

<sup>27</sup> One could conceive a further alternative illustrated in (i), under which there are two distinct occurrences of *me* and *te* : *me*<sub>1</sub> / *te*<sub>1</sub> [A] and *me*<sub>2</sub> / *te*<sub>2</sub> [O]. The evidence against this view, which is related to the positive imperative verbs, where a strong pronoun appears instead of the expected weak form, as illustrated in (i), will be discussed in the next chapter.

(i) a. Regarde-moi/\***me**!  
b. Montre-le-moi/\***me**!

there are four distinct morphemes, three of which are Case-specified, while one is unspecified for Case, parallel to the third person singular pronouns. Consider the first alternative. If it were true, one would expect this pronoun always to behave like a strong pronoun. For instance, strong pronouns can never appear in preverbal position, as illustrated in (59). But this is not always true for *nous/vous*, which indeed may appear in preverbal position, parallel to other specified pronouns, as shown crucially in (60). One may therefore conclude that the (a)-alternative is not viable.

- (59) a. Marie *me*/\***moi** connaît.  
Marie knows me.  
b. Marie *les*/\***eux** a vus.  
Marie saw them.
- (60) a. Pierre **nous** a invités/ \*Pierre a invité **nous**.  
Pierre invited us.  
b. Est-ce Marie qui **vous** a donné ce livre? / \*Est-ce Marie qui a donné **vous** ce livre?  
OP-wh is it Marie who gave you this book?

As for the (b)-alternative, the formal syncretism does not permit us to provide direct evidence against it. Nevertheless, in clitic clusters these pronouns pattern exactly like their singular counterparts, as illustrated in (61). That is, they must obligatorily precede the accusative clitic. As we will see in the next chapter, clitic order depends upon the features of each clitic. Thus, if the plural and the singular pronouns exhibit the same order with respect to other clitics, one may conclude that they are specified for the same features, except of course number features.

- (61) a. Marie **me/nous le** montre.  
Marie shows it to me/us.  
b. Marie veut **te/vous le** donner.  
Marie wants to give it to you.

To summarize, it has been shown that weak personal pronouns, which are traditionally referred to as clitics, are crucially specified for various Case features. In

constrast, strong pronouns, which are not clitics, are shown to be unspecified for Case.<sup>28</sup>

#### 1.6.4 On the argument structure of verbs

As noted above, I assume that verbs are specified for Case. As a general rule, a tensed transitive verb may be specified for both nominative and accusative Cases, a ditransitive verb may be specified for three Cases: nominative, accusative and oblique, each associated to a thematic role. Thematic roles include the following: agent, benefactive (or malefactive), goal, locative, etc. (cf. Gruber, 1976; Jackendoff, 1972). I take the Case specification, the thematic roles, and the number of arguments, to be given in the argument structure of each verb. For instance, the argument structures of the verbs *dormir* 'to sleep', *aimer* 'to like' and *donner* 'to give' is given in (62), where X, Y, and Z represent the arguments.<sup>29</sup> Furthermore, as discussed below, the syntactic tree encodes the sequential nature of the computation. Internal arguments must be computed first, if there are any. Thus, in (62c) for instance, the object X is first computed, then the object Y and finally the external argument Z.<sup>30</sup>

- (62) a. [X [dormir]]  
       b. [Y [aimer X]]  
       c. [Z [[donner X] Y]]

It should be noted that the number of arguments that a verb takes cannot be predicted from its Case specification. Indeed, there is no reason to assume a one-to-one equivalence between the Case specification of a verb and its number of

---

<sup>28</sup> I am aware of an evaluator's objection, as to this conclusion 'is not justified' (see fn. 24 and 26).

<sup>29</sup> Notice that I do not use the type of notation currently encountered in the literature, for instance Verb (x,y), which requires some diacritic (for instance underlining) to distinguish the object from the subject (see for instance Di Sciullo, 1995). See Grimshaw (1991) for a survey of the notations used for the argument structure of verbs.

<sup>30</sup> Occasionally the terms *internal argument* and *external argument* are used descriptively to refer to the object and the subject respectively. But they do not carry here all the implications they have in GB theory. In particular the term external argument, in my view, is not intended to mean that this argument is outside the VP; in fact, I take the subject (or the external argument) to be outside the complement domain of the verb (see below).

arguments.<sup>31</sup> Infinitive verbs, for instance, are not specified for nominative Case (see below), but their argument structure must nevertheless be satisfied. Also, as I will argue in chapter 3, most transitive verbs (perhaps all) in French latently bear the oblique (benefactive) Case. Notice further that intransitive verbs are likely to be specified for accusative Case, even though they do not have an internal argument. As will be discussed extensively in this thesis, a criterion for an argument to cliticize onto a verb is that both must be specified for a common Case. Now, since intransitive verbs like *dormir* 'to sleep' and *manger* 'to eat' may take a pronominal clitic aside from their obligatory subject argument, as illustrated in (63), I conclude that they are specified for Case.<sup>32</sup>

- (63) a. Paul y dort.  
           Paul there sleeps.  
           ' Paul is sleeping there.'  
       b. Paul y mange.  
           Paul there eats.  
           ' Paul is eating there.'

Let us consider now the nature of the argument in order to provide some substance to the discussion of Case in section 1.6.1. In current analyses, it is assumed that the type of argument that a verb takes must be specified in its subcategorization frame. On this view some verbs may take as internal argument either an NP, a PP or a clause, to take the simplest case. For instance, for a verb like *donner* 'to give' it must be specified that one of the internal arguments is an NP and the other a PP (*à*-phrase), as given in (64). I would like to argue that such categorial information is not needed, limiting the discussion to internal arguments such as PP' s and NP' s. More

---

<sup>31</sup> An NP argument, as opposed to a clitic, need not be licensed by a verb with a Case. In French a nonargument clitic may cooccur with an NP argument (cf. *Paul y a rencontré Marie* 'Paul met Mary there'), in which case the NP is not assigned a Case (see below). Argument and nonargument clitics must be licensed by a verb with an available Case.

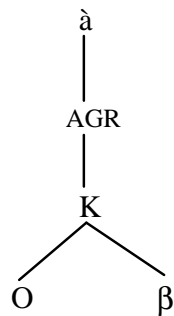
<sup>32</sup> Again, one of the evaluators rejects this conclusion. She writes: "these all refer to assumptions that the author makes rather than to demonstrations." At this point, I have to point out that this reasoning is current in theoretical linguistics. More generally in science, it is correct to draw conclusions from the premises we posited. However, whether the premises are correct or not is another story (see footnotes 24 and 26).

precisely, the categorial nature of the argument depends upon the Case-specification of the verb.

(64) [X [[donner NP] PP]]

Suppose that this verb bears the features [A], [O,β]. Given the Constraints on Merge, which are discussed below, each argument which is being computed must agree with the verb as much as possible with respect to the relevant feature (see (76) below). Assuming that the morpheme *à* has the feature structure shown in (65), the relevant argument of an [O,β]-verb must be first merged with this morpheme if it is unspecified for Case and thematic features so as to agree with the verb. Since accusative Case is not related to a thematic role, as suggested above, it will be computed as an NP. This is shown in (66).

(65)



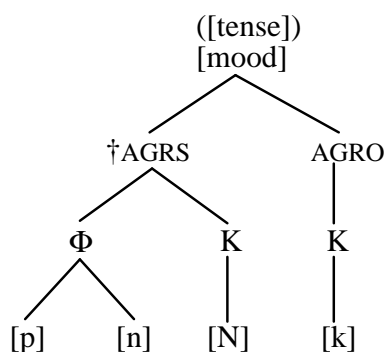
(66) [X [[donner NP] [*à* NP]]]  
 <A>;<O,β> <O,β>

This analysis predicts that if the argument is already specified for the same features as the verb, it may not be computed with the agreement marker. This prediction is indeed true: when the argument is a weak pronoun the preposition is not needed (67a), and indeed yields ill-formedness if present, as expected (67b).

- (67) a. Marie **me** donne **le livre**.  
 b. \*Marie [[donne **le livre**] **à me**].

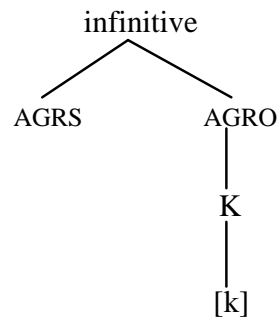
I further take verbs to be defined by the tree structure in (68), which is analogous to the one presented above for pronouns. The root node may contain other features, including tense and mood, the former being optional, and can be realized as infinitive (if tense is lacking), imperative, subjunctive, etc. Under the root node appear the AGR nodes, one for each argument, so that a three-argument verb will have three AGR nodes (the AGR' s are identified with grammatical relations subject (AGRS) and object (AGRO)). AGRS dominates a  $\Phi$  node and a K node; the former has the features person and number, while the latter contains the feature nominative. In French, AGRO dominates only a K node which may have either the feature accusative or oblique in most cases.

(68)



The presence of the dependent nodes of AGRS depends upon whether there is a tense feature in the root node. If there is no tense in the root node, AGRS will not expand to  $\Phi$  and K. This follows from the fact that in the normal case only tensed verbs can take an overt subject. Under this view, a (transitive) infinitive verb may be represented as in (69), where AGRS does not dominate any feature (and perhaps AGRS is missing). In order to express the correspondence between AGRS and tense in the feature tree, I will use the symbol '†' as a diacritic before AGRS. (This is equivalent to the arrow pointing from the root node to the major articulator in the Halle-Sagey model of feature geometry (see Sagey, 1986; Halle, 1992, and Kenstowicz, 1994, for discussion).)

(69)



There is empirical evidence supporting the tree structure in (68). When a verb is tensed, it must have a subject with which it agrees in person and number. If the subject is a pronoun, it must be specified for nominative Case; it can never be specified for another Case, nor can it be a strong pronoun unspecified for Case (70a,b). (This outcome is in fact forced both by the Constraints on Merge (see (76)) and the Blocking Principle (85), which are discussed below.) Unlike tensed verbs, infinitive verbs are incompatible with an overt subject (70c,d) (at least in root clauses). Nevertheless, in some marked contexts (with a special intonation contour), an overt subject is allowed (71). Interestingly, the overt subject may be either a strong pronoun or an NP, but it can never be a Case-specified pronoun, nominative or other (I will return to this point below). Given this, one can conclude that a tensed verb, but not an infinitive verb, is specified for nominative Case.

- (70)
- a. \***Moi** connais Marie.
  - b. **Je** connais Marie.  
I know Marie.
  - c. \***Il** manger une pomme.  
He to eat an apple.
  - d. \***Marie** manger une pomme.  
Marie to eat an apple.

- (71)
- a. Moi/Marie voyager en première classe!
  - b. \***Je** voyager en première classe!  
Me, travel in first class.

On the other hand, transitive verbs (finite or not) can license an internal argument, be it a clitic or not, as illustrated in (72) with an infinitive and an imperative verb. They may not agree in  $\Phi$ -features (person and number) with the internal argument.

- (72) a. Paul voudrait **la** rencontrer. / Paul voudrait rencontrer **Marie**.  
           Paul would-like to meet her / Marie.  
       b. Tenez **le livre**! / tenez-**le**!  
           Hold the book! / hold it!

Summing up, the lexical structure specifies the Case, the thematic feature and the number of arguments that a verb takes. I have established a relationship between the nature of the complement (*à* NP or bare NP) and the features of the verb under a general constraint on agreement at the computational stage, which will be discussed below.<sup>33</sup> I have proposed, furthermore, that verb features are organized in a hierarchical tree structure analogous to the one presented for pronouns.

### 1.6.5 Properties of the Computational Space

The Computational Space (cf. Fig. 1.2), as indicated by its name, is analogous to Chomsky's Computational System, but does not include the concept of numeration (see Di Sciullo, 1996, etc.). I will first take a fresh look at this module, attempting to reduce it to its simplest expression (Merge) along the lines of the bare phrase structure theory, and then I will suggest a general constraint in order to limit its power (the Constraints on Merge).

#### 1.6.5.1 Merge

Let us assume that the computational space builds a bare syntactic structure in the sense of Chomsky (1994). That is, syntactic structure is constructed by recursive application of the operation *Merge*. The latter consists of attaching together two lexical elements, for instance A and B, an operation which forms a unit dominated by

---

<sup>33</sup> Similarly, it will be shown in chapter 3 that some verbs take a *de* complement which corresponds to the clitic *en*, and some others take an *à* complement corresponding to the clitic *y*.



a projection of either A or B: [A A B] or [B A B] (the projected element is the head and also the label). A structure constructed in this way can be further expanded by merging with another element drawn from the lexicon, or with a unit already constructed in the same way. For instance, the structure [A A B] may be merged with either a lexical element C or with a structure [E D E], yielding respectively the structures shown in (73).<sup>34</sup>

- (73) a. [C/A C [A A B]]  
 b. [A/D [A A B] [D D E]]

Chomsky further defines the operation *Move*. It is a singular application of *Merge*, i.e. a structure is merged not with an independent unit already constructed or a new element drawn from the lexicon, but with one element drawn from within itself. By virtue of the extension requirement, such an element (obligatorily) targets the maximal projection of the structure which is being constructed. The targeted node projects and the moved element is attached as a daughter to this new projection. To take an example, suppose that the structure in (74a) is being constructed by merging. The object of the verb must check its features with that of AGRO. The constituent AGRO' is targeted and a new category AGROP is projected from it. Then the object of the verb is raised to Spec-AGROP, yielding the structure in (74b). After this singular transformation, the computation may continue by further applications of *Merge* and *Move*; in particular (74b) may be merged with Tense, or some other functional heads, and the subject may raise to AGRS.

- (74) a. [AGRO' AGRO [VP Mary [V' likes beer]]].  
 b. [AGROP beer [AGRO' AGRO [VP Mary [V' likes t]]].

It appears that the singular transformation as well as the various functional projections complicate Chomsky's system. A natural way to avoid this complication is to assume the simplest hypothesis, that syntactic structure is constructed at one go in the computational space in order to satisfy the syntactic component, which requires that lexical elements be assembled together, forming larger units under conditions

---

<sup>34</sup> See Chomsky (1998) for further development of *Merge*.

that will be discussed below. In my view the phonological component modifies the structure only once it is completely built up. That is, lexical elements forming the phrase marker may move from their initial position, or their feature may be manipulated (delinking process). It is important to note that this does not entail an explicit ordering of the components, a feature which is banned anyway by Di Sciullo's (1996) MCS; indeed, at each step of the derivation, the structure which is being built deals with the relevant level. That is, by the stage where words are merged together, the phonology deals with smaller elements, namely the segments forming the words. Once the syntactic structure is assembled, the phonology comes to deal with a string of words.

As mentioned above, I further claim that no functional head is merged with lexical elements. In other words, I assume a strong lexicalist hypothesis, where derivational and inflectional morphology occurs strictly in the morphology component (see among others Di Sciullo and Williams, 1987). It is thus worth saying that the computational space builds a syntactic structure with a single layer, namely the VP, which comes to be its maximal projection. While I will refer sometimes to the maximal projection of the syntactic structure as **S** (for sentence), this is not to be understood as a category, as in the Standard Theory. Rather it is the highest projection of the verb (VP).

A further consequence is that the number of projections of the verb may not be limited to a fixed value, as it is the case in the GB framework. Indeed, this will depend upon the number of times that a head is merged. Thus, for a ditransitive verb with two arguments the maximal projection of the structure will be  $V'''$  (=S). Given that only binary nodes are allowed, the computation of an adverb will increase the number of projections of the verb. This is illustrated in (75) (details set aside at this point). Further consequences of this claim will be pointed out when relevant to the discussion.<sup>35,36</sup>

---

<sup>35</sup> Note that I assume that the verb always projects when it is merged with another element; such an assumption is not crucial here. However, when a determiner is merged with an N, it is unclear whether the projected element is the determiner or the N. As this question is irrelevant for the analysis to be developed, I leave it open. More generally, in the thesis, no transformation will

- (75) a. [<sub>V'''</sub> Paul [<sub>V''</sub> [<sub>V'</sub> offre [des fleurs]] à Marie]]].  
           Paul   offers DET flowers   to Mary.  
       b. [<sub>V''''</sub> Paul [<sub>V'''</sub> [<sub>V''</sub> [<sub>V'</sub> offre souvent] des fleurs] à Marie]]].  
           Paul                   offers often    DET flowers to Mary.  
           ' Paul often offers flowers to Mary.'  
           (where in (a)  $V''' = VP=S=offre$ , and in (b)  $V'''' = VP=S=offre$ )

Moreover in GB/MP the operation Move allows for making chains. A chain consists of a displaced element and the trace it leaves behind. Traditionally the trace is required in order for some relations to be interpreted locally. For instance, when an element moves to an A-bar position, it must be ensured that it is still assigned a theta role by the verb. Thus, theta-marking a trace, which is a member of a chain, is equivalent to the theta-marking of the antecedent of the trace. In the context of the present proposal, the notion of trace is not necessary. Indeed, movement is generally a repair strategy for the OCP: an element may move to another position. If it bears a Case, it need not be assigned Case by the verb; if it is unspecified for Case, given the nonlinear representation, it can still be assigned Case by spreading. Therefore, I assume the null hypothesis according to which there is no trace in syntactic structure. (This obviously does not entail that there is no empty lexical element.)

To conclude, I assume a modular theory of grammar, specifically that proposed in Di Sciullo (1996), as depicted above in Figure 1.2. The components are autonomous but interact in the computational space to create an optimal derivation by conceptual necessity. I assume that the computational space generates bare phrase

---

refer to the label of the constituent of the syntactic tree, and for simplicity I will therefore omit the labels of the nodes in the bracketing structures.

<sup>36</sup> In GB theory, where the phrase-marker preexists independently of lexical items, initial word order is provided by the phrase-maker (X-bar), which is governed by the LCA (Kayne's 1994). The dynamic conception of the structure adopted here leaves us without an account of linear order. While this issue cannot be discussed here in detail, the following remark is in order. Since the verb (V) is merged with its object (O) before its subject (S), the computational component may yield four types of orderings for a three-word structure: either the V precedes or follows the object, i.e. [V>O] or [O>V], and either this complex follows or precedes the subject, as given in (i). Thus other attested types of linear order must be derived from these four basic types by syntactic process. For a given language, the choice of either of these possibilities might depend upon the lexical and feature inventories. I assume that in French-like languages the initial order is always SVO.

(i) a. [S [V O]]      b. [[V O] S]      c. [S [O V]]      d. [[O V] S]

structures, including strictly syntactic objects, as opposed to abstract functional heads. One notes that in Chomsky's theory, lexical elements freely selected from the lexicon are merged, and it is assumed that the derivation will crash if a wrong choice of lexical items is made. In the next subsection I will attempt to capture the conditions under which Merge takes place, focussing on clitic selection in French.

### 1.6.5.2 Constraints on Merge

Consider the first person singular pronouns in French: *je*, *me*, and *moi*. As discussed above, *je* can only be used as subject, while *me* can be used either as direct object or indirect object of a verb. As for *moi*, it may be used in stressed contexts, and as the object of a preposition. I have described this fact in terms of features: *je* is specified for [N], *me* for both [O], [A], whereas *moi* is unspecified for Case (see (47a-c)). Now, during the building of a syntactic structure containing such elements, one must ensure that the desired element is merged with the verb. That is, one must ensure that *je* will never be computed instead of *me* as object of a verb, nor *moi* in a non-stressed context. In current theories, there is no principle governing the selection of a lexical category; indeed, in MP and some works in GB, it is simply assumed that the derivation will crash if a wrong choice is made (cf. Bouchard, 1984, 1995; Chomsky, 1992, 1994, 1995).

In fact, in GB theory, where the phrase-marker is predetermined and contains explicit grammatical positions, one may say with more or less accuracy that the subject pronoun must be inserted in subject position in the structure and the object pronoun in object position, whereas the strong pronoun might bear some indication in the lexicon, ensuring its use in a stressed context only. In a dynamic conception of the phrase-marker, more can (and must) be said about this matter. Indeed, as there is no predetermined structural position, it is no longer possible to say simply that the object clitic must be inserted in the relevant position of the tree-structure.

I will argue that when two elements are to be merged, the minimal requirement is that they be compatible. Obviously, since each lexical element in a given paradigm bears inherent features that make it distinct from others, it must be

the case that compatibility between two elements (to be merged) depends upon whether their features are compatible in the sense defined below.<sup>37</sup>

On purely conceptual grounds, compatibility may be defined as follows. An element X can be said to be compatible with an element Y with respect to a feature F if both are specified for F. If, instead, X is specified for F and Y is specified for another feature T, they clash with each other. Conversely, if both X and Y are underspecified, they may not clash, and they are compatible in some way. However, if X is specified for F and Y is underspecified, there is no feature clash. Logically, one must expect that the computation of items drawn from a particular paradigm be governed by the same reasoning. Specifically, I propose that merging is governed by the following constraints, which I refer to as the Constraints on Merge (COM; see also Desouvrey, 1995/1996).

(76) Constraints on Merge (COM)

A head X and an element Y can be merged if either (a), (b), or (c) holds:

- a. X and Y bear the same features.
- b. X bears some features, but Y does not bear any feature.
- c. Neither X nor Y bear any feature.

Merger is impossible if:

- d. X and Y each bear different features.

Condition (a) is the normal (and preferable) case in the paradigm, and I will refer to it as *strong agreement*. By virtue of (76a), if the complement Y does not have the feature of the head, it must merge first with some agreement marker, usually a preposition, whenever such an element is available (recall the discussion under § 1.6.4 above). When two elements are merged according to condition (b), they will be said to agree weakly, and therefore I refer to this as *weak agreement*. Condition (c) entails the negation of (a) and (b), as the elements to be merged are feature-free; I thus refer to this case as *default agreement*, since they are compatible through the

---

<sup>37</sup> The notion of feature compatibility is used in unification-based approaches to grammar (see, for instance, Shieber, 1986), in which features are defined as having attributes and values (cf. fn. 22).

absence of features. As for condition (d), which yields a feature conflict, it will never appear in any structure; the grammar will use a strategy to avoid it, as I will show (cf. the discussion of clitic doubling in chapter 4).

Furthermore, (76) is consistent with Di Sciullo's (1998/to appear) Local Asymmetry Hypothesis, an instantiation of which is given in (77) (her (10a)). Indeed, conditions (a), (b) and (c) entail that the head X includes the features, if any (in my view), of the complement Y. We will see later that (76) involves the notion of locality.

- (77)  $R_{\text{compl}}$  (Relation Complement)  
A head properly includes the features of its complement.

Another case, not included in (76), arises when the verb does not bear any feature, while the complement bears some feature. This situation is the reverse of (76b); call it (76b') for ease of reference. (76b'), as is (76d), is ruled out by virtue of (77), since the head does not include the feature of its complement. Moreover, (76b') may be interpreted as a case of feature conflict, parallel to (76d), once one assumes that the absence of a feature on a selecting head (a verb) is equivalent to a null feature [ $\emptyset$ ] (see the discussion of imperative verbs in chapter 4).

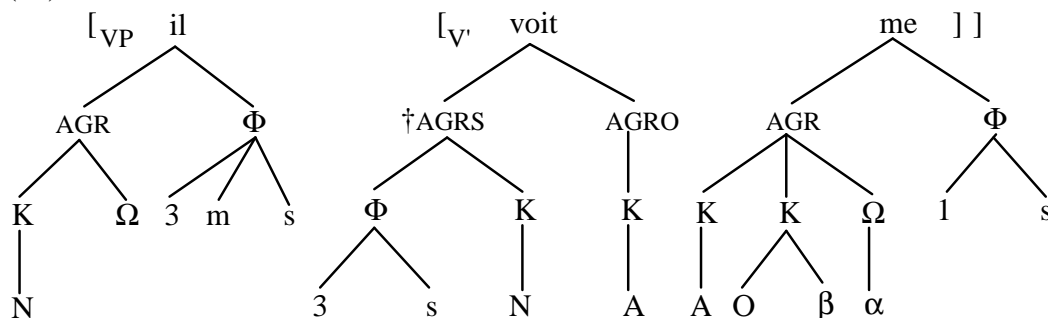
I turn now to evidence to the effect that compatibility of features between two elements which are being merged is consistent with either (76a), (76b) or (76c). Keeping in mind the fact that in French verbs agree in  $\Phi$ -features (person and number) with their subjects, consider the derivation of the simple sentence in (78).

- (78) Il me voit.  
He sees me.

The inflected verb *voit* 'sees' is drawn from the lexicon, as well as the pronoun *me*. Since both are specified for accusative Case, merger is possible by virtue of (76a), yielding the V' (79). It is true that both are specified for an additional Case, for instance nominative for the verb, and oblique for the argument. But given the Local Asymmetrical Hypothesis, and the generally admitted view that the internal argument is more deeply embedded than the subject (cf. Grimshaw, 1991; Di Sciullo, 1995;

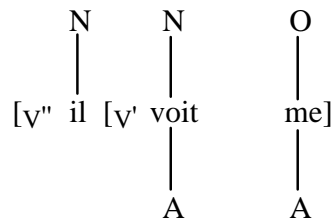
etc.), these extra Cases are irrelevant at this step of the derivation. Although the clitic *me* is also specified for  $\Phi$ -features, there is no clash since the verb does not have this node under the relevant AGR. The computation continues; the subject is drawn from the lexicon, and is merged with the already constructed  $V'$ , yielding the VP (=S). This merger is well-formed, since the verb and the subject are both specified for nominative Case, accusative Case being irrelevant at this step. The pronoun *il* and the head agree also in  $\Phi$ -features: the verb has the following  $\Phi$ -features: third person, singular, as does the pronoun. Although the latter contains a gender feature (masculine), it is compatible with the verb, which does not have gender morphology, by virtue of (76b) (weak agreement). More generally, for each node of the tree structure, there is compatibility between the verb and its arguments, in the sense of the COM. (The next step of the derivation, where movement of the object clitic takes place, will be discussed in chapter 2.)

(79)



Notice that the representation in (79) will not be repeated for each example. In fact I will adopt a more simplified version where class nodes,  $\Phi$ -features, as well as extra features of one of the interacting elements (for example animacy) are set aside. A simplified representation of (79) is given in (80).

(80)



It is clear that all constituents of the structure contain only compatible elements, in the sense of the COM. If (76) were not included in the grammar, one would expect to find a plural pronoun as subject of a singular verb, a second person pronoun as subject of a third person verb, and a nominative pronoun as object of a verb. As another piece of evidence for the COM, consider the ill-formed sentences in (81). One can show that each of these sentences is ruled out by virtue of (76d). In (81a), the verb is third person singular, while the subject is third person plural; in (81b), the subject is second person while the verb is third person singular; finally in (81c) the object of the verb is a nominative pronoun. Here again, if the computational space were allowed to freely merge two elements drawn from the lexicon, these sentences would be grammatical. Since they are not, the COM is supported.

- (81)    a.    \*The dogs sleeps.  
           b.    \*You sleeps.  
           c.    \*They saw he.

Consider another piece of evidence involving a previous morphological treatment of the words to be merged. The morphology component may add agreement affixes to stems, forming grammatical categories fully inflected for  $\Phi$ -features (or inflectional features; see also Ralli (1997)). But the combination of the grammatical categories thus formed in the syntax is ensured by the COM. For example, in the French phrase *une actrice charmante* 'a charming actress', each element may have been constructed in the morphology by adding feminine suffixes to two masculine stems: *acteur* 'actor' and *charmant* 'charming', but the merging in the syntax of the words thus created, which yields [actrice charmante], obeys the condition (a) of (85) (strong agreement). This structure is further merged with the indefinite determiner *une* [Feminine], which is previously created by adding a feminine suffix to the stem *un*, under the same condition. In other words, if merger were free, grammatical objects would not have to agree with each other in the sense just described. One would further expect to find structures such as *\*un charmant actrice* or *\*un charmante acteur*.

In sum, I have shown that feature clash does not occur during the building of the structure: a verb cannot be merged with an element if it does not include its



feature at the relevant step of the computation of its argument structure. That is, given that the verb must first merge with its object, the latter must be in conformity with (85) as regards the features relevant at that step, namely Case and thematic features. In the computation of the subject, Case-features and  $\Phi$ -features are relevant and they must be consistent with (85).

Consider again the conditions (76b) and (76c), which I refer to as *weak agreement* and *default agreement* respectively. Default agreement will arise only when both the verb and the argument are unspecified for Case. Thus, this case occurs with infinitive verbs, which are not specified for nominative Case, as discussed above. Infinitive verbs may take as subject either the abstract element PRO, which I take to be unspecified for Case (see also Bouchard, 1984; Koster, 1984; Manzini, 1983; etc.), a strong pronoun, or a nominal, but never a Case-specified pronoun (clitic), as illustrated in (82).<sup>38,39</sup> Notice that in root sentences such as (82a,b) exclamative intonation is absolutely required, as mentioned earlier, while this is not the case in embedded contexts (82c). This is due to the fact that, normally, sentences must be tensed. Untensed root sentences are less frequent in French (and comparable languages), and indeed, such sentences convey some marginal meaning with a special intonation contour, which is indicated by the exclamation mark. Embedded infinitive verbs do not need a special stress pattern presumably because they are usually dependent on a tensed matrix verb.

- (82) a.   Moi voyager en première classe!  
           Me, travel in first class.  
       b.   Paul épouser Marie!  
           Paul, to marry Marie.  
       c.   Paul aime PRO nager.

---

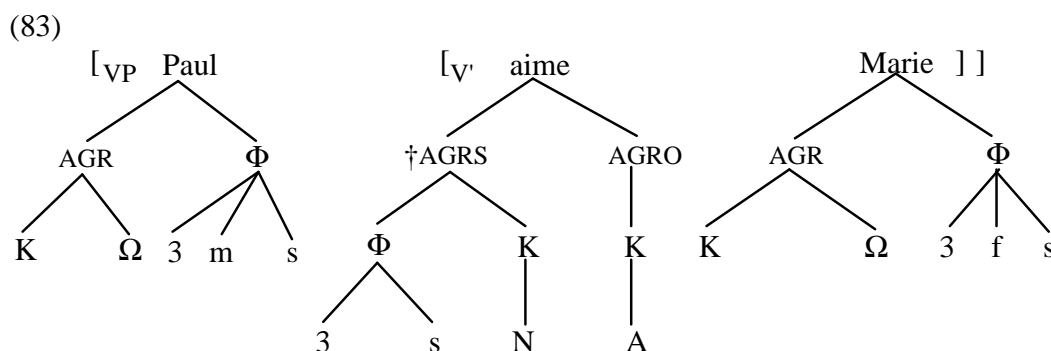
<sup>38</sup> In GB whether PRO lacks Case or not is controversial. To avoid circularity, I will not attempt to provide empirical evidence for my assumption that it is Caseless. However, it will become clear, in the light of my treatment of cliticization, that PRO cannot be specified for Case.

<sup>39</sup> Exceptional Case marking constructions in English may also illustrate the issue at hand. In such constructions, the infinitive verb has no nominative Case, as in French, and its subject may not be specified for Case, as shown in (i). I argue in a later chapter that in English, object pronouns are unspecified for Case.

(i)       John believes [Paul/him/\*he to like Mary]

Paul likes swimming.

As for weak agreement (76b), one may note that it is the normal case in the noun paradigm, which contains no Case-specified element. That is, the force of the constraints must be relativized with respect to the paradigm (see below). As exemplified in (83) by the derivation of the sentence *Paul aime Marie* 'Paul likes Marie', both the object and the subject yield a weak agreement configuration, but the sentence does not necessarily convey an exclamatory nuance.<sup>40</sup>



Clearly the personal pronoun paradigm includes elements which are specified for Case and others which are not, as discussed above. However, it is not the case that the Caseless pronoun may be used freely with Case-specified verbs, even though this does not violate the COM (weak agreement). Indeed, if such a pronoun is used instead of a clitic (Case-specified pronoun in my viewpoint), an ill-formed structure obtains, as exemplified in (84). Kayne (1975) rules out this outcome by positing that the rule of clitic placement is obligatory. We would like to find a more general principle to exclude such structures. A natural principle which comes to mind is the Blocking Principle (BP) (see Aronoff, 1976; Lumsden, 1987, 1992). Expanding from Lumsden (1992: 480), I state this principle as in (85).

- (84) a. \*Marie connaît toi.  
Marie knows you.

<sup>40</sup> The  $\Omega$  node of the NPs is empty, but they may contain some features including animacy, specificity, etc. These features are not crucial in French, and indeed supporting evidence is lacking. However, these feature play an important role in Spanish, as I will show in chapter 4.

- b. \*Toi verras moi demain.  
You will see me tomorrow.

(85) Blocking Principle (BP)

A form which is specified for a relevant feature must be computed before a form which is unspecified.

The BP correctly rules out structures such as (84) because there exists in the paradigm a better form, i.e. a form which is specified for the same Case as the verb, specifically a clitic. However, we may note that in some cases (76b) (weak agreement) is allowed, in violation of the BP, as illustrated in (86a,b). The sentences are acceptable because they contain a nuance of meaning (indicated by the exclamation mark),<sup>41</sup> which is impossible to obtain with the clitic form (which obeys (76a)); the same is the case for (86c) although its meaning is somewhat more lexical: in effect, the expression *ne...que* cannot be constructed with a clitic at all (cf. \**Marie ne le voit que* ).

- |      |    |                        |                       |
|------|----|------------------------|-----------------------|
| (86) | a. | Tu as vu MOI!          | ' You saw ME'         |
|      | b. | LUI connaît le chemin. | ' HE knows the road'  |
|      | c. | Marie ne voit que LUI. | ' Mary sees only HIM' |

It is convenient to interpret the facts in (86) as follows: the BP is violated and is repaired by emphatic stress, which falls precisely on the offending element. If this view is correct, the nuance of meaning conveyed by these structures can be directly attributed to the violation of the BP. We will see later further cases where a violation of a principle is accompanied by emphatic stress.

Summarizing, my objective in this section was to show that lexical elements, which are drawn from the lexicon, are not merged freely; instead this operation is governed by the Constraints on Merge. As the latter express a natural relation of compatibility between two objects, it must be the case that they are universal, and

---

<sup>41</sup> Put another way, it may be that the unmarked form is used in focus, which requires a pitch accent (see Selkirk, 1995). I thank the evaluator referred to above for bringing the notion of pitch accent to my attention. However, she does not necessarily agree with my interpretation of this notion.

therefore no variation is expected cross-linguistically. From this point of view, languages are predicted to vary only according to their lexical inventory and their feature specifications.

## 1.7 Summary

In this chapter, I have attempted to show that there are no clitics per se. In fact, the elements referred to as clitics exhibit a set of properties which can be found to varying degrees in many grammatical categories, including verbs and NPs. Many researchers have already made similar claims (cf. Everett, 1996; Zwicky, 1994; Halpern, 1995; etc.), but in most theoretically oriented works the phenomena characterizing cliticization are taken to be axioms from which the analysis proceeds, the analysis being reduced to the determination of the manner of representation in the syntactic tree.

Indeed, I have discussed the basic tenets of the Principles and Parameters Theory, parameters and structural constraints, showing that the former may not be taken as satisfactory explanations, and that the latter are quite arbitrary in the sense that they are not the only logical possibility allowed by the representation. For instance, I have shown that the constraints of leftward movement only and left adjunction only, although restrictive, are highly arbitrary in that they do not follow from the logic of the representation which is being used.

From this perspective, I have presented an overview of nonlinear phonology, borrowing its structures and governing principles, and applying them to the representation of syntactic features. Nonlinear phonology exhibits some hallmarks of a good theory in that its axioms are logically constrained by the representation. These are: the basic idea that features are autonomous and are located on distinct tiers, (b) a universal association convention, and (c) a well-formedness condition. Once these are adopted, they logically entail operations such as spreading of features, multiple linking, floating features, rightward or leftward association, etc. It would be arbitrary to postulate autonomous tiers for features and to absolutely reject multiple linking or rightward association. The theory would be also considerably weaker if one were to allow parametric variation of the initial stipulations, namely that some languages

should not allow an autosegmental treatment of their phonological features, or that the prohibition on crossing of association lines does not hold for all languages. However, it allows linguistic variation: for instance, association may be leftward or rightward, and multiple linking may not occur in some languages.

The type of reasoning that underlies nonlinear phonology, if applied to syntactic representations, will have the effect of constraining the theory by rendering it clearly falsifiable and less arbitrary. Indeed I make the strong claim that syntactic representations are much more similar to phonological representation than it has been generally supposed. With others (e.g., Bonet, 1991; Yip et al. 1987), I suggest that syntactic features can (and should) be best analysed similarly to their phonological counterparts. As such, they appear on different tiers, and are presumably subject to well-known autosegmental operations, including spreading, delinking, etc.

My claim that syntactic features can be treated in a manner similar to phonological features is framed in the model of grammar presented in Di Sciullo (1996). I also present some key concepts of the Minimalism Program (Chomsky, 1995) which I will use. Specifically, I assume the bare phrase structure theory with the proviso of a strong lexicalist hypothesis, according to which inflectional morphology is dealt with in a separate morphological component. I provide evidence that personal pronouns are specified for various features, including Case and phi-features. While many other authors assume that pronouns do indeed have such features, they do not give any further details on the nature of this specification, that is, whether a pronoun is specified for a single feature or for two features. I further argue that the computation of lexical elements is not free, as it is in the Chomskyan model. I show that it obeys certain strict principles, namely the Constraints on Merge, which are conceptually and empirically grounded. In the next chapter I will consider in some detail the syntax of personal pronouns in French, showing that the fact that they are special clitics follows from their Case features.

## **CHAPTER II**

### **AUTOSEGMENTAL CONSTRAINTS ON SYNTACTIC STRUCTURES**

#### **2.1 Introduction**

In this chapter, I will attempt to provide an answer to the following questions, focussing on French: (a) why the strong form of the first or second person pronoun is used instead of the weak form in positive imperative verbs; (b) why weak pronouns are enclitic in positive imperative sentences and proclitic in negated imperatives; (c) why weak pronouns are proclitic with all other verbs; (d) why the accusative clitic follows the dative clitic, except in the third person; and (e) why a weak pronoun cannot be separated from the verb that selects it by any intervening element other than another clitic.

As discussed in the previous chapter, most traditional analyses take clitics to be a genuine category, with their own defining properties. Under these analyses, Romance languages are assumed to include such a category in their pronoun system, but not English, for instance. This chapter provides evidence to the effect that most personal pronouns in French are clitics because they are specified for Case, consistent with the Condition on Cliticization formulated in the preceding chapter. I will argue that Case is the crucial feature involved in cliticization, and a Caseless argument may not be a special clitic. Assuming that clitics are generated in the same syntactic position as any other arguments, along the lines of Kayne (1975) and many others, I will show that the preverbal position of object clitics on the surface is a consequence of the Obligatory Contour Principle (OCP), which forces any Case-specified element to move outside some syntactic domain, and forces the relevant feature of the clitic to fuse with that of the verb, both under conditions that will be discussed below.

The following section discusses an analysis of feature checking, which is an enforcement of the Constraints on Merge (see the previous chapter) at a further stage of the derivation. By the time that the structure is being constructed, strong agreement requires that both elements under merge have a feature in common, while other irrelevant features of both are ignored (and may clash). After the computation, the structure comes to deal with the phonological component, which, as I argue, will eliminate irrelevant features in the structure if they clash. The strong form of the pronouns which appears in the imperative is due to the delinking of irrelevant features up to structure preservation. Section 3 discusses the effects of the OCP on syntactic features. It is argued that the OCP has two complementary effects on syntactic structures: on the one hand, it forces the clitic to move and to adjoin to the first element to the left of the verb, or the auxiliary+verb complex; on the other, it triggers fusion of Case occurrences in a given tier. It is shown that the fact that clitics may not be separated from the verb is a consequence of the fusion process. A new structure for adjunction is proposed, where the adjunct and the head comes to be on a distinct timing tier. In section 4 the ordering of transformations and cooccurrence restrictions on clitics are discussed. The ordering of transformations is not independently formulated, allowing the order of the clitics to be derived from the application of various constraints including the Animacy Hierarchy and the Clitic Licensing Constraint. The so-called *me-lui*/I-II Constraints will be shown to follow from these two constraints. Section 5 concludes the chapter. Some facts related to auxiliaries are considered in the appendix.

## 2.2 Feature checking

In the Minimalist Program (MP) syntactic elements are drawn from the lexicon fully inflected with all relevant features, including  $\Phi$ -features and Case features. The features are checked against the relevant functional elements; that is, it is verified whether each lexical element really bears the correct features required by the functional element. If there is a correspondence between the features of the lexical element and those of the functional head, the derivation converges, otherwise it crashes (Chomsky, 1995).

On this view, there is no difference between a nominal like *Marie* in the sentence *Paul aime Marie* 'P. likes M.' and the pronoun *la*; both are inflected in the lexicon with features, e.g., accusative, feminine, singular, third person. Although appealing, such an approach might run into the following descriptive problem: it fails to capture the fact that *la*, unlike *Marie*, is always an object pronoun and hence its use as subject yields ill-formedness, as discussed in the previous chapter. Moreover, under this approach, it is predicted that these elements, because they are inflected for the same features, should have the same syntax. However, this prediction is not verified.

As discussed in the previous chapter, most personal pronouns are inherently specified for Case features, while nominals are not. With this in mind, I will defend a different view of the feature-checking process (see also Brody, 1997, for another view). Specifically, I will argue that feature checking is a constraint of the phonological component, which is given in (1).

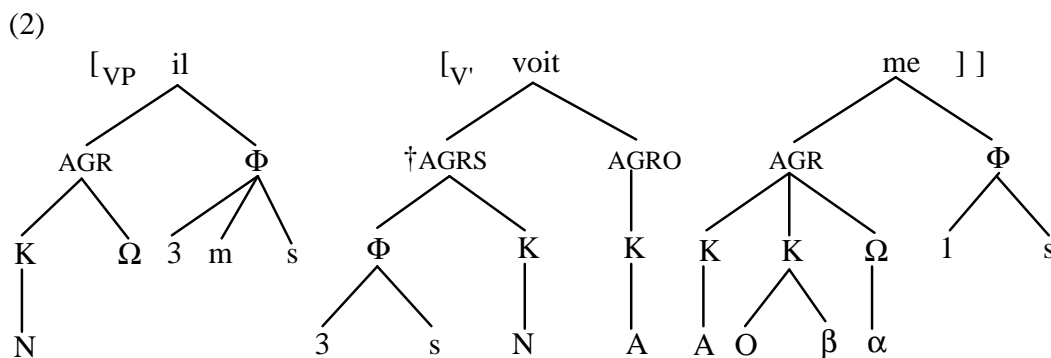
- (1) Feature checking requirement  
Given two merged elements A and B, all features of A must agree with all features of B in the sense of the Constraint on Merge.

Given that the COM governs the building of the structure in the computational space, how can any feature mismatches appear in the structure thus constructed? In fact, in the computational space, structure is built sequentially so that the internal argument is computed before the external argument. As seen in the previous chapter, at each merger only the relevant feature is considered, given the Local Asymmetrical Hypothesis (Di Sciullo, 1998/to appear). For instance, for a verb specified for both nominative and accusative Cases, only the accusative Case of the verb is relevant during the computation of the internal argument; the latter may be specified for more than one Case, but only its accusative Case is relevant. Thus, after the computation of all arguments, the remaining features of the verb and the argument become relevant and may eventually clash with each other.

Now, once the structure-building phase is over, all the features of the arguments must match with all the features of the verb in the sense of the Constraints on Merge



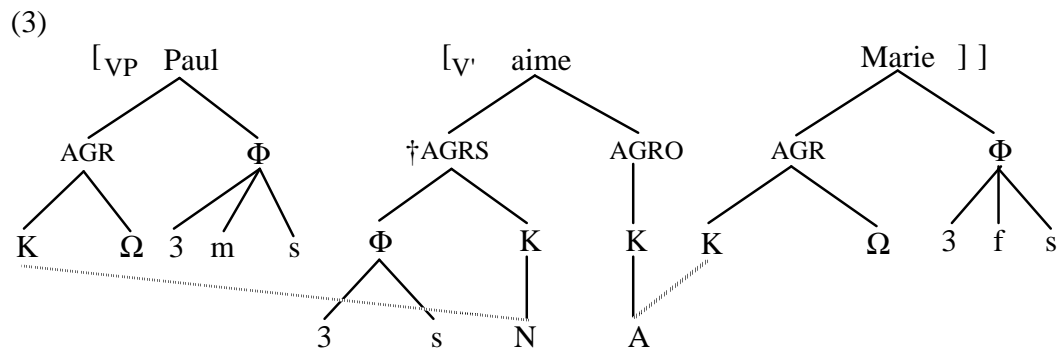
(COM). To see this, consider again the initial structure of the sentence *Il me voit*, see (2), which I discussed in the preceding chapter. One can see that both the arguments and the verb are specified for nominative and accusative Cases; the subject is specified for person and number, as is the verb. Other features of the arguments are not matched with a feature of the verb, but nevertheless, this is consistent with the COM: the object is specified oblique-benefactive  $[O, \beta]$ , while the verb is not specified for these features; the object is also specified for person  $[1]$  and number  $[s]$ , as well as for animacy  $[\alpha]$ , but the verb is not. The structure in (2) is therefore perfect at this step of the derivation.



However, if the checking is not consistent with the COM, the structure must be repaired, as I will show. This occurs when the verb and the argument are each specified for an extra feature and the extra features clash with each other. Before providing evidence for this mismatch, something must be said about the treatment of the features by the phonological component.

First of all, when the verb and the argument are specified for identical features, I assume that those features are assigned a common tier. Thus, in (2) the accusative Case of the verb and the accusative Case of the argument constitute a tier, and both appear in a single plane, which does not include the other features. The same is true for nominative Case. Further, Case-specified arguments are not assigned Case by the verb; nevertheless the Cases of the latter are saturated or absorbed, i.e. they cannot spread onto an unspecified argument. The extra features of the argument, namely oblique, benefactive, animacy, and gender are ignored because they do not interact with the verb, which is not specified for such features.

If an argument is not specified for Case, the verb Case-marks it by spreading, as illustrated in (3). In accordance with the Principle of Spreading (see the preceding chapter), the nominative and the accusative Case spread onto the NPs *Paul* and *Marie* respectively.<sup>1</sup> This process saturates the Case of the verb, parallel to the case where the argument is Case-specified (see more on that matter below). Now, since the Cases of the verb are related to grammatical roles, the assumption about the directionality of the spreading process, namely that nominative Case spreads leftward and accusative Case rightward (see (41) in chapter 1), is not needed. One may rather assume that the accusative Case will spread onto the intended direct object and similarly the nominative Case will spread onto the intended subject.



Let us return now to the Checking Constraints. In the previous structures, there is no feature mismatch between the verb and the arguments, whether they are Case-specified (2) or not (3). Now consider imperative verbs. In French, imperative verbs do not have any special morphology. Indeed, the three imperative forms of any verb are similar to some other tense, most often the present tense; compare the imperative conjugation in (4) with the present indicative in (5) (where the forms used in imperative verbs are emphasized). Given this syncretism (and absence of counterevidence), I take imperative verbs in French to be specified for the same features as all other tensed verbs.

- (4)     Chante!     ' sing!'     (2 singular)  
           Chantons!   ' Let' s sing!' (1 plural)

<sup>1</sup> The featureless NPs are shown with bare class nodes. However, there is no evidence that such nodes are really present. This issue is not crucial here.

Chantez! ' sing!' (2 plural)

- (5) Je *chante*. ' I sing.'  
 Tu *chantes*. 'you sing.'  
 Il *chante*. 'h e sing.'  
 Nous *chantons*. ' we sing.'  
 Vous *chantez*. 'you sing.'  
 Ils chantent. ' they sing.'

Thus a transitive imperative verb must be specified for both nominative and accusative Cases. Now, a characteristic of imperative verbs (in Romance) is that they generally lack a subject.<sup>2</sup> One could say either that imperative verbs have a non-overt subject or that they have no subject at all. If there is a non-overt subject, it must be either PRO or *pro*. Assuming that PRO is restricted to infinitive verbs, one remains with the possibility that the null subject in imperative verbs is small *pro*. But since French is not a *pro*-drop language, this empty pronominal is not expected to be present in the inventory of pronouns. Thus, one may conclude that imperative verbs are subjectless. (An additional piece of evidence comes from the lack of proclitics in positive imperative verbs; see below.)

Now consider the facts in (6). A first (or second) person weak pronoun cannot be the object of an imperative verb, while a third person weak pronoun can. To the best of my knowledge, this fact is not accounted for in traditional analyses. I argue that it is due to the violation of (1), the checking requirement. I will first show how this problem arises and afterward I will turn to show that the use of the strong form, in fact a spurious strong form, is the repair strategy which allows the failure of the checking requirement to be circumvented.

- (6) a. Regardez-moi/\***me**!  
 Watch me.  
 b. Obéissez-moi/\***me**!  
 Obey me.

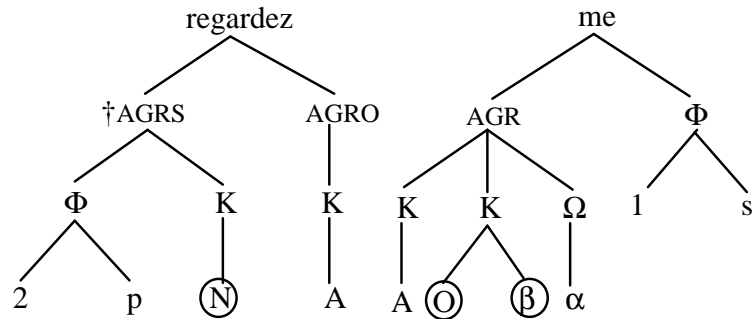
---

<sup>2</sup> I am referring here to imperative verbs, not to imperative utterances in general. Thanks to an evaluator for pointing out this possible source of confusion to me.

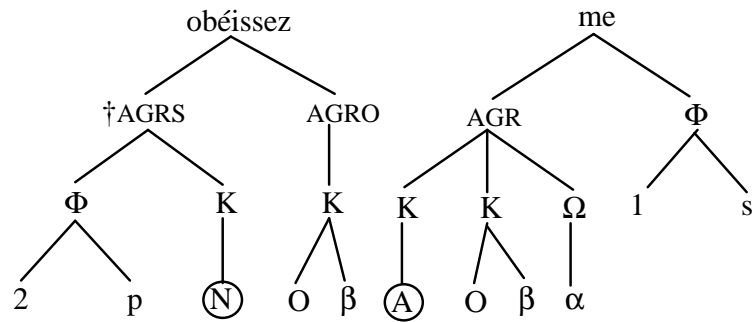
- c.    Regardez-**les**/\*EUX!  
      Watch them.
- d.    Obéissez-**leur**/\*EUX!  
      Obey them.

The derivation of (6a) and (6b) proceeds from the structures given in (7) and (8) respectively, where the Blocking Principle and the COM force the best pronoun to be merged with the verb, that is, a pronoun which is compatible with the verb in the sense of the COM (clause (a)). In these structures, feature checking takes place: the accusative Case of the clitic matches the accusative Case of the verb in (7), while the oblique Case of the clitic matches the oblique Case of the verb in (8). Thus, in both examples, the common features are assigned a tier. However, both the verb and the clitic bear an extra Case in these structures. In (7) (and (8)), the verb bears an additional nominative Case, while the clitic bears an extra oblique Case. In (8), the extra Case of the clitic is accusative, since the verb is specified for oblique. In either case, the nominative Case of the verb and the extra Case of the argument do not match and ill-formedness results. (Extra Cases are circled.) Notice that The  $\Phi$ -features of the verb and those of the argument also clash; indeed, the second person feature of the verb is not checked by an argument in subject position, and thus seemingly conflicts with the first person feature of the object. I claim, however, that only Case features must undergo the checking requirement. This is due to the fact that Cases have different properties from other features. Indeed, they can spread and be saturated by an argument, because probably they are related to grammatical relations; other features are either present or absent in lexical items, but they do not spread onto another element and they have no effect on the distribution of the elements bearing them.

(7)

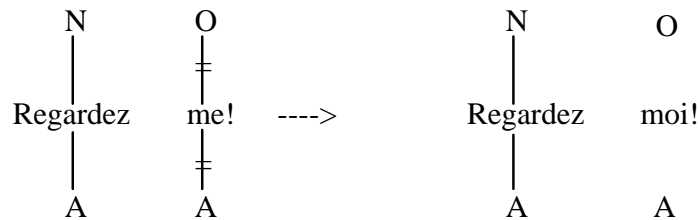


(8)

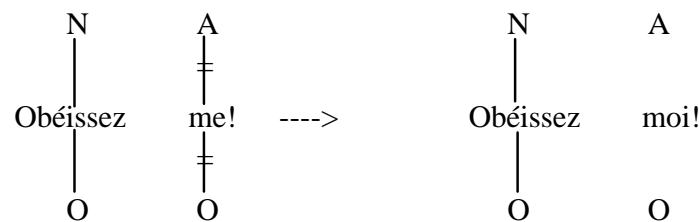


Clearly, (7) and (8) are ill-formed because they contain a Case-mismatch: [N] vs. [O,β] (7) and [N] vs. [A] (8). I would like to suggest that they are repaired by a delinking process, a current operation in nonlinear morphology and phonology (see also Bonet, 1991). The delinking process must affect the argument and not the verb, as discussed immediately below. Given that lexical elements are distinguished from one another by their features, if an element loses some features by a delinking process, it is no longer the same element. Then the question is: what is it? Assuming the Structure Preservation Constraint of Kiparsky (1985), if an element is delinked from its features, one must automatically obtain another existing element of the same paradigm. Thus, in order to repair these structures, one may delink the irrelevant Case of the first person clitic, namely the oblique Case in (7) and the accusative Case in (8). However this operation is incorrect, since it yields two non-existent forms: \*[2,A] and \*[2,O,β] respectively. But by delinking both Cases of the first person pronoun, one obtains the strong pronoun *moi*, as illustrated in (9) and (10) respectively, where class nodes and Φ-features are set aside for simplicity.

(9)



(10)



One may now wonder what happens to the delinked features. In phonology, it is generally assumed that such features are deleted in the representation by what is referred to as the Stray Erasure Convention. If this were to apply here, one would say that the mutation of *me* in *moi* is total, and therefore the resulting form may not have clitic-like properties. However, this is not so. As noted by Kayne (1975, fn 24: 91), the strong forms used in imperative verbs exhibit all clitic properties; in particular they may not be separated from the verb, as shown in (11a-c). (11d) further shows that a derived strong form may intervene between an imperative verb and a genuine clitic.

- (11) a.      Regardez-moi bien!  
               \*Regardez bien moi!  
               Watch me well.
- b.      Écoutez-moi attentivement!  
               \*Écoutez attentivement moi!  
               Listen to me carefully.
- c.      Raconte-moi une histoire!  
               \*Raconte une histoire moi!  
               Tell me a story.
- d.      Raconte-moi-le! (colloquial)

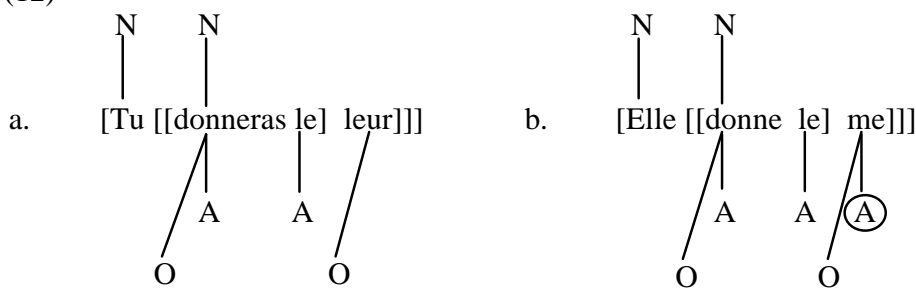
Raconte-le-moi! (standard)

Tell it to me.

Since in my view the clitic-hood of an element depends upon whether or not it bears an inherent Case, it is then clear that the delinked Case is not further deleted in the structure. I thus assume that only the offending Case is deleted by Stray Erasure, namely the oblique-benefactive in (9) and the accusative in (10). Even when floating, the relevant Case is active in the structure, and is responsible for the clitic-like behaviour of the strong form in this context.<sup>3</sup> (We will see in the next section why a Case-specified element must be adjacent to its verb.)

The checking requirement is further illustrated with the examples in (12). In (12a) each Case of the verb correctly matches the Case of each argument, and there is neither an extra Case on the verb nor on the arguments. These structures are therefore well-formed (at this point of the derivation). Similarly, in (12b) each Case of the verb matches the Case of the arguments. However, here the pronoun *me* (doubly specified for Case) bears an extra accusative Case, which is not checked off with the verb, given that the accusative Case of the latter is already saturated by the clitic *le*. The stranding of a single Case does not yield a feature clash, and the structure is well-formed at this point of the derivation.

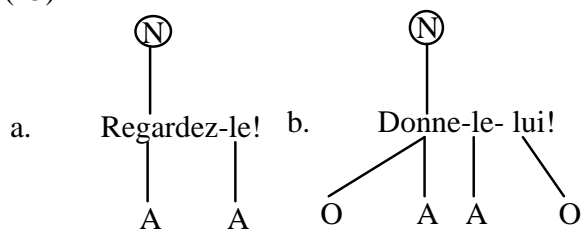
(12)



<sup>3</sup> An evaluator of this thesis claims that it is problematic how a floating feature remains active in a structure. It should be noted that in non-linear phonology, floating features are allowed. Such features are either deleted by Stray Erasure Convention or left floating, in which case they may combine with another feature. Obviously, one knows of the existence of such features iff they affect the output by combining with an adjacent feature. For example, as discussed in chapter 1 a floating low tone (L) may combine with an adjacent high tone (H), yielding a contour tone, (LH). Similarly in syntax a floating Case is assumed to be present if the element bearing it still behaves as other Case-specified elements.

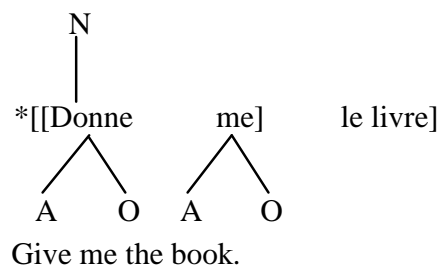
In (12b) the stranded Case belongs to the pronoun, but it could belong to the verb as well. Thus, in the imperative verbs in (13) the nominative Case of the verb does not match with any other Case, and therefore, is stranded. The structure is nevertheless well-formed, since this Case does not clash with another stranded Case.<sup>4</sup> We will show in chapter 4 that feature checking is also confirmed by negative imperatives.

(13)



Consider now an apparent problem for the analysis. The strong form of the clitic appears to be obligatory even in contexts where there are no residual features conflicting with each other. As illustrated in (14), the oblique and the accusative Cases of the imperative verb are saturated by the double Case clitic *me*, and the nominative Case is stranded alone in the structure, parallel to the example in (13). However, (14) is ill-formed.

(14)



This structure is significantly different from the previous ones. In (13) the stranded nominative Case is not saturated, and is ignored; in (14) the nominative Case cooccurs with a (Caseless) NP which saturates it. Clearly this is incorrect, since the

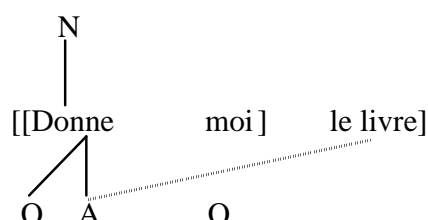
---

<sup>4</sup> The same evaluator observes: "It appears that the author introduces here a specification on what constitutes a clash without explanation." In my view, the notions of feature clash and feature mismatch are synonymous and are already discussed in chapter 1 and made more precise in (1) above.



NP *le livre* is not intended to be the external argument. That is, there is a clash between grammatical role and Case. Therefore the repair strategy must take place in order to prevent the intended direct object from being assigned nominative Case by the verb. The delinking process ensures this outcome. Both Cases of the clitic are deleted, yielding the strong form. Since the clitic is the indirect object, the accusative Case must be deleted by stray erasure, which allows the verb to spread its accusative Case to the NP. Since the latter is properly Case-marked, it no longer interacts with the nominative Case of the imperative verb; this is shown in (15).<sup>5</sup>

(15)



As suggested above, the delinking process does not affect the features of the verb. Now, let us motivate this claim. Since the verb bears tense morphology which contributes to the meaning of the structure, it is natural to assume that it may not be modified by feature manipulation. If, indeed, the nominative Case of the verb were deleted, one would obtain an infinitive verb, since there is no tensed verb which does not have nominative Case. As the meaning of an infinitive verb is significantly different from that of a tensed verb, feature manipulation on the verb would result in a loss of semantic information, unlike feature manipulation on the pronoun. We can then conclude that the delinking process may not affect the features of the verb.

---

<sup>5</sup> With respect to this discussion, the evaluator notes: "[h]ere we are confronted with a series of statements that at best appear to be arbitrary [...]. What role does 'intension' play in grammar?" To repeat, in the present theory there is a relationship between Cases and grammatical roles (GR), such as subject, direct object, etc. (see chapter 1). Thus, each argument enters the derivation with a grammatical role. Since each GR is related to a particular Case, an argument must be assigned the relevant Case in order to obtain a well-formed structure. On this view the term 'intended GR' refers to the GR attached to the element at the beginning of the derivation. The offending term is attested in the literature, in syntax as well as in phonology.

She concludes that "even if one accepts the author explanation for *donne-moi le livre*, his analysis does not explain why *donne-moi* is possible." Since *\*donne-moi* is not a well-formed sentence (perhaps possible in child language), I discard this conclusion.

To conclude, the problems addressed in this section are not accounted for in traditional analyses. In GB theory, they are treated as idiosyncrasies of French grammar. In MP, it is difficult to account for these facts, as it is assumed that all arguments, whether NPs or clitics, are fully specified with relevant Case. With such a global approach to features, one cannot see that first and second person pronouns in French are specified for both oblique and accusative Cases, while third person object pronouns are specified either for accusative or oblique Case. Note that the machinery used here is related to that advocated in Bonet (1991, 1995). However, her system cannot explain why such processes must take place in the grammar, for she does not extend this type of representation to other grammatical categories, verbs in particular. In the analysis suggested here, delinking is required in order to avoid feature clash. Without the conception of checking theory argued for here, one could delink features from a morpheme as a function of the desired output, but one could not predict when this process will take place. In the next section, I turn to consider why clitics appears mainly before the verb. Encliticization with positive imperative verbs will be discussed as well, while similar phenomena in other Romance languages will be set aside until chapter IV.

### **2.3 OCP effects**

I have proposed, along with Bonet (1991), that a well-known phonological process, namely delinking of features, applies also to syntactic features. In this section I will attempt to show that the Obligatory Contour Principle (henceforth OCP) (cf. Goldsmith, 1976; Leben, 1973; MacCarthy, 1986; etc.) is also at work in syntactic configurations. In particular it will be argued that the movement of clitics from their original position, as well as their obligatory adjacency with the verb, is due to the OCP.

In nonlinear phonology and morphology, the OCP is a structural constraint on feature cooccurrence within a domain, generally the morpheme (see chapter 1). It generally has the effect of triggering the fusion of adjacent identical features in phonological representations, hence the definition in (16), which is due to McCarthy (1986: 208). Notice that in phonology, the melodic level as opposed to the skeletal



the inner domains: [<sub>V</sub>' voit le], [<sub>V</sub>' voit te]. In phonology and morphology, the effect of the OCP takes place in a defined domain, generally the morpheme. If the OCP were to apply in syntax also, it would have to be the case that it holds for features belonging to the same syntactic domain. Suppose that the complement domain of the verb, as defined above, is such a domain. Thus, focussing on Case, I formulate the syntactic analog of the OCP as follows.

- (18)           Obligatory Contour Principle (Syntax)<sup>6,7</sup>  
                   Identical Case features are prohibited within the complement domain of the verb.

Given (18), the structures (17a,b), which contain a clitic in the complement domain of the verb, may not be the correct output, unless they are repaired. I argue that the appearance of the clitic to the left of the verb results from the movement of the clitic out of the complement domain of the verb in order to repair the violation of the OCP.

Verb movement being a widely accepted hypothesis in current analyses, one may wonder why the verb is not moved instead of the clitic. This alternative may be rejected on purely conceptual grounds. Since the verb is the central element of the syntactic structure, i.e. the one which projects and ultimately assigns Case, as assumed here, it must be the case that it can't move anywhere. If, indeed, movement

---

<sup>6</sup> The evaluator objects that (18) is not equivalent to the phonological OCP in (16), since there is no contour in the structural description I refer to. Be that as it may, I will continue to refer to (18) as the OCP. (According to McCarthy, 1986: 208, the name OCP (16) is a 'regrettably imperspicuous' one.) Now, the original formulation of the OCP by Goldsmith (1976), drawing on work by Leben (1973), is intended to exclude tone sequences such as HHL and LLH within a morpheme; thus only contour sequences such as HL and LH are allowed. In fact, the OCP does not exclude a morpheme with a single tone, L or H. In the syntactic analog of the OCP (18), a level sequence (as opposed to a contour sequence) of Cases is excluded in the complement domain of the verb, e.g. \*[A A]. Since syntactic strings (sentences) are created productively, as opposed to segmental strings (non-derived words) which are listed in the lexicon, minimally there must be some difference between syntax and phonology with respect to the manifestation of the OCP.

<sup>7</sup> Juan Uriagereka points out (p.c.) that the OCP should be formulated in more general terms, without referring to a specific feature. While I agree with this, I will continue to use the narrow definition in (18), since in the data discussed here there are no other features than Case that are relevant to the OCP. Similarly, the domain of application of (18) is in principle any syntactic domain, but conveniently the formulation refers to the complement domain of the verb, in which the most obvious effect of the OCP, namely movement, appears (see the discussion under 2.3.3.).

of the verb were allowed, in a theory which does not use traces there would result an anomalous structure, such as a VP which contains no verb at all. To take an example, given the structure in (19a), the movement of the verb out of the V' would yield a structure such as (19b), where the V' is no longer headed by a verb, clearly an unwanted result. (The hypothetical landing site of the verb is omitted.)

- (19) a. [VP Marie [V' voit le]]  
 b. \* voit...[VP Marie [V' Ø le]]....

Thus, the only viable alternative is to allow the clitic to move out of the complement domain of the verb, as suggested above. In a parsimonious syntactic structure which contains no functional heads, as assumed here, it is a matter of determining the adjunction site of the moved clitic. I argue that the clitic adjoins to the first element outside the forbidden domain, that is to say the closest element to the verb.<sup>8</sup> For ease of exposition and reference, let us state this as follows:<sup>9</sup>

- (20) Adjunction site of object clitics  
 Object clitics must adjoin to the first element (if any) out of the complement domain of the verb.

In my view, the closest element to the complement domain of the verb (in the sentence above) is the subject. Therefore, the latter must be the adjunction site of the clitic. Notice that (20) says nothing about the manner of adjunction; in fact, as I will show, the grammar need not specify whether clitics are left or right-adjoined. But for the time being, let us assume that clitics are right-adjoined to the subject. Thus, in the structures above, adjoining the clitic to the right of the subject readily yields the correct output as shown in (21). Notice that in my perspective, as there is no trace, when an element is moved, the branch linking it to the tree is assumed to be deleted by stray erasure.

---

<sup>8</sup> To the extent that the derivation is governed by economy, a clitic must at any rate make the shortest move necessary to satisfy the OCP (see also Chomsky, 1995).

<sup>9</sup> While a morpheme-specific formulation is used here, for the sake of simplicity, it will become clear that every element specified for the same Case as the verb will move out of the complement domain of the verb. The adjunction site will depend on other features; for a clitic it is the first element outside the complement domain, but for scopal elements it is different (see chapter 4).

- (21) a. [[Paul le] voit]  
 b. [[Paul te] voit]

If this analysis is correct, any element which legitimately may appear before the verb is predicted to be an adjunction site for the clitic. Negated sentences strongly support this view. Indeed, as shown in (22a), the negative morpheme *ne* appears between the verb and the subject. If the object of the verb is a clitic, it always appears to the right of the negation morpheme. Thus, one may conclude that the clitic is adjoined to *ne*, as illustrated in (22b).

- (22) a. Paul **ne** voit pas son ami.  
           Paul NE sees not his friend.  
           ' Paul does not see his friend.'  
 b. Paul [**ne** le] voit pas.  
           ' Paul does not see him.'

With infinitive verbs, both the negative morphemes (*ne* and *pas*) precede the verb, as shown in (23a). Consistent with (20) the clitic adjoins to the null subject PRO, as can be seen in (23b).<sup>10</sup>

- (23) a. Ne pas voir Marie...  
           NE not to see Marie...  
 b. Ne pas [PRO la] voir...  
           NE not her to see...

With compound tense verbs, clitics always precede the auxiliary verb, which is fully inflected for Case and phi-features (on auxiliaries, see appendix 1). Assuming with Rizzi (1982) that the auxiliary and the participle form a complex verb, the clitic must move out of the complement domain of this complex verb. Thus, as in the previous cases, the landing site is the first element to the left of such a domain. This is illustrated in (24).

---

<sup>10</sup> Given this type of representation, an abstract syntactic element such as PRO must be formally distinguished from a non-existent one. A natural way to do so is to assume that the abstract element PRO is a bare anchor, that is an X-slot without melody, consistent with current practice in phonology (see for instance the discussion in Goldsmith, 1990, and below).

- (24)            [[Paul les] [a reconnus]]  
                  Paul them has recognized.  
                  ' Paul recognized them.'

Our claim that the clitic may be adjoined to any element preceding the verb is seemingly challenged by the traditional view that the adjunct and the head form a syntactic constituent. Traditionally two elements are taken to be a syntactic constituent if some process affects them jointly, for instance movement. Since there is no evidence that the subject and the clitic may move together, under traditional analyses it is legitimate to conclude that they do not form a syntactic constituent, ruling out the idea that they are in an adjunction relationship. In fact, as I will show, the adjunct and its head do not form a constituent in the traditional sense, as far as terminal elements are concerned.<sup>11</sup>

The adjunct and a segment of the head form a constituent only if one uses so-called Chomsky-adjunction. This type of adjunction requires that a segment of the targeted element be created, and the adjunct is attached to it so as to become the sister

---

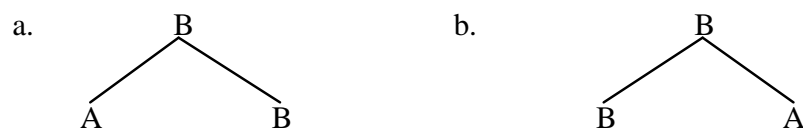
<sup>11</sup> In GB analyses, the clitic and the verb are assumed to form a syntactic constituent. For proponents of the movement analysis, the clitic is moved to the lower projection of the verb  $V^0$ , while for proponents of the base-generation analysis the clitic is generated under the  $V^0$  node. Both analyses crucially rely on the fact in (i), (ia) is a simple affirmative sentence, where normally the clitic appears between the subject and the verb; (ib) on the other hand, is an interrogative sentence, where the subject is inverted with respect to the complex clitic+verb. In traditional analyses, the sentence in (ib) is assumed to be derived from the same D-structure as the one in (ia), precisely by movement of the constituent Clitic+Verb. However, I will argue in chapter 4 that (ia) and (ib) are unrelated.

- (i)     a.        Elle le voit.  
                  She him sees.  
                  ' She sees him.'  
              b.        Le voit-elle?  
                  Him sees she?  
                  ' Does she see him?'

Furthermore, the GB-type analysis is apparently supported by the fact that a proclitic forms a group clitic with the accompanying verb. The clitic group behaves as a distinct phonological unit, as Nespor and Vogel (1986) argues. In fact, this does not contradict my proposal; rather, this observation is accounted for in the present theory. A clitic leans on the verb because of Case Fusion (see below), even though it is not adjoined to. Notice that Klavans (1985) observes that there is not always a correspondence between syntactic constituent and phonological constituent, as far as clitics are concerned. In her view, a clitic may attach in opposite directions in syntax and in phonology. Of course, since clitics are in general monomorphemic words, which must be licensed by a verb with a valid Case (see (53), chap. 1 and below), it is not surprising that they trigger phonological alternations in some languages, e.g. vowel deletion, stress readjustment, etc. (Thanks to an evaluator for pointing out these issues to me.)

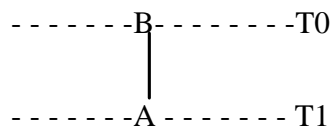
of the original segment. As the operation is binary, it is necessary to specify the linear order of the adjunct with respect to the head; for instance, if an element A is Chomsky-adjoined to an element B, one obtains either (25a) or (25b). One of these alternatives is chosen according to the desired result. In Kayne (1994), for instance, it is explicitly stipulated that clitics must be left-adjoined (like in (25a)) to some functional head.

(25)



Alternatively, suppose that the adjunction operation consists of attaching the adjunct (A) and the head (B) with a single association line, as shown in (26) (I consider shortly T0 and T1). Let us refer to this type of adjunction as Temporal-Adjunction, a term that will become clear.<sup>12</sup>

(26)



Suppose further that the terminal elements of syntactic structures are initially on the same level (or tier). Given that such a level contains morphemes and not abstract features, I take it to be similar to the *timing tier* in nonlinear phonology. Let us refer to this initial tier as T0. Then, a terminal element (of the initial tier) which is adjoined to another one, as in (26), is no longer on the same timing tier as the rest of the structure. Let us refer to the new timing tier, which is dependent of the former, as T1. As can be seen, this type of adjunction gives no linear order between the adjunct and the head. However, it is clear that at the output, a single string is perceived; I suggest that this is due to a buffer that levels (or linearizes) the structure at the end of the

---

<sup>12</sup> According to the evaluator, Temporal-Adjunction is an unfortunate choice of term. I will continue to use it, for want of a better term.



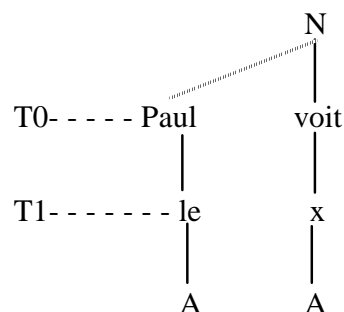


' = ' .) Notice that, as mentioned above, in this representation of adjunction, the head and the adjunct may not be seen as a constituent, since each one lies on its own tier. With respect to the movement process, the complex B=A in (28) may not move under the simplest assumptions, since they are not dominated by the same node, and they are not on the same tier. Furthermore, it is natural to assume that the head B and the adjunct A may no longer move, as their movement would imply that the adjunction operation would have to be undone.

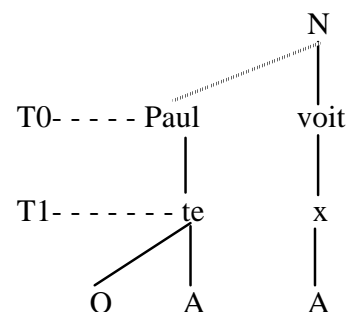
(28)        - - - - - B=A - - - - - T

Let us return now to the structures in (21a,b). Given the proposal regarding adjunction just discussed, these must be more accurately represented as in (29). Notice that there is no distinction between the adjunction line and the line associating features with morphemes. Further, I assume that a dummy node (x) is generated under the verb, creating a path in order to maintain the accusative Case of the verb and the moved argument in the same plane. This is reminiscent of the Node Generation Convention of Archangeli and Pulleyblank (see, for instance, Paradis and Prunet, 1993).

(29) a.

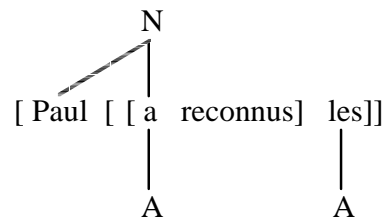


b.

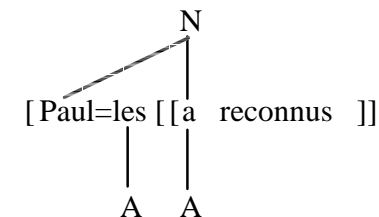


Similarly, the derivation of example (24) is given in (30). The adjunction line is replaced by the symbol ' = ' for convenience.

(30) a.



b.



It is clear that by virtue of the OCP a clitic (i.e. a Case-specified pronoun) may not remain in the complement domain of the verb. And, consistent with (18), it must adjoin to the first element outside this domain, whether such an element is an NP, a non-overt pronoun (PRO), the morpheme of negation, or any element that can legitimately occupy such a position. I now discuss a first piece of independent evidence for this, as well as two apparent counterexamples.

Indeed, if it is true that the OCP forces Case-specified pronouns to move outside the complement domain of the verb, it is expected that a Caseless element remains in situ. This prediction is borne out: all things being equal, NP arguments never move, as illustrated in (31). Similarly, if the object of the verb is a Caseless pronoun, there is no movement, as expected; compare (32a) and (32b).<sup>16</sup> (Notice that in (32a) the strong pronoun must bear a pitch accent, otherwise the structure is ill-formed; see below.)

- (31) a. Marie aime Paul.  
       b. \*Marie Paul aime.  
           Marie likes Paul.
- (32) a. Tu as vu MOI!  
       b. \*Tu MOI as vu!  
           You saw ME!

---

<sup>16</sup> For reasons of space, only pronouns are discussed. Thus, I will have nothing to say about morphologically rich languages where DPs/NPs bear overt morphological Case. An evaluator points out that in Greek, for instance, personal pronouns are in most cases proclitics to verbs, as in Romance, but NPs/DPs, which are declined for Case, remain to the right of the verb, i.e. they do not move in my view; compare (i) and (ii) (Angela Ralli, p.c.).

- (i) I        Maria    djavazi    to        vivlio.  
       the-FEM Mary-NOM reads        the-MAS book-ACC  
       ' Mary reads the books.'
- (ii) I        Maria to        djavazi.  
       the-FEM Mary it-MAS.ACC reads  
       ' Mary reads it.'

In favor of the OCP, however, we may note that most languages which have Case on the object argument are of the SOV type, as already noted by Greenberg. Furthermore, as is well known, scrambling processes occur mostly in morphologically rich languages, a phenomenon which looks like a series of gemination effects (see below). The Greek case may be related to the exceptional instances of infinitives and gerundives in Spanish and Italian, where procliticization is not allowed (see chapter IV), and therefore can arguably be handled by the same type of analysis.

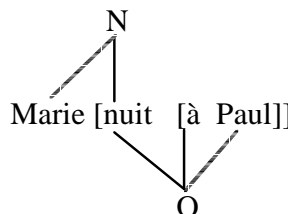
Besides the strong forms of the paradigm of personal pronouns, there is another pronoun in French which is arguably unspecified for Case. This is the pronoun *ça* 'that'. This pronoun may be used either as subject or object of the verb, thus displaying the same distribution as an NP; indeed in (33a) it is the subject of the verb, while in (33b) it is the object of the verb. As predicted by the OCP, such a pronoun may never appear in a preverbal position when it is the object of the verb, as illustrated by the ill-formedness of (33c).

- (33) a. *Ça marche pas!*  
That does not work!  
b. *Il connaît bien ça.*  
c. *\*Il ça connaît bien.*  
He knows that well.

Consider the case where the argument of the verb is an *à*-phrase, as illustrated in (34a). Since the morpheme *à* is a Case-marker, it is expected that the *à*-phrase should move outside the complement domain of the verb, like clitics. However, this is not the case; the movement of the *à*-phrase yields ill-formedness, as shown in (34b). Two analyses are possible here: (a) the OCP is not violated in (34a) because it is irrelevant, or (b) the OCP is relevant and is violated because it cannot be correctly repaired. Consider the first analysis. The preposition and its complement form a distinct domain, as can be seen in (34a). The preposition, being a Case-assigner, may not saturate the Case of the verb; rather, the whole *à*-phrase saturates the Case of the verb. If one assumes that fusion of the Case of the verb and that of the preposition takes place, then it remains a single Case in the complement domain of the verb, and the NP *Paul* saturates it as shown in (34c). Under this analysis, the argument *Paul* saturates both the verb and the preposition (Case fusion is discussed below).

- (34) a. *Marie [nuit [à Paul]].*  
b. *\*Marie=[à Paul] nuit.*  
Marie disturbs Paul.

c.

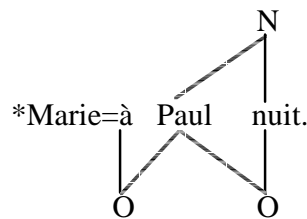


As stated in (b) above, another analysis is possible. Suppose that the *à*-phrase is moved to the subject *Marie*, as shown in (35) (= (34b)). One can see that this structure is ill-formed because it yields further constraint violations. Indeed, a Case conflict and a mismatch between the nominative Case of the verb and the grammatical function of the *à*-phrase appears in the structure: on the one hand *Paul* is assigned oblique Case by the preposition (and logically by the verb also), and on the other it is assigned nominative Case by the verb. In addition, the indirect object comes to intercept the nominative Case assigned to the subject. If movement does not take place such a problem will not arise. Furthermore, the oblique Case of the verb and that of the preposition are adjacent in (34c), but not in (35). As I will show in the next section, a further effect of the OCP is that Case-specified elements must be adjacent in the structure, insofar as possible. Indeed, even after movement the object clitic remains adjacent to the verb, as can be seen in the structures discussed above. Thus, since the movement of the *à*-phrase would trigger more problems than it would solve (namely Case conflict, lack of adjacency of the preposition and the verb), it is blocked. Notice that under the first analysis, if an *à*-phrase moves to clause initial position (for reasons to be discussed in chapter 4), a well-formed structure is expected, since there is no Case conflict; cf. *tu parles à qui / à qui parles-tu?* 'to whom do you speak?' (see chapter 4).<sup>7</sup>

---

<sup>17</sup> I should note another objection. An evaluator notes: " [h]ere we find again a statement which is surprising. It appears that the author has no qualms about recurring to ad-hoc assumptions to handle problems with his analysis. It now appears that Desouvrey assumes the existence of some superordinate mechanism that evaluates the number of outcomes to a particular problem and chooses that which is associated with the least number of problems. What is the nature of this mechanism? When does this evaluation takes place.' ' It appears that I did not propose some superordinate mechanism. My analysis of the ill-formedness of (35) relies on independently needed assumptions and constraints which are used throughout the thesis. To repeat my assumptions, all transformations apply simultaneously and neither transformations nor constraints are explicitly ranked, in contrast to Optimality Theory, for instance. It follows that a constraint may not be repaired at the expense of a

(35)



It should be noted that the fact that two analyses are possible does not entail that the theory must explicitly stipulate the ordering of transformations. Recall that our crucial assumptions are the following ones: transformations (and constraints of the same type) are not ordered, and they apply simultaneously; and a maximal number of constraints must be satisfied in a given structure. Under these requirements, if an element may undergo two different transformations, say A and B, and there is an asymmetric relationship between them such that A feeds B but B does not feed A, it is clear that in order for the maximal number of constraints to be satisfied, A must apply before B. But if there is no feeding relationship between A and B, there can be no ordering, and two different derivations may take place, one for A and the other for B, yielding eventually different outputs. As for simultaneity, it will take place when in a given structure two elements x and y undergo the transformations A and B respectively. For instance, if x and y must move to different targets, their movements take place at the same time (but see below).

Another apparent piece of counter-evidence for the movement-under-OCP analysis is the fact that in imperative verbs, which are fully inflected for Case, as discussed above, weak pronouns are enclitic, that is, they remain in situ. I would like to suggest that clitics do not move in positive imperative verbs because there is no available landing site outside the complement domain of the verb. Indeed, as

---

further constraint. For instance, suppose that an element A has been moved to a position B by virtue of a constraint C. If A violates a further constraint C' in B, a further transformation is required to repair C', since it is not lower ranked than C. This transformation will eventually move A back to its initial position. At this point the derivation is cut off by virtue of the ban on string-vacuous operations; otherwise the derivation would loop at infinitum. Similarly in (35) the à-phrase would move again, ending up to its initial position. In other words, movement in this case creates more problems, hence the offending element remains in situ with a constraint violation. This is an explanation which follows from my assumptions, not an ad-hoc mechanism.

discussed earlier, imperative verbs do not have null subjects, in contrast with infinitive verbs. Some examples are given in (36).<sup>18</sup>

- (36) a. Admirez-le!  
Admire him!
- b. Obéissez-lui!  
Obey him!
- c. Donnez-lui un livre!  
Give him a book!

Notice that the violation of the OCP in imperative verbs triggers the appearance of a pitch accent, as noted in chapter 1 in the context of the Blocking Principle. The pitch accent falls precisely on the offending element, namely the clitic, which is usually incompatible with pitch, as standardly assumed in the literature. Thus while the third person accusative clitic *le* is usually spelled out without the schwa (*le-->l'*), in positive imperative structures, the vowel must be spelled out, as shown in (37) (phonetic spelling in brackets). Notice that one could not explain (37a) by saying that the clitic is in a stressed position, that is the last syllable of the phrase in French. In fact, if that were the case, the clitic without the schwa would be syllabified with the verb, precisely as the coda of the last syllable, on which the stress would fall.<sup>19</sup> As shown in (37b) this is impossible. Furthermore, additional facts show that the clitic is stressed even in contexts where it is not the rightmost element in the structure. Indeed, when the imperative is followed by an infinitival clause, the clitic complement of the infinitive verb cannot move outside the domain of the imperative verb, and it bears a pitch accent, as in *fais-le/\*l' entre* 'let him come in'. Clearly this supports our interpretation, namely that the pitch accent is related to the failure of movement to take place; it may not be interpreted in the sense that the clitic is the

---

<sup>18</sup> An account of the semantics of imperative verbs fall far beyond the scope of this work. Thus, I have nothing to say about verbs which are incompatible with this mood, including seem-type verbs.

<sup>19</sup> Notice that in Haitian Creole, the object pronoun has a strong form and a weak form, *li* vs *l'* 'him, her, it', both being unspecified for Case (see chapter 4). In this language, stress patterns as in French, that is it falls on the last syllable (see Cadely, 1994). Interestingly, in imperative verbs, the weak form is usually used, and it is syllabified as the coda of the stressed syllable of the verb; thus: *achte-l* [aʃ-**tel**] 'buy it', *li/l antre* 'let him come in'.

focus of the structure, although such an interpretation is not systematically excluded (see Selkirk, 1995).

- (37) a. Achetez-**le**! [aʃte**l**ə]  
 b. \*Achetez-**l'**! [aʃte**l**]  
 Buy it.

This analysis of imperative verbs in terms of OCP violations due to the absence of a possible landing site is strikingly confirmed by the fact that in negated imperatives, clitics normally move. The part of the morpheme of negation which precedes the verb is a landing site for clitics. (I will return to these structures in chapter 4.)

- (38) a. Ne=le regardez pas!  
 NE him look-at not.  
 ' Do not look at him!'  
 b. Ne=lui montrez pas ce livre!  
 NE him show not this book.  
 ' Do not show him this book!'

Interestingly, in dialects of French where the morpheme *ne* is not used, clitics do not move in negated imperative verbs, as expected under our analysis. Thus compare:

- (39) a. Jean=**le** voit pas.  
 Jean him sees not.  
 ' Jean does not see him.'  
 b. Regarde-**le** pas!  
 Look at him not.  
 ' Do not look at him.'

In some *ne*-less dialects, the object clitic may precede the negated imperative verb, as illustrated in (40). Julie Auger suggests (p. c.) that this may be due to the fact that the morpheme *ne* is present underlyingly but not phonetically in such dialects, while in other dialects there is no *ne* at all. A natural way to implement Auger's suggestion is to assume that French dialects vary according to whether or not each



segment of *ne* is linked to a skeletal slot (see below on skeletal slots). Thus, while in the *ne*-dialect, only the vowel is floating (41a), in *ne*-less dialects of the type illustrated in (40) both the vowel and the consonant are floating (41b).<sup>20</sup> Thus in the dialect which has (41b) the clitic is adjoined to an empty skeletal slot, which is never associated with a melody.

- (40) a. Le regarde pas!  
Him look-at not.  
' Do not look at him!'
- b. T' en fais pas!  
You EN worry not.  
' Don' t worry for that.'

- (41)
- |   |  |
|---|--|
| <p>a.      n    e</p> <p style="margin-left: 100px;"> </p> <p style="margin-left: 100px;">X X</p> | <p>b.      n    e</p> <p style="margin-left: 100px;">X   X</p> |
|---|--|

Note that our account of the imperative does not complicate the system, as no further claim is introduced. Indeed, no constraint may refer especially to imperative verbs. What is needed is the natural assumption that a structural operation (movement) applies whenever it is possible.

In sum, I have argued that in French, object clitics (or Case-specified pronouns) are generated to the right of the verb, like ordinary NP arguments, and that their movement is intended to repair the violation of the OCP. By assuming the simplest hypothesis, namely that the landing site of clitics is the first element outside the complement domain of the verb, the correct result is derived for many types of sentences. Now, it must be shown that this adjunction site of clitics, which is rather

---

<sup>20</sup> That the representation in (41a) is correct for the *ne*-dialect (mainly the standard dialect) is confirmed by the fact that in current speech, the *e* of *ne* is always silent, unless the other segment *n* cannot be syllabified with an adjacent word. Thus:

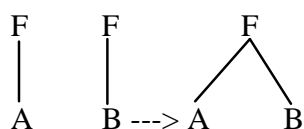
- |     |  |  |
|-----|--|--|
| (i) | <p>a. Marie n connaît pas Paul.</p> <p>b. Marie n' a pas connu Paul.</p> <p>c. Paul *n/ne connaît pas Marie.</p> | <p>' Marie does not know Paul.'</p> <p>' Marie did not know Paul.'</p> <p>' Paul does not know Marie.'</p> |
|-----|--|--|

unfamiliar to many, satisfies all tests of cliticization. In particular, I need to explain why no element can intervene between clitics and verbs, the main feature which motivates the analysis that a clitic form a constituent with the verb it is an argument of. In the next section, I show that this property of clitics follows from another manifestation of the OCP in syntactic structure, namely the fusion of adjacent identical features.

### 2.3.2 Fusion

Another interaction between syntactic features is what we take to be the equivalent of fusion, a process common in nonlinear phonology and morphology. Fusion is the process by which two adjacent identical features come to merge, yielding a single feature associated with two elements, as illustrated in (42). (F stands for features, and A and B are lexical elements.)

(42)



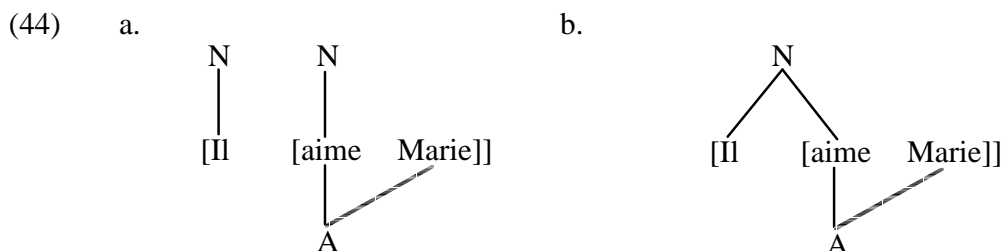
Fusion is the second effect of the OCP that will be discussed in this section. I will refer to it as OCP-2, as opposed to OCP-1, which triggers movement. In the context of syntactic structure, I propose the following formulation of OCP-2.

(43)

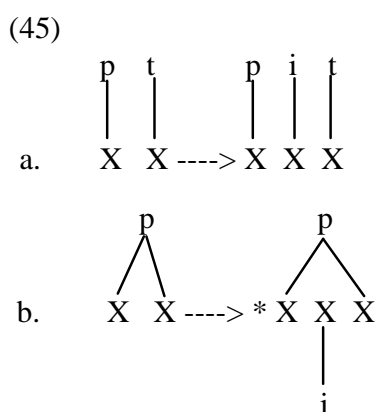
#### Obligatory Contour Principle-2

In a given tier, all occurrences of a Case feature must be reduced to a single one by fusion.

Now the two versions of the OCP are seemingly in conflict. One may wonder, indeed, why fusion does not take place when the occurrences of a Case feature belong to two elements of the same domain. In fact, it appears that there is no conflict: the object clitic moves and its feature fuses with that of the verb, both processes occurring simultaneously under my assumptions. Notice that when the clitic is not within the complement domain of the verb, only fusion occurs, as illustrated by the derivation in (44).



As this process is not directly observed on the surface, what type of evidence can we find to justify it? A piece of evidence is provided by the fact that fusion of Case features, as illustrated in (44b), is much more like a geminate segment in nonlinear phonology and morphology. A geminate segment is a vowel or a consonantal melody which is linked to two X-slots. It is observed that geminate consonants are *inseparable* in that they repel the insertion of an intervening segment by rules which are otherwise possible in nongeminate contexts. If a language has a rule that breaks the sequence of two consonants by inserting a vowel, as in the hypothetical example (45a), this rule will fail to apply when the consonants arise from gemination, as illustrated in (45b). Besides *inseparability*, another well-known characteristic of geminate segments is their *inalterability*: a rule cannot delink a branch individually without affecting the other also. Thus, in the hypothetical word *batta*, a rule may affect both parts of the geminate, for instance *tt* may become *pp*, yielding *bappa*; but the first *t* alone may not be altered to derive the word *\*bapta*.<sup>21</sup>



<sup>21</sup> With respect to geminates in phonology, see Guerssel (1977, 1978), Hayes (1986), Schein and Steriade (1986), or a textbook such as Kenstowicz (1994).

Similarly in syntax, as I will show, the process whereby two occurrences of a Case feature are reduced to a single one which comes to be associated with two morphemes, as in (44b), presents many similarities with geminate segments in phonology. Indeed, clitics and verbs behave as if they were inseparable: in most cases they cannot be separated by an intervening element, except another clitic (see below).<sup>22</sup>

Furthermore, verb-clitic clusters are also inalterable, like geminate segments in phonology. Kayne (1975: 92, fn. 26) observed the intriguing facts in (46). As can be seen, an intervening clause is allowed between the subject NP and the verb (46a). But if the subject is a clitic, the sentence is ill-formed, as expected (46b). The interesting point here is the fact that even a strong pronoun, which is Caseless, yields an ill-formed sentence, as shown in (46c).

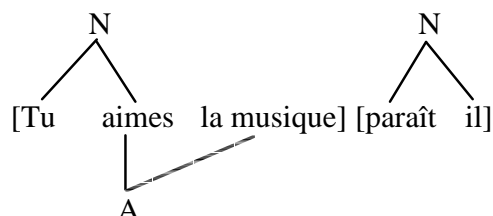
- (46) a. Jean, paraît-il, est heureux.  
           Jean, it appears, is happy.  
       b. \*Tu, paraît-il, aimes la musique.  
       c. \*Toi, paraît-il, aimes la musique.  
           You, it appears, like music.

Under the analysis of relative clauses advocated in Desouvrey (1996, 1997b) (see footnote 13), the relative clause and the main clause are generated separately as two independent and unordered structures. Suppose that this type of analysis holds for any complex structure with a root clause and a subordinate clause. If so, (46a) may have been generated as shown in (47) (the order of the substructures is irrelevant as well as their internal constituency). In the left-hand structure, the subject being a clitic, fusion takes place automatically, and therefore the adjunction of the right clause to an X-slot between the clitic and the verb is prohibited, given the inseparability effect of fusion.

---

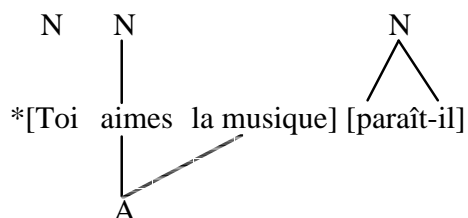
<sup>22</sup> As discussed in chapter 1, traditional analyses take this fact to be a defining property of clitics.

(47)

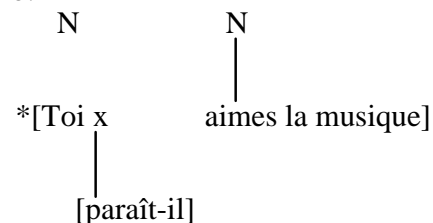


Now as discussed above with imperative verbs, delinking may take place to eliminate an offending feature in the derivation. Thus, one would expect that in the main clause, the clitic *tu* would be delinked from the nominative Case in order to allow the subordinate clause to be inserted. If such a transformation were possible, the derivation would be as in (48): the second person pronoun *tu* becomes a spurious strong form *toi* by losing its Case feature (48a), and then the right-hand clause is adjoined to a bare X-slot (48b) (accusative Case set aside for simplicity). But such a derivation is incorrect precisely because fusion is inalterable. Put another way, fusion may not be undone by delinking an association line from one of the involved elements.

(48) a.



b.



Notice that if the nominative Case is delinked from both the verb and the clitic in order to overcome the inalterability of fusion, one would obtain the structure in (49). In effect, one cannot just delink the nominative Case from the verb, for there is no verbal form with tensed morphology which is not specified for nominative Case. Thus, given structure preservation the tense feature must be also eliminated, yielding the infinitive form of the verb. Of course, since infinitive verbs and tensed verbs are obviously not equivalent semantically, this transformation is blocked.

(49) \*Toi=paraît-il aimer la musique.

To conclude, fusion of Case features under OCP-2 yields a structure similar to geminate segments in phonology: the morphemes whose Cases have undergone fusion are inseparable in that they repel the adjunction of an intervening structure, and inalterable in that they cannot undergo a delinking process separately. However, this is not always true; the negative morpheme *ne* as well as object clitics, which are moved under OCP-1, may intervene between the verb and subject clitics. On theoretical grounds this may not be a problem because constraints are not absolute (see also Prince and Smolensky, 1993). In the next subsection, I will attempt to account for the fact that object clitics may intervene between the verb and the subject clitic, postponing the discussion of negation until chapter 4. Also, a formal account of the geminate-like properties of fusion will be proposed, and an analysis of the fusion of accusative features after the movement outside the complement domain of the verb will be given.

### 2.3.2.1 X-slot theory

As noted earlier, in nonlinear phonology each morpheme is linked to a skeleton or timing tier, which may be either a sequence of undifferentiated positions (i.e. neither vowel nor consonant), referred to as X-slots (cf. Halle and Vergnaud, 1980; Kaye and Lowenstamm, 1984; Goldsmith, 1990; etc.), or a prosodic position referred to as a mora in what is called the moraic model (cf. Hyman, 1985; Hayes, 1989; etc.) (see Kenstowicz, 1994, for a comparison of these two theories). Let us assume the X-slot theory of the skeleton. Under this theory, a simple sentence such as the one in (50) will have the representation in (51), where each segment is linked to an X-slot. (Phonetic spelling is required with this representation.)

(50) Paul connaît Marie.  
Paul knows Marie.

(51) pɔl kɔnɛ mari  
| | | | | | | |  
xxx xxx xxxx

Now, if in (51) one replaces the object NP by a clitic, one obtains the structure in (52). As discussed above, the clitic must move out of the complement domain. I have argued that the adjunction site of the clitic is the first element outside this domain, here the subject. The latter being composed of a sequence of timing units, the X-slots, it must be the case that the clitic adjoins to the nearest edge, namely the rightmost segment of [pɔl], as illustrated in (53).

(52)            pɔl kɔnɛ la  
                  ||| |||| ||  
                  xxx xxxx xx

(53)            pɔl      kɔnɛ  
                  |||      ||||  
                  xxx      xxxx  
                  |  
                  xx  
                  ||  
                  la

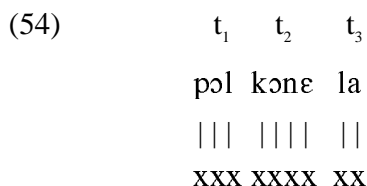
As can be seen, the adjoined element and the head come to constitute a single bimorphemic element, although each morpheme is on a distinct timing tier.<sup>23</sup> This representation permits us to capture the traditional stipulation that the adjunct and the head are tightly tied in such a way that no further element can be inserted between them.

Furthermore, as each X-slot is a timing unit, an element associated with two X-slots is twice as long as another one which is associated with a single X-slot. Similarly, if a morpheme contains three X-slots, it is longer than another which contains only one or two. Moreover, phonetically, each segment of a morpheme is articulated sequentially such that in a word containing more than one segment, the leftmost segment is articulated, and hence perceived, before the rightmost one. Given that each audible segment is associated to a skeletal slot, the X-slots can be thought of

---

<sup>23</sup> Temporal adjunction is somewhat analogous to the segregation of morphemes onto distinct tiers under the Morpheme Tier Hypothesis (see McCarthy, 1986).

as linearly arranged in time. Since words are formed out of segments, it must be the case that each word in a syntactic string is in the timing tier with respect to another word. Thus, suppose that the leftmost word in a syntactic string is at time  $t_1$ , the second one is at time  $t_2$ , and so on until the rightmost word which is the last element in the time axis. In the structure in (52), it follows that the subject is at time  $t_1$ , the verb at time  $t_2$ , and the object clitic at time  $t_3$ , as shown in (54).

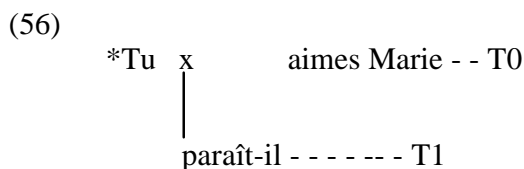


From this perspective, each morpheme in a structure corresponds to a time-interval. Thus, when the clitic is adjoined to the subject, as discussed above, the time axis comes to be reduced, as a time interval is suppressed. In effect, as can be seen in (53) the clitic and the subject form a single (bimorphemic) element and therefore occupy the same time-interval  $t_1$ , which is the sum of their respective X-slots, namely five units after linearization. It appears then that adjunction of one head to another implies the merging of two time-intervals, hence the term of *temporal adjunction*. Put another way, the adjunct and the head are contemporary in the time axis. If this view is correct, one may assume that adjunction of a single element to a specified head does not break the inseparability of fusion, as each part remains temporally adjacent to each other, i.e. at contiguous times.

The result is quite different, however, if one adjoins a clause to a Case-specified head. As mentioned above, we take a clause to be adjoined to a bare temporal slot inside a root clause. Thus the ill-formed sentence in (46b), repeated below as (55), would have the structure shown in (56). As can be seen, unlike the adjunction of a simple clitic to the subject, the adjunction of a clause to the Case-specified head breaks the temporal adjacency of the latter with the verb, since additional time-intervals are introduced in the structure.

(55) \*Tu, paraît-il, aime la musique.



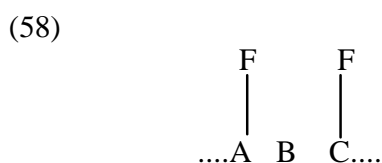


In the light of this temporal relation, a clearer picture of the inseparability-inalterability of fusion emerges. Case fusion repels the insertion of an intervening morpheme between the Case-bearing morphemes because the latter must be temporally adjacent to each other. The adjunction of a single element to the relevant Case-bearing morpheme does not affect this temporal adjacency, since the adjunct and the head come to form a single time-interval. The features under fusion are not affected by the process of adjunction, and hence remain inalterable. To conclude, Case fusion requires the temporal adjacency of the Case-specified morphemes on the timing tier. Let us refer to this property as a constraint on the fusion process.

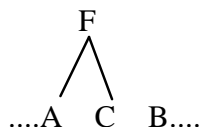
(57) Constraint on fusion

When two occurrences of a Case undergo fusion, the elements bearing these Cases may be separated from each other by at most one time interval.

Clearly, (57) rules out the adjunction of a string of many time-intervals to a member of a cluster of Case-specified elements. Also, as I will show, elements which are base-generated between Case-specified elements must move in order to allow them (the Case-specified elements) to be temporally adjacent. That is, given a string of the type illustrated in (58), where A and C are specified for the feature F, this string must be modified by movement of B, as shown in (59). I will refer to this kind of movement as a *gemination effect*.

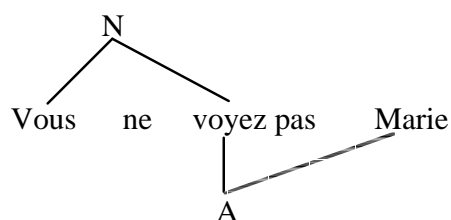


(59)



However, as with OCP-1 and other structural constraints, the gemination effect does not occur automatically in all contexts. For a particular reason that will be discussed in chapter 4, the structure in (60), where the negative morpheme intervenes between the verb and the subject clitic, is well formed, while those in (61), which are expected under the analysis in (59), are not.

(60)



You NE see not Marie.

' You do not see Marie.'

(61) a. \*Vous voyez ne pas Marie.

You see NE not Marie.

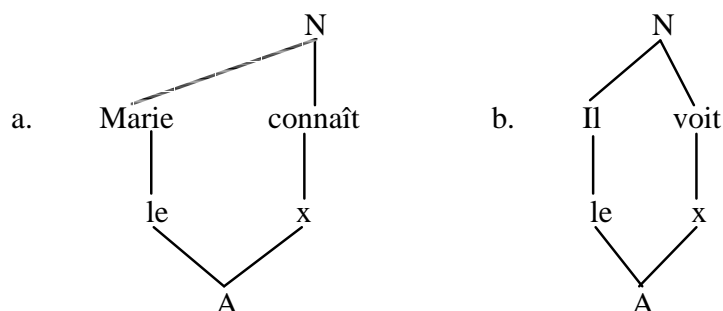
b. \*Ne vous voyez pas Marie.

NE you see not Marie.

Furthermore, as noted above both OCP-1 and OCP-2 apply to object clitics. In fact, nothing in the representation prevents object clitics from undergoing fusion with the verb after moving outside the complement domain of the verb. As illustrated in (62a) and (62b), the accusative Case of the object clitic and the accusative Case of the verb fuse, yielding a single occurrence of this feature (notice that in (62b) there is fusion of nominative Case also). As in the case of subject clitics, no sequence of elements (such as a relative clause) may intervene between the object clitic and the verb, as expected under the fusion analysis. Nevertheless, a single element such as

another clitic may intervene between the object clitic and the verb, as illustrated in (63). (The examples in (63), which contain a clitic cluster, will be discussed below.)

(62)



- (63) a. Marie=le=lui a donné.  
Marie it him has given.  
' Marie gave it to him.'
- b. Elle=lui=en a offert une.  
She him of-it has offered one.  
' She offered one to him.'

To conclude, the notion of time is introduced into syntactic structures. An adjoined element and its head constitute a single time-interval and therefore both are temporally adjacent with the verb. By the fusion process, pronominal clitics, and more generally all specified elements, must be as close as possible to the verb.

### 2.3.3 Discussion of the OCP

Why does the OCP exist? The OCP could be thought of in terms of economy of representation, as opposed to economy of derivation. Clearly a grammar is more highly valued if it contains a smaller number of occurrences of any feature in the tiers. In both of its manifestations (movement and fusion), the OCP significantly simplifies the structure: under movement, morpheme concatenation takes place via

temporal adjunction, yielding a structure with fewer time-intervals on the timing tier; under fusion, two occurrences of a given feature are reduced to a single one.<sup>24</sup>

A more technical issue is why resolution of the OCP by movement is limited to the complement domain of the verb, and not extended to a larger domain, the sentence for example. In fact OCP-1 and OCP-2 turn out to be two sides of the same constraint. Indeed, when an element moves outside the complement domain of the verb under OCP-1, it must land at the nearest element out of this domain so that fusion can take place under OCP-2. For instance, in the structure in (64) suppose that D and E are specified for the same Case feature. Under OCP-1, E must move to the nearest element out of the domain *d*, namely C. If E were to move to B, one would obtain a structural description similar to (58), as shown in (64b), and C would be forced to move from its position between E and D (gemination effect). Assuming the minimal hypothesis that if a transformation may be avoided it should be, the moved element must target the nearest head outside the forbidden domain so that it will not trigger further transformations. This is equivalent to the *Shortest Move* of Chomsky (1992).

- (64) a. [a A [b B [c C [d D E]]]]  
 b. \*[a A [b B=E [c C [d D t]]]]

Now suppose that in (64a) C and D are specified for the same feature. As can be seen, the domain *d* is embedded within the domain *c*, and thus C and D can be

---

<sup>24</sup> Note that as an economy constraint, it is expected that the OCP may be violated in some contexts where it would compromise the intended meaning. Indeed, the OCP does not move a clitic outside of the complement domain of the verb when the resulting structure would be uninterpretable. For instance, in causative structures the clitic argument of the embedded verb must move to the causative verb, as shown in (i). However the reflexive clitic fails to move, as illustrated in (ii). Even though a full account of causative structure falls far beyond the scope of this thesis, I would like to point out that the appearance of *se* to the right of the matrix verb is required because it must be bound by the object *Paul*. If *se* is adjoined to the subject of the matrix verb, as is the case of (ia), it could no longer be bound by the intended referent; rather the subject of the matrix verb will bind it. In other words, in such a structure, the OCP is in conflict with another constraint relating to the interpretation of anaphors. Since the OCP is a phonological-type constraint, it fails to apply.

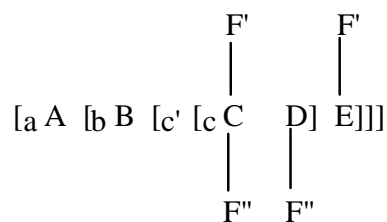
- (i) a. Marie **lui** fait lire un livre.  
 b. \*Marie fait **lui** lire un livre.  
 Marie makes him read a book.  
 (ii) a. Marie fait **se** tuer Paul.  
 b. \*Marie **se** fait tuer Paul.  
 Marie makes Paul kill himself.

taken to be within the same domain  $c$ . Under this view, if  $C$  moves out of  $c$ , it must land at the nearest element  $B$  in order to fuse its feature with the feature of  $D$  without triggering the movement of  $B$ . In such a case, one would obtain the structure in (65), once it is assumed that when one element moves, the branch linking it to the rest of the structure, as well as the intermediate dominating projection, is deleted by Stray Erasure Convention. As can be seen, it results in a structure with a smaller number of time-intervals, but at the output there is no visible variation, as all elements still appear in the same original order. If this transformation were possible, it would be difficult for a child to detect it. That is, given that the input and the output are identical, the young learner cannot infer the existence of any transformation. Hence, I adopt the null hypothesis that string-vacuous movement is not allowed; that is to say, movement occurs only when the element to be moved will pass over a head.<sup>25</sup>

(65)  $*[a\ A\ [b\ B=C\ [d\ D\ E]]]$

It is clear from this discussion that moving the subject clitic under OCP-1 yields string-vacuous movement, where the input and the output are similar. Now consider the structure in (66), where the head  $C$  projects twice. Suppose that  $C$  is specified for two Case features,  $F'$  and  $F''$ , and that  $D$  and  $E$  are each specified for a feature of  $C$ ,  $F'$  and  $F''$  respectively. In this configuration, there are two overlapping domains:  $c$  and  $c'$ , where both  $D$  and  $E$  are to the right of the head  $C$ . Under OCP-1, both  $D$  and  $E$  may move outside their respective domain, since in doing so they will pass over the head  $C$ . Such a structure will be discussed in the next section.

(66)



Another question that arises is why fusion does not take place without moving the object clitic. That is, why does fusion not occur in the complement domain of the

---

<sup>25</sup> With respect to string-vacuous movement, see Chomsky (1973), who rejects it, and Clements et al. (1983), who argue for it.

verb? Consider the schematic representation of the fusion configuration and the movement configuration, illustrated in (67a) and (67b) respectively. As can be seen, in (67a) both elements are adjacent but they are not dominated by the same node, while in (67b) the elements are adjacent and daughter to the same node. Since there is a structural difference between these structures, it is expected that this difference is reflected in the output.

- (67) a. ... A [ B ....    b. ...[A    B]...

I will indeed argue that fusion is stronger in configuration (67b) than in (67a). When fusion is stronger, the two morphemes should be more tightly bound to each other. If so, it is expected that the inseparability of fusion is such that it repels the insertion of even a single element by temporal adjunction. Imperative verbs provide evidence for this analysis. An adverb may intervene between the imperative verb and its NP argument, as illustrated in (68a). But the adverb may not intervene between the imperative and the argument when the latter is a clitic, as can be seen in (68b,c). Assuming that adverbs are initially generated to the left of the verb (see chapter 4) they are expected to move rightward, adjoining to the right edge of the verb. Under this analysis, the correct result is obtained in (68a), but not in (68b), where the adverb lands instead at the right of the clitic (68c). Since a clitic may appear between a verb and another clitic, that is, another clitic may adjoin to a clitic (temporal adjunction), as exemplified in (69), it must be the case that the ill-formedness of the (temporal) adjunction of the adverb to the imperative verb is due to the fact that fusion of the imperative verb and its clitic occurs in a structure like (67b), instead of the one in (67a).

- (68) a.    Regarde souvent Marie!  
           Look often (at) Marie.  
       b.    \*Regarde souvent le!  
           Look often (at) him.  
       c.    Regarde-le souvent!  
           Look (at) him often.

- (69)        Il=le connaît.

He knows him.

A further example comes from *ne*-less dialects. The negation marker *pas* appears between the imperative verb and the NP argument, as illustrated in (70a). But when the argument is a clitic, the negation morpheme appears to the right of the clitic (70b), and not between the clitic and the verb as in the ill-formed (70c). That is, fusion of the clitic with the imperative verb occurs under (67b), and repels the adjunction of the negation marker to the verb. We interpret this as evidence that fusion is stronger when the verb and the argument are dominated by the same node.

- (70)    a.    Regarde=*pas* Max!  
              Look (at) not Max.  
              b.    Regarde le=*pas*!  
                      Look (at) him not.  
              c.    \*Regarde=*pas* le!  
                      Look not (at) him.

More generally, I take these facts to be the reason why fusion generally does not occur in the complement domain of the verb. In effect, it may be that when fusion occurs under (67b), independent morphemes are so inseparable that they become affix-like elements in the syntax. Under this view, the OCP triggers the movement of the argument outside the complement domain of the verb in order to maintain the autonomy of the verb with respect to the Case-specified argument. Thus the verb may be a landing site for other elements, which may therefore avoid a longer movement.

To conclude, it is suggested that the OCP (in both of its manifestations) functions as a constraint favoring an economy of representation. Indeed, given a syntactic structure, it allows the number of time-intervals to be reduced by temporal adjunction. It significantly reduces the number of Case occurrences in a given tier by a process of fusion. Further, it is suggested that the OCP induces movement out of the complement domain of the verb in order to prevent the concatenation of independent morphemes in the syntax.

## 2.4 Transformation ordering and Clitic clusters

Naturally, the question arises as to whether or not transformations are ordered. In the data discussed so far, this does not matter since there is no conflict between OCP-1 and OCP-2, as just discussed. More generally, as the discussion proceeds, it will become clear that the ordering of transformations is not necessary in the present perspective. A theory which explicitly ranks all transformations is more complicated than a theory which does not. To repeat my assumptions (see chapter 1), I take all transformations (movement and fusion) to apply without explicit ordering. Let us assume furthermore that in order for a structure to be well-formed, a maximal number of constraints must be satisfied, or to put it another way, constraint violations must be minimal.

Now, when there are two clitics in the complement domain of the verb, the question arises as to which one must move first, as clitics are not freely ordered at the output. I will show, in fact, that the linear order is due to the claim that a maximal number of constraints must be satisfied for the derivation to be successful, and not to the ranking of the constraints involved. This analysis will make it possible to account in a simple fashion for the conundrum of dative inversion, and the well-known restrictions on clitic cooccurrences.

### 2.4.1 Clitic order

When several clitics appear in the complement domain of the verb, we must know the initial position of each of them. Since clitics do not move in imperative sentences, one may use this mood to diagnose their initial position. Consider an imperative verb with two object clitics, as illustrated in (71). As can be seen, the direct object clitic always precedes the indirect object clitic in standard French, an order which parallels that of nonclitic arguments, as exemplified in (72). Given this observation, we may assume that in standard French, the accusative (direct object) clitic is computed before the oblique one (indirect object) (but see the discussion below).<sup>26</sup>

---

<sup>26</sup> This is not the case in other standard Romance languages such as Italian, Portuguese, and Spanish, in which the indirect object clitic always precedes the direct object clitic. In these languages,



- (71) a. Donne-le-moi!  
Give it me.  
' Give it to me.'
- b. Montre-le-lui!  
Show it him.  
' Show it to him.'
- c. Raconte-le moi!  
Tell it me.  
' Tell it to me.'
- (72) a. Marie donne le livre à Paul.  
Marie gives the book to Paul.
- b. ?Marie donne à Paul le livre.  
Marie gives to Paul the book.

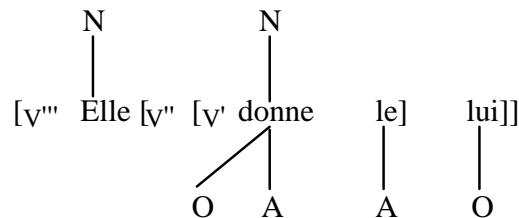
If this view is right, the derivation of a simple sentence such as (73) must proceed from structure (74).

- (73) Elle le lui donne.  
' She gives it to him.'

---

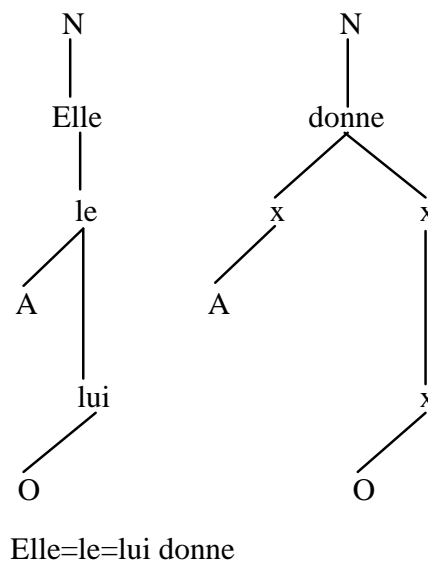
the order of the complement may be explained by the Animacy Hierarchy, as discussed in chapter 4 with Spanish data. In Spanish, verbs and indirect object clitics are specified for the animacy feature [animate]. French is different from these languages in that the verb is not specified for any animacy features (see section 2.4.2.2). It seems that French exploits the unspecification for animacy by computing the direct object before the indirect object in most clusters, a move which allows the language to avoid use of further repair strategies, hence the number of transformations (see chapters 3 and 4). When violation of a constraint cannot be avoided by any initial order (DO > IO or IO > DO), clitic order is somewhat free. Thus, in imperative verbs, one has either *donnez-moi-le!* or *donnez-le-moi!*, because the delinking process cannot be avoided. In the first case, it must take place for two reasons: the second clitic cannot be licensed and the nominative Case of the verb clashes with the accusative Case of the second clitic; in the second case, the two clitics are licensed, but the extra Case of the second clitic (accusative) conflicts with the nominative Case of the imperative verb. Notice, however, that prescriptive grammarians prohibit the order IO > DO. This relative freedom of ordering is possibly related to the violation of many constraints in imperative verbs. For instance, in a cluster of two third person clitics, the accusative clitic always precedes the other clitic: *donnez-le-lui!* vs. *\*donnez-lui-le!*. (But see Terzi, 1999, for another view.)

(74)



In the inner domain, the  $V'$ , there are two occurrences of accusative Case, and in the next domain, the  $V''$ , two occurrences of oblique Case. It is thus expected that the clitics move outside their respective domains in order to satisfy the OCP. Now since both clitics target the same landing site, it is a matter of knowing which one moves first. Given that the phrase-structure tree by its very nature reflects certain locality conditions, it is reasonable to assume (in the absence of any other constraints) that the clitic which is nearer to the landing site moves first. Thus, the accusative clitic must adjoin to the subject *elle*, while the oblique clitic is adjoined to the accusative clitic, yielding the structure in (75) (fusion is omitted for simplicity).

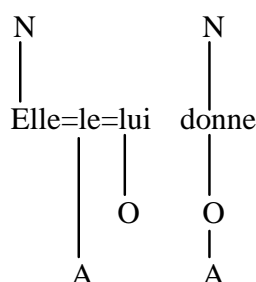
(75)



A question related to this state of affairs is the following: what happens to the tiers after the application of the Levelling Convention to the structure? I assume the null hypothesis that there is no deletion of features; thus each element in the

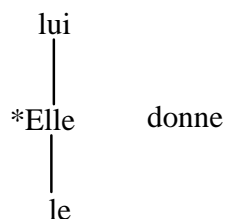
preceding structure keeps its features, as shown in (76). Recall that the symbol '=' indicates the adjunction relation, and that the head-adjunct complex forms a single time-interval. (For clarity, each element appears with its feature, instead of the single feature due to fusion, and dummy x-nodes are omitted.)

(76)



It is important to note that in the above example, the oblique clitic must obligatorily adjoin to the accusative clitic: multiple adjunction to the same head is excluded by the Levelling Convention. If both clitics are adjoined to the subject, the resulting structure, given in (77), may not be spelled out. For even though the three clitics occupy the same time-interval, the Levelling Convention cannot successfully linearize the structure. As can be seen, from the initial element, *elle*, the LC cannot apply in a continuous fashion, as it must first process either of the clitics and return to the other before processing the verb. Under double adjunction to a single head, one would obtain two zigzag paths, yielding the structures in (78). In addition, a further stipulation would be required to rule out (78a), but not (78b).

(77)

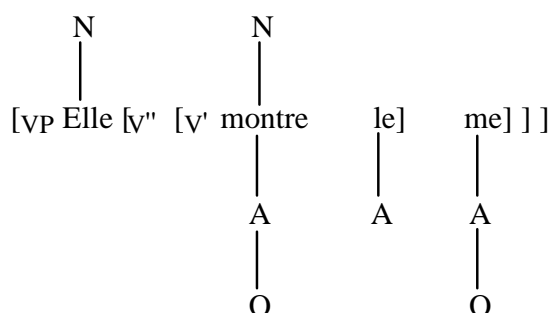


- (78)
- a. \*Elle=lui=le donne.
  - b. Elle=le=lui donne.

Consider now the case where the indirect object is a first (or second) person clitic, as in (79). This structure significantly differs from the previous one in that the indirect object clitic precedes the direct object clitic. There is no reason to postulate that the clitics in (79) are initially generated in a different linear order from those in (73). Thus, the initial structure of (79a), for instance, must be as shown in (80). The initial order of the clitics is the same in both (74) and (80). However the oblique clitic in (80) is intrinsically different from the one in (74). Indeed, while in the latter the third person indirect object pronoun is specified for a single Case (oblique), in the former the first person pronoun is specified for both accusative and oblique Cases. Thus, these two configurations being different, it is expected (and highly desirable) that they yield a different linear order in the output.

- (79) a. Elle me le montre. / \*Elle le me montre.  
           She me.DAT him.ACC shows.  
       b. Il vous le donne. / \*Il le vous donne.  
           He you-DAT he-ACC gives.

(80)

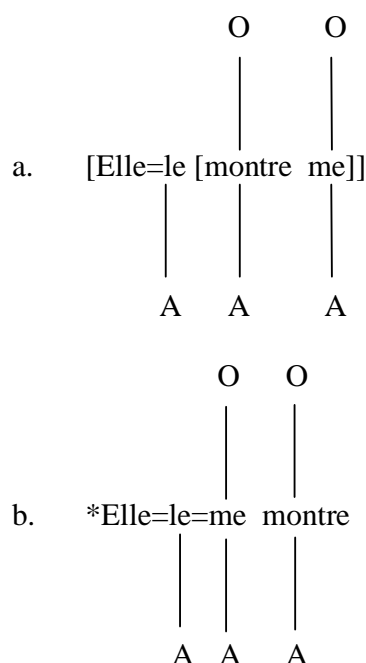


This type of representation allows the difference between the structures (80) and (74) to be captured. Since in (80) the clitics are not equal in features, the hypothesis that the most embedded clitic moves first under some locality condition cannot hold. Indeed, if the single-Case clitic moves before the double-Case clitic, one obtains an ill-formed structure, as exemplified in (81).

- (81)           \*Elle=le=me montre.

As can be seen in (80), within the complement domain of the verb, there are three occurrences of accusative Case and two occurrences of oblique Case. If the first clitic moves first, the movement will be string-vacuous in that it does not modify the structural description. Indeed, the OCP violation would remain in both the oblique and accusative tiers after the movement of the single-Case clitic, as shown by the derivational step in (82a). Movement of the second clitic leads to the ill-formed structure (81), see (82b). (Nominative Case, which is not relevant to the discussion, is set aside here and below.)

(82)



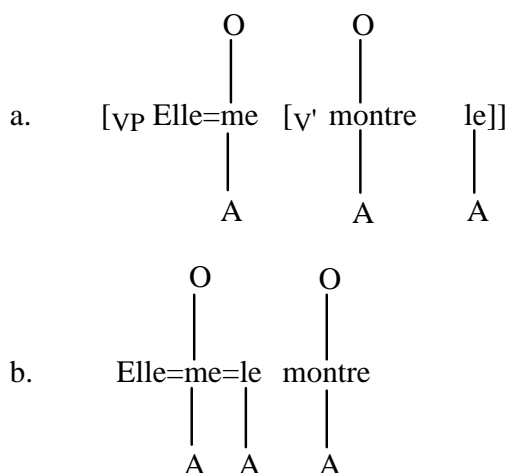
The ill-formedness of the derivation in (82) is thus due to a type of string-vacuous movement. Given that this sort of movement is not allowed, it must be the case that the grammar requires that the more offending element for a given constraint moves first. Let us refer to this requirement as the *Worst Evil (moves) First Constraint* (WEFC), where the term *worst evil* refers to the more offending element.

(83) Worst Evil First Constraint

Given a structure where two elements independently violate a constraint, the more offending element must move before the other.

With this constraint in mind, consider the structure in (80). The double-Case clitic is the worst evil, and hence must move first, yielding the intermediate step in (84a). Then the other clitic is adjoined to it, yielding the desired structure in (84b).

(84)



Below, I will provide additional evidence for the WEFC. I will also show that there are other constraints which intervene in the above structure; they are not discussed now because they do not play a crucial role, as they are satisfied independently and taken for granted.

We may note at this point that in the derivation in (84) no constraint is violated. Since the constraints are not ranked, a violation of a constraint must yield an ill-formed structure, unless it is repaired. In this respect, it is important to notice that the movement of the clitic under the locality condition, as discussed in (74), should not be taken to be a constraint. That is, there is no constraint stating that in the configuration in (74) the first clitic must move first. At best, the configuration (74) may be seen as a default case, where there is no constraint conflict. Indeed, the OCP is violated independently in each domain of the verb, and it is repaired independently in each domain by movement.

In sum, in the analysis presented so far the ordering of transformations may not be stipulated. When two elements equal in features compete for the same adjunction site, the one which is closer to the target moves first. However, when one of the

clitics gives rise to a double violation of the OCP, it must move first. We will see later that when each element targets a different landing site, the question of which one moves first is irrelevant.

#### 2.4.2 Clitic cooccurrence restrictions

We have seen that third-person object pronouns are specified for a single Case feature, which is either accusative (*le, les*), or oblique-benefactive (*lui, leur*), while first and second person pronouns are specified for both accusative and oblique Cases. Thus, with a simple transitive verb a double-Case pronoun may be either the direct object or the indirect object depending upon whether the verb is specified [A] or [O,β]; see (85). Similarly, with a ditransitive verb, the double-Case pronoun may be either the direct object or the indirect object, (86). However, in ditransitive verbs both arguments cannot be doubly specified for Case (87a). Moreover, a double-Case pronoun may not cooccur with a third person oblique clitic, (87b) (see Perlmutter, 1971; Kayne, 1976; Bonet, 1991).

- (85) a. Marie=me voit.  
Marie sees me.  
b. Marie=me nuit.  
Marie disturbs me.
- (86) a. Marie me présentera à Paul.  
Marie will present me to Paul.  
b. Marie me donnera un livre.  
Marie will give me a book.
- (87) a. \*Marie me vous avait recommandé.  
Marie recommended me to you.  
\*Marie vous m' avait recommandé.  
Marie recommended you to me.  
b. \*Marie me lui présentera.  
Marie will introduce me to him.

Bonet (1991) attempts to account for the facts in (86) and (87) by identifying two constraints, which she refers to as the *me-lui/I-II Constraint*. They are given in (88) and (89) respectively.<sup>27</sup> Bonet points out that (88) is universal and exceptionless, while (89) may vary within a given language from speaker to speaker with various degrees of acceptability.

- (88) The \**me-lui* Constraint (universal) (Bonet, 1991: 181)
- a. In a combination of a direct object and an indirect object, if there is one third person, it has to be the direct object.
  - b. Both the direct object and the indirect object are phonologically weak.
- (89) The \*I/II Constraint (Bonet, 1991: 181)
- a. In a combination of a direct object and an indirect object, the two objects cannot be first and second person.
  - b. Both the direct object and the indirect object are phonologically weak.

The constraints (88) and (89) are only descriptive, and explain nothing in themselves, as Bonet acknowledges. I will attempt to account for these constraints in a principled way. Specifically, I will show that the *me-lui/I-II* effects are due to the Animacy Hierarchy. The version of the Animacy Hierarchy (AH) which I will assume is given in (90) below.<sup>28</sup> In French, other constraints, in particular the

---

<sup>27</sup> In fact, Bonet suggests that the combination of a DO and an IO is governed by a single constraint, which has a strong version and a weak version, as given in (i) and (ii) respectively (Bonet, 1991: 182).

(i) The direct object has to be third person. (Strong version)

(ii) If there is a third person, it has to be the direct object. (Weak version)

For ease of exposition, I adopt the formulation in (88) and (89) in the present discussion.

<sup>28</sup> Notice that the Animacy Hierarchy refers to features, and not to any grammatical category. It is a well-known constraint which is found in other frameworks, e.g. functional grammar (see Dahl and Fraurud, 1996, for a discussion and further references). To the best of my knowledge, this constraint is not used in the GB (and MP) framework, even though this theory recognizes that certain processes, e.g. agreement, are triggered by animate feature. Nevertheless, GB uses an analogous hierarchy, the thematic hierarchy, which ranks thematic roles as follows: Agent > Theme > Goal > etc. Many researchers have proposed their own hierarchy of thematic features (see Larson, 1988; Speas, 1990; and Baker, 1997, for some discussion). In the OT framework Grimshaw (1997b) refers to a Universal



licensing of clitics (see (53) of chapter I and below), also intervene in the structures displaying me-lui/I-II effects, but I will show that they are not crucial.

- (90)           The Animacy Hierarchy (AH)  
                   In the complement domain of the verb, an argument  
                   specified as animate must precede another one which is  
                   specified as inanimate or as a thing.

#### 2.4.2.1 Clitic licensing

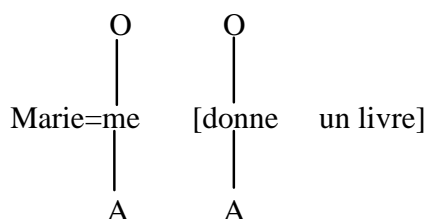
Having already discussed the syntax of sentences such as (85), which have a simple clitic as argument, I turn to (86), in which one of the internal arguments is an NP and the other a clitic. The derivation of (86b) is given in (91) below. The clitic being in the complement domain of the verb, it normally moves under OCP, as shown in (91b). Fusion also takes place, yielding a single Case occurrence on each tier, although for clarity it is not indicated in the schema. As can be seen in (91b), both Cases of the clitic fuse with the Cases of the verb, which is therefore saturated. Thus the NP *un livre* cannot be assigned Case by the verb. The structure is nevertheless well-formed, for, I argue, the NP is licensed by the lexical structure of the verb. Indeed, since the lexical structure of this verb requires two internal arguments, it is not crucial for the NP to be assigned Case by spreading. In other words, Case-marking of an NP is not a necessary condition.

- (91)   a.
- 
- b.

---

Markedness Hierarchy, given in (i), which includes a person hierarchy and a Case hierarchy. Both may be derived from the Animacy Hierarchy, as I will show in chapter 4.

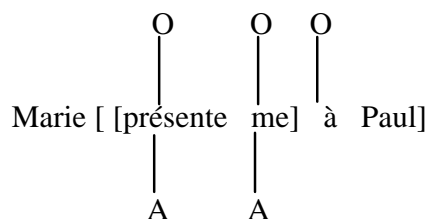
- (i)       \*2 > \*1 > \*3.  
           Dative > Accusative.



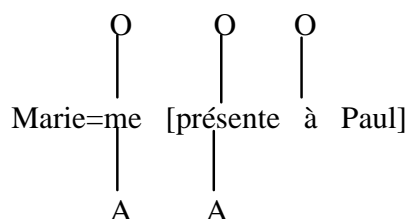
Notice that the initial order of the clitic and the NP is not crucial in (91a), as it has no effect on the output. Whether or not the clitic is in the inner domain of the verb, the NP will be interpreted as the direct object, for if it were the indirect object (oblique-benefactive) it would agree with the verb via the preposition  $\grave{a}$ .

Indeed in (86a), the indirect object NP is introduced by the preposition  $\grave{a}$ , expressing Case and thematic agreement as discussed in chapter 1. Here again, the clitic saturates both the accusative and the oblique Cases of the verb, as exemplified in (92a). But contrary to the previous example, the NP argument (the indirect object) is assigned Case by the preposition  $\grave{a}$ . The  $\grave{a}$ -phrase is reanalysed as a PP, since the preposition is no longer licensed by the verb, the relevant Case and thematic features of the latter being saturated by the clitic. Under the OCP, the double-Case clitic moves outside the complement domain of the verb and adjoins to the subject, as illustrated in (92b). Thus, it appears that if the clitic is doubly specified for Case, it absorbs both Cases of the verb, and the NP is licensed not by Case but by the argument structure of the verb. If the NP is introduced by the morpheme  $\grave{a}$ , in which case it is the indirect object, the  $\grave{a}$ -phrase is reanalysed as a PP.

(92) a.



b.

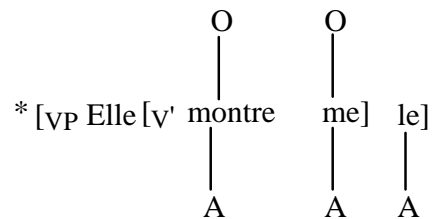


While NPs may be licensed by the argument structure of the verb, in my view a clitic must be obligatorily licensed by Case, i.e. the verb which selects the clitic must have a relevant and valid Case. This in fact follows from the Constraint on Merge and the local asymmetrical head-complement relation ( $r_{\text{compl}}$ ), as discussed in the previous chapter. I will refer to this as the Clitic Licensing Constraint, a preliminary version of which is given in (93).

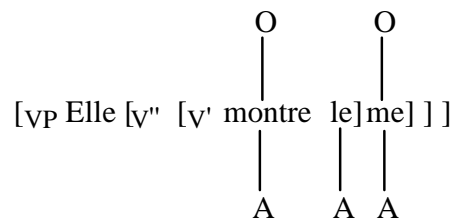
- (93) Clitic licensing Constraint (to be modified)  
 A clitic is licensed in a structure iff (a) and (b) hold:
- a. The clitic is specified for the same Case as the verb; and
  - b. the Case of the verb is not saturated by another Case-specified element.

Given (93), it is expected that a double-Case clitic completely saturates the verb, preventing another clitic from being licensed. The structure (94a) would thus be ruled out by virtue of (93), since the clitic *me* saturates both Cases of the verb. In fact, this configuration may not be generated as such, for the initial order of the complements in French is direct object > indirect object, as suggested above (but see below). That is, the direct object clitic *le* must be computed before the indirect object clitic *me*, as shown in (94b) (already discussed in (80)). In the configuration (94b), the first clitic saturates only the accusative Case of the verb, which may thus license the second clitic with its oblique Case. In languages such as Italian and Spanish, where the indirect object clitic is computed before the direct object clitic, a configuration such as (94a) yields an ill-formed structure, which triggers a repair strategy. For instance, Spanish spurious *se* is the consequence of repairing structures of the type (94a), as I will show in chapter 4.

(94) a.



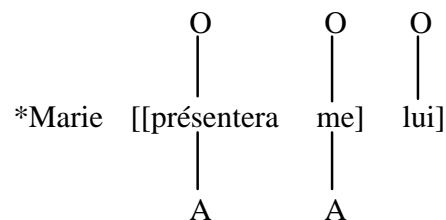
b.



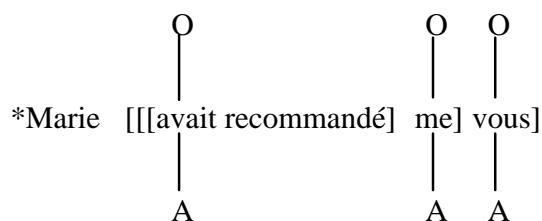
Summing up, clitics are licensed differently from NPs. If an NP fails to be assigned Case, it can still be licensed by the lexical structure of the verb. But a pronominal clitic necessarily requires that the verb bear a relevant and valid Case (i.e., a Case which is not already taken by another clitic).

Now, if the double-Case clitic is the direct object, under (93) it is predicted that another clitic may not be licensed by the verb. Indeed, the structures in (87) are ill-formed precisely because the direct object, which is computed before the indirect object, saturates both Cases of the verb, as shown in (95a) and (95b), from which the ill-formed structures in (87a) and (87b) are derived, respectively.

(95) a.



b.



In fact, the structures in (95) are not crucially ruled out because the second clitic is not licensed. If this were the case, one would expect the grammar to use a repair strategy to circumvent the problem. We discussed above a type of repair strategy, where the offending features of the clitic are deleted up to Structure Preservation. If this transformation were possible here, one would obtain for a structure like (95b) the output shown in (96), where the Case features of the first clitic have been delinked. Since (96) is ill-formed, it is clear that feature manipulation does not apply here.<sup>29,30</sup> I will indeed argue that another constraint is crucially responsible for the ill-formedness of these structures. Such a constraint is the Animacy Hierarchy, as defined in (90) above. The next subsection is devoted to this goal.

<sup>29</sup> I argue further on that Italian uses such a strategy. In this language the indirect object clitic, which is doubly specified for Case, is computed first. Thus it always saturates the verb, and the NP argument is licensed by the argument structure of the verb. When the second argument (the direct object) is another clitic, the features of the doubly specified clitic must be deleted up to Structure Preservation, which yields the strong form in a clitic cluster. Thus, in (i) the weak form of the first person pronoun is used, while in (ii) the strong form of this pronoun is used.

- (i) **Mi** regalò il libro.  
Pro.1.sg.A.O gives the book.  
' He gave me the book.'
- (ii) **Me/\*mi** lo regalò.  
Pro.1.sg. Pro.3.sg.A. gives  
' He gave it to me.'

<sup>30</sup> According to Bonet (1991), the delinking strategy is adopted in Catalan in order to avoid the effect of the *me lui*/I-II constraint. In Catalan, Bonet argues, a locative clitic is obtained from the dative pronoun by feature manipulation, that is by delinking process. But in French, she suggests that another strategy is used, which she dubs the *Spell-Out Elsewhere*. That is, given that in her analysis a clitic, which is in Infl, is coindexed with a phrase in argument position, in order to avoid the *me lui*/I-II Constraint, one of the clitics is not spelled out; instead the coindexed phrase is spelled out. Under this view, the grammar generates both clitics, but the second one is not spelled out, as illustrated in (i).

- (i) Marie m<sub>j</sub>' (\*vo<sub>u</sub><sub>j</sub>) a recommandé (\*mo<sub>i</sub><sub>j</sub>) [à vous<sub>i</sub>]

Unfortunately, the question as to why Catalan uses the delinking process strategy and French the Spelled-Out Elsewhere strategy does not matter for Bonet, who summarily attributes it to a parametric variation.

(96) \*Marie **moi** lui avait recommandé.

To conclude, clitics and NPs are licensed differently. Both of them saturate the Case of the verb, as well as its argument structure. But while an NP may fail to be assigned Case by the verb, and still be a legitimate object in the structure, a clitic must obligatorily saturate the Case of the verb in the structure it appears in. It follows from this that if an NP and a clitic compete for the same Case of the verb, the verb will be saturated by the clitic and not by the NP, as shown above. Also, two clitics may not compete for the same Case of the verb; a structure containing more clitics than the verb has Cases cannot be grammatical.

#### 2.4.2.2 Animacy effects

First of all, I will attempt to motivate the Animacy Hierarchy, and I will show how it correctly excludes the configurations (95).

Suppose that UG allows some elements to be specified for the feature [animate] or its contrary, which we may call [inanimate] or [thing]. The latter would be equivalent to the feature [-animate] in an equipollent theory of feature specification. In the paradigm of personal pronouns, the elements which are likely to be specified for the feature [animate] are first and second person pronouns, which refer to human. As a matter of fact, let us assume that these pronouns are specified for the feature [animate]. On the other hand, third person pronouns may refer to either feature, animate or inanimate. Indeed, many languages have different morphemes for animate and inanimate in the third person. For instance, subject pronouns in English display an opposition between 'thingness' and 'humanness' in third person singular (*it* vs. *he*). In French while the subject pronouns *il* 'he' and *elle* 'she' and the object pronoun *le* 'him' and *la* 'her' may refer either to animate beings (including human) or things, the object pronouns *lui* (indirect object) may refer mostly (perhaps only) to animate objects. Suppose then that indirect object pronouns such as *lui/leur* are specified for [animate], and the direct object pronouns are unspecified for animacy features. That is, they are specified neither as [animate] nor as [thing].

Now given the Animacy Hierarchy (see (90)), it is expected that in a pronoun system which includes both the features [animate] and [thing], an element specified for the feature [animate] will always precede an element specified for the feature [thing]. But if a pronoun system contains an opposition between the feature [animate] and nothing, it follows that the Animacy Hierarchy may not apply. That is, the linear order of the two elements may not be governed by the Animacy Hierarchy. These two possibilities are singled out in (97) for clarity (where 0 stands for no feature). ((97b) means that the order is either  $0 > \text{animate}$  or  $\text{animate} > 0$ .)

- (97) a. [animate] > [thing]  
 b.  $0 < \text{animate} < 0$

The strongest evidence for this analysis comes from the linear ordering of internal arguments in French and in other standard Romance languages. With respect to NPs, French patterns like other Romance languages: NP order is usually DO>IO; the reverse order, which is not systematically bad, is less natural, as exemplified in (98). These facts are consistent with the Animacy Hierarchy (97). With NP arguments, the order is somewhat free, because such elements are neither specified for [thing] nor for [animate].<sup>31</sup>

- (98) a. Paul donne le livre à Marie.  
           'P aul gives the book to M<sup>arie</sup>.'  
 b. ?Paul donne à Marie le livre.  
       Paul gives to Marie the book.'

Similarly, if both complements are clitics, their ordering is not absolute: either the indirect object precedes the direct object or the indirect object precedes the direct object. As discussed above, the first alternative occurs when the indirect object clitic is not third person, while the second takes place if both the clitics are third person. For instance both the clusters *me-le* (IO>DO) and *le-lui* (DO>IO) are possible, as

---

<sup>31</sup> The unnaturalness of the order IO > DO may be due to a Case conflict. In this order the NP indirect object is assigned accusative Case by the verb and oblique Case by the preposition. On the other hand, in the order direct object > indirect object, the NP indirect object is assigned Case only by the preposition, because the NP direct object absorbs the accusative Case of the verb. See also the discussion of (34).

illustrated in (99) and (100) respectively (where  $\alpha$  stands for animate). These facts are consistent with (97) since in each possible cluster a single clitic is specified for animacy (*me* and *lui*). (Note that the ungrammaticality of the (b)-sentences follows from other independent facts, WEFC (99b) and the locality effect on movement (100b), as discussed above.)

- (99) a. Paul **me le** donne.  
           Paul CL.1.sg.O/(A). $\alpha$  CL.3.sg.A gives.  
       b. \*Paul **le me** donne.  
           'P aul gives it to me.'
- (100) a. Paul **le lui** donne.  
           Paul CL.3.sg.A CL.1.sg.O/(A). $\alpha$  gives.  
       b. \*Paul **lui le** donne.  
           'P aul gives it to him.'

Other Standard Romance languages such as Italian and Spanish contrast sharply with French with respect to the linear order of clitic arguments. Indeed in these languages, clitic arguments obligatorily show the order IO > DO. I tentatively argue that this is due to the fact that in these languages verbs are specified for the feature [animate]. The existence of this feature is more transparent in Spanish than in Italian, for example. That Spanish verbs are indeed specified for [animate] is supported by the fact that all animate NP objects must be introduced by the animacy morpheme *a*, see (101). Thus since the verb is specified as [animate] and one of the clitic arguments is also specified as [animate], it is clear that the animate clitic, i.e. the IO clitic, must be computed first, and then the DO clitic. This has a strong empirical basis: whatever the person feature involved in a clitic cluster in Spanish, the order is always IO > DO, as exemplified in (102).

- (101) a. Veo \*(a) Juan.  
           I see A Juan.  
       b. Pedro quiere \*(a) sus padres.  
           Pedro likes A his parents.



- (102) a. (Maria) me lo dio.  
           Maria Cl.1.sg.O.α Cl.3.sg.A.τ gave  
           ' (Maria) gave it to me.'
- b. (Maria) se lo dio.  
           CL.3.O.α Cl.3.sg.A.τ gave  
           ' (Maria) gave it to him.'

Thus on my view, French displays both orders, IO > DO and DO > IO, because the Animacy Hierarchy is irrelevant and the verb is not specified for animacy. As assumed above, in French the direct object clitic must be computed before the indirect object clitic: that is, at the initial stage of the derivation the order of the clitics is DO > IO. Why should this be the case? It may be that in the absence of the Animacy Hierarchy a language chooses the order which allows the correct output to be obtained with a minimal number of transformations. Given that in French most indirect object pronouns are doubly specified for Case, if the initial order were IO > DO, the CLC would be always violated, except when the indirect object is third person. When this constraint is violated, feature delinking must take place, a process which significantly increases the complexity of the structure (see chapter 4). On the other hand, with the sequence DO > IO the CLC is never violated, and the correct linear order is directly derived by movement (under OCP), with the proviso of the WEFC. To conclude, in French, clitics are in most cases generated in the order V > DO > IO, which allows the correct output to be derived with a smaller number of transformations.<sup>32,33</sup>

Having motivated the Animacy Hierarchy, let us turn now to see how it rules out structures displaying the me-lui/I-II effects such as (87) repeated in (103) below for convenience. Since the AH requires an opposition between the features [animate] and [thing], it is reasonable to assume that it is a general constraint on structure interpretation, which discriminates arguments. Under this view a structure where both

---

<sup>32</sup> Notice that the WEFC does not count in the choice of the order of the argument at the initial stage since it always applies, independently of the Animacy Hierarchy.

<sup>33</sup> In the next chapter, I will show that the clitic *en* is specified for [accusative] and [thing], and therefore when it is involved in a clitic cluster, it must be ordered after the animate clitic, consistently with the AH.

arguments are specified for [animate] (or [thing]) may not sound natural any more than a structure displaying two direct objects or two themes, as given in (104) (see also Burston, 1983; Laenzlinger, 1993).<sup>34</sup> Now in the ill-formed structures in (103), we find a configuration precisely of the type illustrated in (104a): both the arguments are specified for the feature [animate].

- (103) a. \*Marie me vous avait recommandé.  
           Marie recommended me to you.  
           \*Marie vous m' avait recommandé.  
           Marie recommended you to me.  
       b. \*Marie me lui présentera.  
           Marie will introduce me to him.
- (104) a. \*animate > animate.  
       b. \*thing > thing.

Consider now the repair strategy which is used to avoid the AH in *me-lui* and I-II contexts. As mentioned above, feature manipulation is not possible here. The reason is that the animacy feature, unlike Case-features, may not be deleted. Since the strong pronouns are also specified for animacy, given Structure Preservation, the result of the delinking operation cannot yield a form which lacks the feature animacy. If indeed one delinks the feature [animate] as well as any other Case feature in the

---

<sup>34</sup> Expanding on Burston (1983), Laenzlinger (1993) suggests that all weak pronouns are associated with an individuation feature, which has either the values [animate], [inanimate] or [Ø]. In his perspective, these feature values are 'determined either on the basis of the governing verb or on the basis of the context.' (p.256) He then proposes the constraint in (i) (his (35)), which is similar to the filter in (ii) presented in Burston (1983), although he attempts to derive it from Rizzi's (1990) Relativized Minimality.

(i) \*CL<sub>1</sub><sup>0</sup> CL<sub>1</sub><sup>0</sup> [XP e] [YP e]  
       where both clitics are lexically specified for the same individuation feature value, namely <animate>, <inanimate> or <Ø>.

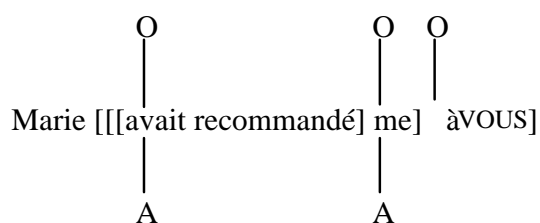
(ii) \* [+Individuation, +Individuation]

(ii) is equivalent to (104), although Burston does not correlate it with the AH. Bonet (1991) rejects Burston's account of the *me-lui*/I-II constraint (ii) because according to her it is too strong. She points out that a sequence of two I-clitics (in my view, animate) is possible if one of them is an ethical dative. Unfortunately, this is not a strong reason, for ethical datives possess a set of properties which are not shared by nonethical clitics. Hence the two types of clitics may not necessarily be accounted for with the same apparatus. I will not discuss ethical datives in this thesis, but according to some preliminary observations, they are licensed by the latent Case of the verb, as discussed in chapter 3.

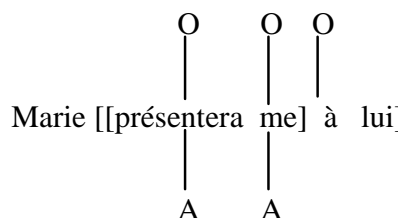
examples above, the remaining features would not correspond to any pronoun in the paradigm.

I would like to suggest that in structures such as (103) the grammar does not in fact generate two clitics. This is obviously the minimal assumption since there is no delinking process here as discussed above. Indeed, instead of the second clitic a pronominal *à*-phrase is generated, as illustrated in (105). (This claim contrasts with the analysis discussed in Bonet (1991), who suggests that the output structure is derived from the ill-formed structure not by feature manipulation, but by another mechanism she refers to as *Spell Out Elsewhere* (see footnote 31)). Now, the AH conflicts with the Blocking Principle. At the initial step the latter must rule out the structure (105) in which a Caseless form is selected over a Case-specified pronoun. However, since the AH is an interpretive constraint, it wins out, consistently with my assumptions. Furthermore, since the strong pronouns are specified for animacy, one must explain why the AH does not rule out these structures, unlike the sequence of two clitics. Under the assumption that animacy effects occur only when both internal arguments are of the same nature, i.e. they are categorially identical, the structures in (105) do not violate the AH. In other words, a sequence of two Case-specified pronouns may violate the AH not because they are phonologically weak, as stated by Bonet, but because they belong to the same syntactic category. Notice that the clitics move normally, yielding the well-formed structures in (106).

(105) a.



b.



(106) a. Marie=m' avait recommandé àVOUS.

- b. Marie=me présentera à LUI.

This analysis predicts that if both internal arguments are of the type NP, the structure may not be perfect. This is borne out. In Spanish, all animate direct objects must be introduced by the animacy marker *a* (see chapter 4). In ditransitive verbs, the direct object NP is not introduced by *a*, as illustrated in (107).

- (107) a. Recomendaron Paul a Maria.  
 b. \*Recomendaron a Paul a Maria.  
 They recommended Paul to Maria.

To summarize, the well-formedness of a clitic cluster is regulated by several constraints, which are: the Clitic Licensing Constraint, the Animacy Hierarchy, the Worst Evil First Constraint. The Animacy Hierarchy is a semantic-like constraint which overrides the effect of the Blocking Principle, and it rules out any structure in which both internal arguments are specified for the same animacy feature. The Clitic Licensing Constraint and the Worst Evil First Constraint do not interfere with the Animacy Hierarchy, and thus always apply. The former is not violated in the French data discussed so far, since the single Case clitic (direct object) is computed before the double Case clitic (generally the indirect object). The WEFC forces the double Case clitic to be ordered before the single Case clitic. However, if both the direct object and the indirect object clitic are specified for a single Case, the former is ordered before the latter, by virtue of a locality effect. It is important to note (and to recall) that the word 'first' in Worst Evil First Constraint is not intended to mean that this constraint must be ranked before any other constraint. Indeed, this constraint refers to a configuration with two clitics, one of which violates the OCP on two different Case tiers. The clitic which yields the violation of the OCP twice is the worst evil and hence must move before the other clitic. It is not in conflict with any other constraint, since the locality effect in syntactic structure may not be taken to be a constraint; it is rather a default situation where there is no constraint in action (except of course the OCP).

## 2.5 Summary

In my view a pronominal element is a clitic if it is specified for Case. The cliticization process is the result of the OCP, which moves the Case-specified argument outside the complement domain of the verb. The movement is not absolute, and occurs if the structure contains a possible adjunction site for the clitic. Indeed, in imperative verbs there is no movement under OCP because there is no element to the left of the verb, which may serve as an adjunction site. The OCP has a second effect on Case-specified arguments: feature fusion. Whether the clitic has been moved or not, its Case feature must fuse with that of the verb, yielding a configuration where a single Case occurrence appears on a given tier.

It has been shown that these two effects of the OCP allow the various defining properties of clitics to be accounted for in a principled way. Specifically, the fact that clitics must be adjacent to their head follows from Case fusion. Under the analysis of adjunction I proposed, where the adjunct and the head are in a temporal relationship, it is explained why another clitic may intervene between a clitic and a verb.

Furthermore I have attempted to account for clitic clusters in a principled way. It has been shown that the ordering of clitics in the output depends both upon the Case feature of each clitic and the order of the computation. When both clitics are third person, a locality condition allows the most embedded clitic, which is closer to the target, to move first. I have argued that this locality condition on movement is not a constraint, but rather a default case which arises when the elements to be moved under OCP are equal in features. With a cluster of two third person clitics, which are specified as oblique and accusative respectively, the accusative clitic which is more embedded moves first under the locality condition. On the other hand in a cluster of a third person clitic and a first or second person clitic, the latter moves first because it is doubly specified for Case. The Constraint which ensures this outcome is referred to as the Worst Evil (moves) First.

The cooccurrence restrictions on clitic clusters has also been discussed. It has been argued that the so-called *me-lui* /I-II Constraints (Bonet, 1991; Perlmutter, 1971) are the result of the Animacy Hierarchy, which rules out a sequence of two

clitics which are specified for [animate]. In order to avoid a violation of the Animacy Hierarchy, it is argued that the grammar generates the second clitic as an *à*-phrase.

This chapter has mainly discussed French data, although not exhaustively, for other French clitics such as *en*, *on*, *se*, will be discussed in the next chapter. Given that no additional theses will be required to account for the latter, they will serve further as evidence for the analysis presented in this chapter. In chapter 4, I will present cross-linguistic evidence for the OCP, showing that the behaviour of pronouns in English, Haitian Creole, Italian and Spanish follows from the analysis presented above. For instance, it will be shown that the fact that the object pronouns in English appear postverbally is due to their lack of Case specification.

Finally, we may note that the formal apparatus which underlies the analysis is that which is prevalent in generative phonology. A grammar which uses the same set of operations and well-formedness conditions in two different components is clearly more highly valued than one which uses different types of devices, given the Relativized Modularity Hypothesis of Di Sciullo (1996). Therefore the analysis of clitics presented in this chapter is preferable on both theoretical and empirical grounds.

## APPENDIX TO CHAPTER II

### AUXILIARIES

There are two main auxiliary verbs in French: *être* 'to be' and *avoir* 'to have'. These verbs also have lexical uses, exactly like their English cognates. *Avoir* in its lexical use is strictly a transitive verb, taking a subject and a direct object; it may not have two objects, nor can it stand without one, as shown in (108). *Être*, on the other hand, is mainly a copula verb, but it is also used to express possession and location, as illustrated in (109).

- (108) a. Marie a une maison.  
          Marie has a house.  
      b. \*J' ai une maison à Marie.  
          I have a house of Marie's .  
      c. \*J' ai.  
          I have.
- (109) a. Je suis étudiant.  
          I am a student.  
      b. Le livre est à Jean.  
          The book is Jean' s.  
      c. Jean est à Paris.  
          Jean is in Paris.

In (108) and (109) the verbs *avoir* and *être* assign nominative Case to their respective subjects. This is confirmed by the fact that the Blocking Principle excludes

a non-nominative pronoun in the subject position of these verbs (*\*Moi ai une maison* ' Me have a house' *\*moi suis étudiant* ' Me am a student' ). Like any DO verb *avoir* assigns accusative Case to its object in (108a). As can be seen in (109), the complement of the verb *être* can be an *à*-phrase or an adjective. Nevertheless, I argue that *être* is specified for accusative Case. That *être* is specified for accusative Case is supported by the fact that the argument in (109a) may be replaced by the neutral clitic *le*, as shown in (110a). Recall that if a verb does not have a valid (and relevant) Case, it may not license a clitic. Furthermore, the fact that *être* is compatible with the clitic *y*, see (110c), suggests that it is specified for accusative Case, and perhaps for an oblique Case as well, since there are two *y* morphemes, one of which is specified as accusative and the other as oblique (see chapter 3). Both *y* morphemes, as I will argue in the next chapter, bear the semantic features locative and goal, which are marked on NPs with homophonous prepositions, namely *à*. Thus, the fact that in (110b) *être* is incompatible with an oblique-benefactive clitic suggests that the preposition *à* is not the benefactive marker encountered above. It seems that the morpheme *à* in (109b) bears an alienable possessive feature, which is not related to any clitic in the paradigm (see also the discussion of inalienable possession in the next chapter).

- (110) a. Je le suis.  
I am it.  
b. \*Le livre lui est.  
the book him-DAT is.  
c. Jean y est.  
Jean is there.

The lexical uses of these verbs contrast sharply with their auxiliary uses with respect to argument structure. Indeed, contrary to its lexical counterparts, the auxiliary *avoir* is compatible with either an intransitive, a transitive, or a ditransitive verb, as shown in (111). The auxiliary *être* on the other hand may be used to form the passive of transitive verbs, as is well known. Since the verb *être* does not take an internal argument, it is expected that the verb it passivizes does not take an internal argument either. This is borne out, since passives are like intransitive verbs.



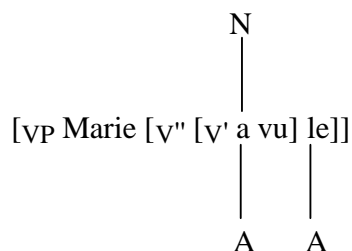
- (111) a. J' ai dormi.  
I have slept.  
b. Marie a vu Paul.  
Marie has seen Paul.  
c. Il a donné le livre à Paul.  
He has given the book to Paul.
- (112) a. Paul a été vu par Marie.  
Paul has been seen by Marie.  
b. Le livre a été donné à Paul  
The book has been given to Paul.

In light of these facts, it appears that *avoir* and *être* have the following properties: *avoir* is a transitive verb which is specified for nominative and accusative Cases; *être* does not take an internal NP, and hence it is not a transitive verb, leaving open the exact definition of its argument structure (but see Tremblay 1991). Used as an auxiliary, *avoir* may become intransitive, while *être* converts a lexical transitive verb into an intransitive verb (passivization). As a minimal assumption, I assume that *avoir* in its auxiliary use has the property of copying the argument structure of the lexical verb. I further assume that participial verbs bear no Case (at least in French). All these operations occur prior to the syntactic computation, that is, in the morphological component.

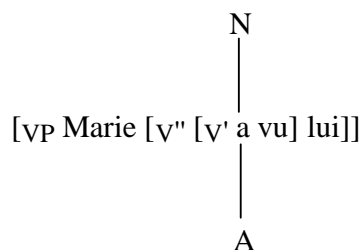
Now, since the participial verb cannot take an argument on its own, one may assume that it forms a complex verb with the auxiliary, following Rizzi (1986). As such, both are merged yielding a structure V' (= auxiliary+verb); the latter is then merged with the arguments, as discussed in the text. This is illustrated in (113a). The other alternative, that the participial phrase is the complement of the auxiliary, as illustrated in (113b), cannot be maintained. This is obvious when the internal argument is a pronoun. If this were the case, one would expect that a clitic would never appear as the complement of a compound tensed verb. The fact is, the participle not being Case-specified, it would select the unspecified form in the paradigm, consistently with the Constraints on Merge and the Blocking Principle. This is illustrated in (114).

- (113) a. [VP Marie [V'' [V' a vu] Paul]]  
 b. \*[VP Marie [V'' a [V' vu Paul]]]

- (114) a.



- b.



Turning now to *être*, as a copula, this verb does not change the argument structure of the adjective, in contrast to transitive verbs. Some adjectives behave like verbs to the extent that they take an indirect object (oblique-benefactive), as illustrated in (115a). One may assume that these adjectives are specified for oblique-benefactive Case. Thus, the *à*-phrase complement of the adjective may be replaced by a clitic as seen in (115b). Assuming that *être* is also specified for a latent oblique (benefactive) Case (see chapter 3), the clitic climbs to the left of the copula, outside its complement domain, as can be seen in (115b). Note that if the verb were not specified for oblique Case, the clitic would not move outside the complement domain; instead, it would adjoin to the verb (or perhaps to PRO if the latter were allowed here), yielding the ill-formed structure in (116), after linearization.

- (115) a. Marie [est [fidèle [à Paul]]]  
 Marie is faithful to Paul  
 b. Marie=lui [est fidèle]  
 Marie is faithful to him.

(116) \*Marie [est=lui fidèle]

To conclude, *avoir* and *être* are specified at least for the following Cases: [nominative] (if tensed), [accusative] and [oblique-benefactive], the latter being a latent Case. I assume that *avoir* has the property of copying the argument structure of the lexical verb, so that it may be the auxiliary of an intransitive or a ditransitive verb. I assume that participles are not specified for Case, and they form a complex verb with the auxiliary (cf. Rizzi, 1986). That is, in my view, the auxiliary is merged with the participle, and the complex thus formed is further merged with the arguments.

## CHAPTER III

### ADDITIONAL CLITICS IN FRENCH

#### 1. Introduction

In this chapter I pursue the analysis of the French pronoun system, focussing on the pronouns *en*, *on*, *se* and *y*. I will show that they obey the same basic principles discussed in the preceding chapters. In particular, all of them but *on* are shown to move from the complement domain of the verb they are argument of by virtue of the OCP. Like other object clitics, they adjoin to the first element outside this domain in order for fusion of their Case feature with the verb's Case to take place. As a nominative clitic, *on* remains in situ but it normally undergoes the fusion process.

The basic claim which underlies this chapter is the Argument Licensing Constraint. In the previous chapter, I show that a clitic must be licensed by a valid Case of its verb, a requirement I referred to as the Clitic Licensing Constraint. In fact Case-licensing appears to be only one condition on the licensing of clitics. Indeed it will be shown that an argument clitic must also be licensed by the lexical structure of a verb, an obvious claim since argument clitics are in complementary distribution with NPs, which are strictly licensed by a verb's lexical structure (see (1b)). In certain cases to be discussed below, a clitic can be licensed by the NP argument it refers to, as stated in (1a).

(1) Argument Licensing Constraint<sup>1</sup>

## a. Clitic licensing

A clitic is licensed (i) by a verb with a valid Case and (ii) either by the lexical structure of the verb or by coindexation with an argument of the verb.

## b. NP licensing

An NP is licensed by the lexical structure of a verb.

A unitary analysis of the clitic *se* will be provided. It is argued that whatever the type of construction it appears in, *se* is generated to the right of the verb selecting it, i.e., within the complement domain of the verb like any internal argument, and is moved outside this domain under OCP. It is argued that this clitic is specified for accusative and oblique-benefactive Cases, exactly like first and second person clitics (cf. *me /nous*, *te /vous*). Clitic *se* is different from the latter in that it is inherently reflexive. A lexical analysis of middle constructions will be proposed; and given (1) many properties of such constructions will be shown to follow straightforwardly.

The clitics *en* and *y*, as I will argue, may function both as pronouns and anaphors. As pronouns, they have a referent in the discourse, while as anaphors their referent is in the clause. The clitic *en* may replace a *de*-phrase which is either the complement of a transitive verb or the complement of an NP. As an argument, *en* saturates the Case of the verb, and occupies a slot in its argument structure, just like the clitics discussed in the preceding chapter. However, when it replaces a noun complement, it does not fill a slot in the argument structure of a verb; rather, it is licensed by the verb's Case and by coindexation with the head noun. As I will show, a clitic will be licensed by coindexation when it is coindexed with an NP which bears a grammatical function compatible with its Case.

With respect to features, it will be shown that *en* is specified for accusative Case, and bears the semantic feature associated with the preposition *de*; the latter may express various semantic roles, but only a subset of which are compatible with *en*.

---

<sup>1</sup> Note that the first part of (ii) is obvious, since it is not the case that any verb can license any clitic: cf. \**Paul le lui voit* 'Paul it.acc him.dat sees.'

The subset of semantic features which are compatible with *en* will be collectively referred to as delta ( $\delta$ ). As for *y*, it will be argued that it constitutes two separate morphemes with identical thematic roles, but with different Cases: one is specified as accusative and the other as oblique, each bearing the semantic roles goal ( $\gamma$ ) and location ( $\lambda$ ).

This chapter is organized as follows. The syntax of *se*, *on*, *en*, and *y* will be discussed sequentially. Some related facts, namely the double clitic problem, will be considered in section 3.6. In section 3.7 I discuss the hypothesis put forward in chapter 1, according to which most French verbs bear a latent oblique-benefactive Case. It is shown that this latent Case licenses non-argumental clitics, including extended datives. Section 3.8 summarizes this chapter.

### 3.2 The syntax of *se*

*Se* appears in various types of constructions in French: it is used to express reflexivity (2) and reciprocity (3). It is also used in middle constructions (4) and to form neuter verbs (5), also referred to as ergative verbs. It can also be an obligatory part of the argument structure of some verbs, as exemplified in (6).

- (2) a. Jean se regarde dans le miroir.  
Jean SE watches (himself) in the mirror.
- b. Paul se rase tous les deux jours.  
Paul SE shaves every two days.
- c. Les enfants se lavent dans la rivière.  
The kids SE wash in the river.
- (3) a. Jean et Marie se regardent.  
Jean and Marie SE watch.
- b. Jean et Marie se détestent l'un l' autre.  
Jean and Marie SE hate each other
- (4) a. Ce vin se boit frais.  
This wine SE drinks cold.

- b. Les légumes se vendent bien.  
The vegetables SE sell well.
- (5) a. La fenêtre s' est cassée.  
The window SE broke.
- b. La foule se disperse.  
The crowd SE disappears.
- (6) a. Jean s' est évanoui.  
Jean SE fainted.
- b. Elle s' imagine que Paul l' aime.  
She SE imagines that Paul likes her.

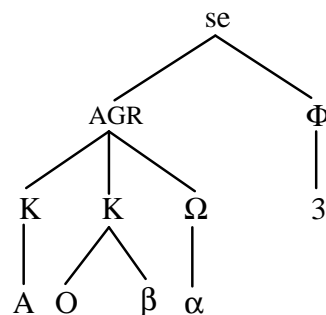
There is an important literature on French *se* in particular and Romance *se* in general (cf., among many others, Bouchard, 1995; Burzio, 1986; Di Sciullo, 1991; Everett, 1996; Grimshaw, 1982, 1991; Kayne, 1975; Lyons, 1995; Ruwet, 1972; Wehrli, 1986; Zribi-Hertz, 1982, 1987). Given the semantic diversity of *se* constructions, many researchers attempt to present a unified analysis of *se* in Romance. Most analyses also acknowledge the idea that *se* is a clitic, as it successfully passes the usual tests for clitichood (see Kayne, 1975).

I will attempt to propose a unitary account of the various types of *se*. However, I will set aside the semantic peculiarity inherent to some types of constructions involving this clitic. For instance middle-*se* is limited to generic or habitual contexts and therefore is incompatible with some tenses and certain types of adverbs; compare the middle construction in (7) and the ergative (neuter) construction in (8) (examples from Ruwet, 1972).

- (7) \*Ces lunettes se sont nettoyées hier à huit heures et quart.  
My glasses SE cleaned yesterday at quarter past eight.
- (8) Cette branche s' est cassée hier à huit heures et quart.  
This branch SE broken yesterday at quarter past eight.

To begin with, following Everett (1996), Lyons (1995) and Wehrli (1986), among others, I claim that there is a single morpheme *se* (in French). Since in its most transparent use as a reflexive it can be either the direct object or the indirect object of a verb, it must be the case that it is specified for both accusative and oblique-benefactive Cases, like first and second person object pronouns (see evidence below). *Se* shows neither number nor gender distinction, and therefore the sole  $\Phi$ -feature which it is likely specified for is (perhaps) third person. As for the  $\Omega$  node, it is probably empty, given that *se* acquires a reflexive reading only with animate referents.<sup>2</sup> The feature structure of *se* is given in (9). Notice that the tree structure displays no reflexive feature; the fact is that there is no evidence that such a feature ever interacts (in the relevant sense discussed above) with any feature of a verb. Therefore, I take the reflexivity of *se* to be marked on the root node, i.e., its defining property.

(9)



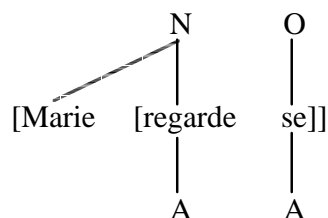
Given this feature structure, I take reflexive/reciprocal *se* to be generated within the complement domain of the verb of which it is an argument, whence it moves under the OCP, as discussed in the preceding chapter. This is illustrated in (10): in (10a) *se* is the argument of a direct object verb, while in (10b) it is the argument of an indirect object verb.

---

<sup>2</sup> Like other clitics in French, *se* has a strong form, namely *soi*, which I do not consider here, since it is no longer in use as a reflexive.



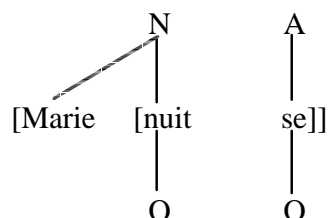
(10) a.



Marie=se regarde.

Marie look at herself.

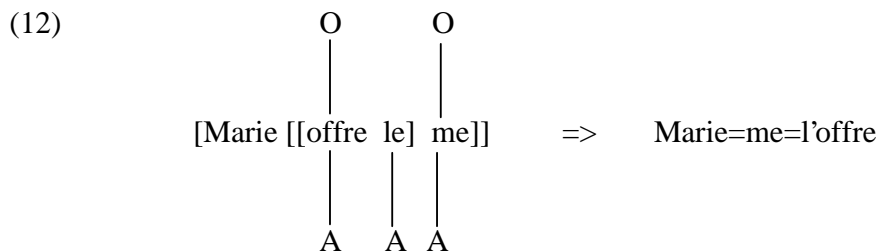
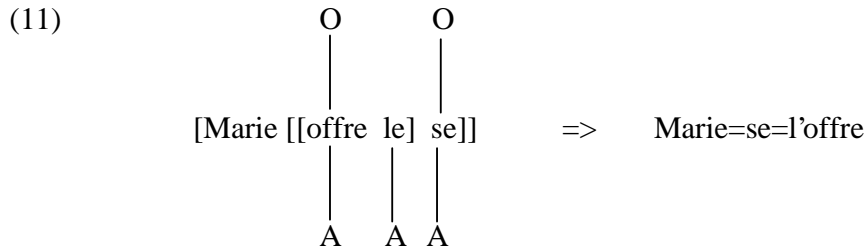
b.



Marie=se nuît.

Marie harms herself.

That *se* is doubly specified for Case is confirmed by the fact that in a clitic cluster, it behaves like other double-Case clitics. As discussed in the preceding chapter, the WEFC forces double-Case pronouns, e.g., *me*, *te*, to move outside the complement domain of a verb before the single-Case clitic *le/la*. Now, *se* is parallel to first and second person clitics in that it is always ordered before other clitics; compare (11) and (12) (non-crucial features are omitted). Therefore, I conclude that *se* is specified for two Cases.



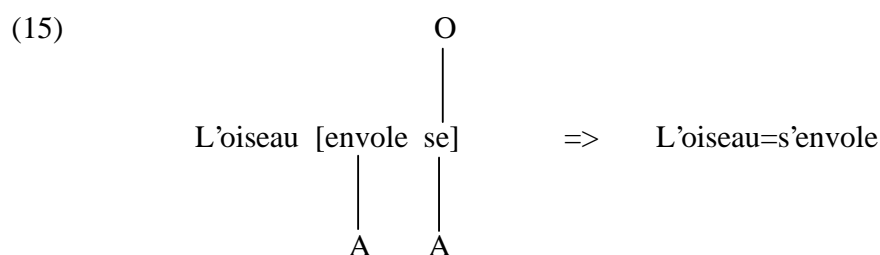
As for inherent-*se*, I claim that it is also generated in postverbal position and undergoes movement under the OCP, like all other clitics. At first glance, this claim may appear very odd, since verbs taking an inherent-*se* do not alternate with an NP argument, unlike verbs which are used with reflexive/reciprocal *se*; compare the reflexive/reciprocal *se* in (13) and the inherent-*se* in (14).

- (13) a. Marie a regardé Paul.  
           Marie watched Paul.  
       b. Marie s' est regardée.  
           Marie SE watched.
- (14) a. Marie s' est évanouie.  
           Marie fainted.  
       b. \*Marie a évanoui Paul.  
           Marie fainted Paul.

Wehrli (1986: 268) points out that "traditional grammarians have noted that many inherent-*se* derive historically from reflexive or reciprocal *se*." He interprets this observation as evidence for his claim that *se* absorbs an argument (see below). According to him, while inherent-*se*, which is generated in clitic position (under the V node), absorbs a lexicalized (or frozen) argument, it is not to be taken to be an

argument. Unfortunately, this is not conclusive enough to consider *se* to be an affix in an underived position. The observation of traditional grammarians may be interpreted as evidence to the effect that in inherent-*se* verbs, *se* is an internal argument which is frozen or lexicalised in the argument structure of the verb, very much like NP idioms (on idioms, see for instance Di Sciullo, 1983). That is, even though the semantics of the construction (i.e. its reflexive or reciprocal aspect, as noted by traditional grammarians) has been lost over time, the syntactic process that underlies it, namely movement by OCP, remains unchanged, as I will show.

Structure (15) illustrates a derivation of inherent-*se* structure; as claimed above, *se* is doubly specified for Case, like first and second person pronouns. If one assumes that this verb is specified for accusative Case, the movement of *se* is accounted for straightforwardly by the OCP.<sup>3</sup>



The claim that inherent-*se* is generated in the complement domain of the verb is supported by its distribution with respect to the verb. Indeed, like other pronominal clitics, inherent-*se* precedes the complex verb+auxiliary in compound tenses, as can be seen in (16). If inherent-*se* were an affix (as suggested in Wehrli's analysis), one could not explain this fact straightforwardly, for in general affixes are rigidly ordered with respect to their host. Under such an analysis, one would be led to assume that all clitics are affixed to their verb, since inherent-*se* alternates with other clitics, i.e., when the subject is first or second person, inherent-*se* becomes respectively *me* and

---

<sup>3</sup> Unfortunately, it seems that there is no way to check whether the relevant Case of the inherent-*se* verb (which licenses the clitic) is accusative or oblique. I assume in the text that the verb is specified as accusative, but the result would not change if it were specified as Oblique. The impossibility of testing the Case is due to the fact that an inherent-*se* verb may never take the other third person clitics as complement, namely *le* and *lui*, which are specified for a single Case.

*te*; compare (17a) and (17b). Under the base-generation analysis of inherent-*se*, all the gains made in the previous chapter would be lost.

- (16) a. Il **s'**[est évanoui].  
           He SE is fainted.  
           He fainted.
- b. \*Il est **s'**evanoui.
- (17) a. \*Je **se** suis évanoui.  
           I SE fainted.
- b. Je **me** suis évanoui.  
           I ME fainted.

To conclude, in my view there is no reason to assume that inherent-*se* is different syntactically from reflexive/reciprocal *se*, or other pronominal clitics. The simplest assumption is that syntactically, it is generated in the complement domain of the verb selecting it, like other object clitics, and that the OCP forces it to move outside this domain. The peculiarity of inherent *se* with respect to other clitics and reciprocal or reflexive uses of *se* is that it is lexicalized in the argument structure of the verb, like an idiom. Below I present an account of middle *se*.

### 3.2.1 Middle *se*

Parallel to inherent *se*, the simplest assumption is that middle *se* also is generated within the complement domain of the verb, as I will show. The peculiarity of this construction is that the syntactic subject is semantically the object of the verb, and in fact, it is traditionally assumed that the subject of such a construction is not a genuine one. In some analyses, the subject of middles is derived by movement of the internal argument to the subject position. For instance, in Wehrli (1986), although *se* absorbs the external argument of a verb, the subject position is not suppressed. Thus, the internal argument is raised to the empty subject position (see also Burzio, 1986; Kayne, 1975; Ruwet, 1972; Zribi-Hertz, 1982). In others, middles do not involve movement; the syntactic subject results from a lexical rule which externalizes the internal argument (Fagan, 1992; Grimshaw, 1982) or from the saturation of the

external argument at Lexical Conceptual Structure, given the Multi-Level Saturation Hypothesis (Di Sciullo, 1991). There are many problems which are inherent to the movement analysis, also referred to as raising, as I show immediately below.<sup>4</sup>

The Raising Transformation is used in the treatment of passive, unaccusative verbs, as well as in *seem*-type verbs, which are referred to as raising verbs. As noted above, it requires that an empty subject position be generated by the grammar. The existence of such an empty position cannot be maintained here, given my assumptions. Recall that the phrase structure tree is built dynamically by successive applications of Merge and that only lexical elements are computed.

Furthermore, in French, verbs agree with their subject in person and number, but never with their object. If in middle constructions the syntactic subject were underlyingly a syntactic object, it would not agree with the verb in person and number. The fact is, in structures with a subject inversion, which is a syntactic process, the verb still agrees with the inverted subject, as exemplified in (19). If the subject of middle verbs were in object positions at some stage of the derivation, its morphological shape would not be affected by the movement process, and would be parallel to subject inversion.<sup>5</sup> Since this is not the case, as can be seen in (18), the raising analysis of middles cannot be maintained.

- (18) a. Les légumes se vendent bien aujourd'hui .  
The vegetables SE sell.3.PL. well today.  
b. \*Les légumes se vend bien aujourd'hui .  
The vegetables SE sell.3.SG. well today.
- (19) a. Que font les enfants?  
What do.3.PL the kids?  
b. \*Que fait les enfants?  
What do.3.SG the kids?

---

<sup>4</sup> On middle constructions in general, see Keyser and Roeper (1984), Fagan (1988, 1992), Authier and Reed (1996), Stroik (1992, 1999), Ackema and Schoorlemmer (1995), etc.

<sup>5</sup> As I will argue in chapter 4, subject inversion of the type illustrated in (19) is not derived by verb movement, but by NP movement, which is triggered by a gemination effect.

Moreover, if the derived subject were underlyingly in object position, one would expect to find an object clitic in subject position. As shown in (20) with a passive structure, this is not the case: the derived subject must be inherently specified for nominative Case, and the appearance of an accusative pronoun yields an ill-formed structure (21).

- (20) a. Il a été tué par un chauffard.  
He was killed by a driver.  
b. Elle semble apprécier la bière.  
She seems to like beer.
- (21) a. \*Le a été tué par un chauffard.  
Him has been killed by a driver.  
b. \*La semble apprécier la bière.  
Her seems to like beer.

Given the problems just discussed, it is more natural to assume that there is no raising of an underlying object to an empty subject position. I will, indeed, assume that the subject of middle verbs results from a lexical rule of externalization of the internal argument, along the lines of Fagan (1992) and Grimshaw (1982). Given a transitive verb whose argument structure is schematized in (22a), the lexical rule of middle transformation will demote the subject (X), replace it by the object Y and insert the reflexive in the place of the latter, as seen in (22b).<sup>6</sup>

- (22) a. [X [Verb Y]]  
b. [Y<sub>i</sub> [Verb SE<sub>i</sub> ]]

Many researchers argue that *se* is not an argument, and therefore middle verbs are intransitive verbs (cf., among others, Wehrli, 1986; Di Sciullo, 1991). In fact, I claim that *se* is an argument of the verb in middle constructions: as a clitic it is

---

<sup>6</sup> As for the interpretation of middles, Fagan (1988), relying on Rizzi (1986a), proposes a rule assigning *arb* to the external theta-role, the latter being the theta-role associated with the external argument. Di Sciullo (1991) states that "when argument saturation occurs at LCS, the interpretation of the implicit argument is fixed. It is arbitrary generic in middles (...)" (p. 17). As stated above, I leave open the semantics of this construction.

licensed by Case and the lexical structure of the verb, as stated in (1). To support this view, one may note that the argument structure of transitive verbs may be modified by new morphology; for instance adding passive morphology to a two-argument verb yields a one-argument verb. But, since *se* is not morphology, it cannot modify the argument structure of a transitive verb; rather it can only fill a slot in a verb argument structure. When *se* occupies a slot in the argument structure of a verb, there is no room for a further object argument. Just like any simple transitive verbs, a middle verb may not take more or less arguments than required by its argument structure. Indeed the demotion of their subjects is concomitant with the appearance of *se*; hence the nature of the arguments changes, but not their number. Middle verbs are misleading because actually they are reflexive verbs whose reflexive interpretation cannot be obtained, mainly because the subject is inanimate. This is supported by the fact that if the externalized argument is animate, *se* is likely to be interpreted as a real reflexive, and hence the middle reading may be ambiguous or even impossible (see also Zribi-Hertz, 1982). For instance, from the structure in (23a), it is difficult to obtain the middle interpretation (23b) because the NP is preferably interpreted as a non-derived subject and coindexed with the reflexive.<sup>7</sup>

- (23) a. L' institutrice aime bien ces élèves.  
The teacher likes well her students.
- b. \*?Ces élèves s' aiment bien.  
These students SE like well.

It turns out that the lexical rule does not violate the argument structure of the verb although it does modify it, allowing a different type of meaning to be conveyed by the verb. In other words, *se* functions as a dummy argument, introducing no semantic information: an internal argument is still present, but it is meaningless, since the reflexive interpretation is not possible with an inanimate subject.

---

<sup>7</sup> Similarly, the reflexive and the reciprocal constructions are potentially ambiguous. A sentence such as *Paul et Marie se détestent* 'P. and MSE hate' may be interpreted either as a reflexive, i.e. Paul hates himself, Marie hates herself, or as a reciprocal ('Paul hates Marie and Marie hates Paul'). It is by virtue of this ambiguity that adverbs such as *mutuellement* 'mutually', *un l' autre* 'each other', are often used in reciprocal constructions. Thus, the above sentence is not ambiguous if it includes such an adverb: *Paul et Marie se détestent mutuellement/un l' autre*

In the light of this reasoning, consider two well-known properties of middle constructions in French: (a) middle *se* cannot cooccur with an infinitival verb, nor in an adjunct clause (24); and (b) it allows an expletive *il* (25b). While traditional analyses need a very complex apparatus, including NP movement and deletion, as well as stipulations on the properties of *se* (see, for instance, Burzio, 1986), under my assumptions these structures can be accounted for straightforwardly.

- (24) a. \*La possibilité de **se** construire des immeubles est limitée.  
The possibility of SE to build any buildings is limited.
- b. \*Beaucoup de livres **s'** achètent sans ~~PRO~~ **se** lire.  
Many books SE buy without SE reading.  
'Many books are bought without being read.'
- (25) a. Beaucoup d' immeuble**se** construisent dans cette ville.  
Many buildings SE build (' are built' ) in this city.
- b. Il **se** construit beaucoup d' immeubles dans cette ville.  
It SE builds many buildings in this city.

In (25b), one must explain (a) why the middle transformation does not externalize the internal argument (*beaucoup d' immeuble*), and (b) how the simple transitive verb licenses three arguments: the expletive clitic *il*, the NP *beaucoup d' immeubles* and the clitic *se*.

Consider the first question. Obviously the NP is not externalized because the expletive clitic is computed as subject. If an expletive subject can be inserted, it must be the case that the externalization rule is not absolute. The question then arises as to why the externalization of the NP is obligatory in the absence of the expletive (cf. \**se construit beaucoup d' immeuble*). In my view, the explanation is simple: *se* being doubly-specified for Case features ([A] and [O,β]), it absorbs the accusative Case of the verb, but remains with an extra Case, namely [O,β]. Since the nominative Case of the verb is not saturated, it clashes with the extra Case of the reflexive ([O,β]), and therefore the structure is ill-formed, as discussed in the preceding chapter. This feature clash could be avoided if the NP were able to absorb the nominative Case of the verb. But this cannot be, since an element which is assigned nominative Case must agree in person and number with the verb, and further the NP is not intended to



be the subject. It turns out that the nominative Case of the verb must be neutralized; this can be done either by externalization of the internal argument or by an expletive subject.

Let us turn now to the second problem stated above with respect to the licensing of three arguments by a simple transitive verb. I show that it is accounted for directly under (1). First of all, notice that in a middle like (25a), the reflexive is overlicensed; that is, (a) it absorbs the accusative Case of the verb (obligatory Case-licensing); (b) it fills a slot in the argument structure of the verb, and (c) it is coindexed with the syntactic subject. Since either (b) or (c) is sufficient, the reflexive is licensed more than is required. Likewise, the NP subject is overlicensed: (a) it fills a slot in the argument structure of the verb and (b) it is assigned nominative Case, while only one of these would be sufficient. I show that it is precisely the overlicensing of both of the arguments that allows a third argument (expletive) to be computed. The reasoning is that since two arguments are overlicensed by some properties, say *x* and *y* respectively, a further argument can be licensed by *x* and *y* provided that it has no consequence for the meaning of the structure. That is, it must be a dummy argument.

Indeed, the expletive pronoun is such an element. Being a clitic, it must obligatorily be licensed by Case and by filling a slot in the argument structure of the verb. It may not be licensed by coindexation, since being non-referential by definition, it cannot be coindexed with any element, like the indefinite clitic *on* (see Rizzi, 1986b; and below). This turns out to be true. Indeed, in (25b) the expletive clitic is licensed by the nominative Case of the verb as well as by its argument structure; the NP remains in object position and is now licensed only by the argument structure of the verb; as for *se*, it is licensed by the accusative Case of the verb and by coindexation with the NP, and no longer by occupying a slot in the argument structure of the verb.

It should be noted that a simple reflexive cannot give rise to an expletive structure. Although *se* in (26a) is overlicensed, the verb cannot take an expletive subject clitic, parallel to (25b), as shown in (26b). This result is in fact expected, since there is no lexical rule that demotes the object, in contrast to the expletive

middle. That is, expletive middles are possible because when the subject is demoted, something else, whether the object or an expletive, may replace it. In a simple reflexive, in contrast, there is no lexical rule of demotion of the subject.

- (26) a. Paul **se** connaît.  
           Paul knows himself.  
       b. \*Il **se** connaît Paul.  
           It SE knows Paul.

In our analysis of expletive middles, the reflexive *se* does not occupy a slot in the argument structure of the verb, and therefore it must be coindexed with another argument. This is confirmed by the fact that in (25) if either of the interacting elements is missing, an ill-formed structure will result. Indeed, despite the fact that all Cases of the verbs are absorbed, (27a) is ungrammatical since *se* is not coindexed with an argument. (27b) is also ungrammatical, but for two related reasons: the nominative Case of the verb is not saturated, and therefore it clashes with the oblique Case of the reflexive. Similarly, (27c) is ungrammatical under the expletive reading of *il*, even though all Cases of the verb are satisfied, for in the normal case (i.e. without the middle transformation) this verb may not take an expletive subject.

- (27) a. \*Il *se*<sub>j</sub> construit dans cette ville.  
       b. \**Se*<sub>j</sub> construit beaucoup d' immeubles dans cette ville.  
       c. \*Il construit beaucoup d' immeubles dans cette ville.

Consider now the structures in (24), repeated below as (28), which contains a nonovert subject. Since in the literature there is no case of an infinitive verb which does not have a subject, overt or null, one may take PRO to be obligatory and automatically generated as the subject of every infinitive verb in the absence of an overt subject. If so, a null subject may not be demoted by the lexical middle rule, since this would be a string vacuous operation. Thus, the ill-formedness of (28a) is accounted for straightforwardly: since PRO cannot be demoted, no element can be inserted in its position, either the NP object or an expletive. Anyway, as a clitic an expletive would have to be obligatorily Case-licensed, a requirement that the infinitive verb cannot satisfy. In (28b), both the main clause and the adjunct clause

contain a middle verb. As in the previous case, the null subject in the adjunct clause cannot be demoted, and therefore the middle rule cannot take place.

- (28) a. \*La possibilité de [PRO **se** construire des immeubles] est limitée.  
           The possibility of SE to build any buildings is limited.
- b. \*Beaucoup de livres s' achètent sans PRO **se** lire.  
           Many books SE buy without SE reading.  
           'Many books are bought without being reading.'

Summarizing, I have suggested that there is a single morpheme *se* in French, which is doubly specified for Case. It is distinct from other double-Case arguments in that it is inherently a reflexive pronoun which must be bound by an argument, usually the subject. It is its reflexive property that makes it able to appear in middle constructions. I have argued that middles can be best explained as the result of a lexical operation which modifies the argument structure of a verb prior to the computation of syntactic structure. All other particularities of this construction follow from the previous assumptions, namely the Argument Licensing Constraint and the OCP, which moves *se* out of the complement domain of the verb in all types of constructions.

### 3.3 The syntax of *on*

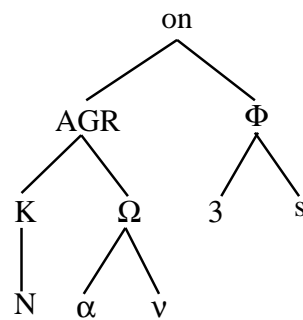
The pronoun *on* can be used only as subject. It is thus syntactically similar to the third person subject pronouns *il*, and *elle*. Some examples are given in (29).

- (29) a. **On** va au cinéma tous les jours.  
           People go to the movie every day.
- b. **On** aime nager.  
           People like swimming.

However, *on* is different from other third person subject pronouns as it provides no indication of number and gender; it refers strictly to an indefinite (non-specific/non-referential) human, individual or collective. Thus the feature matrix of

*on* must contain at least the following features: [animate], [nominative], [indefinite]; the latter will be represented by the Greek character  $\nu$  (nu).

(30)



As seen above, an expletive cannot be the antecedent of reflexive clitic *se*. This is parallel to the following contrast, noted by Kayne (1972): while the third person subject pronoun *il* can be the antecedent of another pronoun (31a), *on* cannot be (31b).

- (31) a. Il<sub>i</sub> dit que je l<sub>i</sub>' ai insulté.  
           ' He says that I insulted him.'  
       b. \*On<sub>i</sub> dit que je l<sub>i</sub>' ai insulté.  
           ' One says that I insulted him.'

The referential inadequacy of *on* is further illustrated in (32) (cf. Kayne, 1972, 1975; Rizzi, 1986). Unlike other subject clitics, *on* is incompatible with some types of conjunction, as can be seen by comparing (32a) and (32b); however, with a compound tense, conjunction is possible (32c).

- (32) a. Il mange de la viande et boit du bon vin.  
           ' He eats meat and drinks good wine.'  
       b. \*On mange de la viande et boit du bon vin.  
           ' One eats meat and drinks good wine.'  
       c. On a mangé de la viande et bu du bon vin.  
           ' ~~On~~e has eaten meat and drunk good wine.'

Attempting to puzzle out these facts, Rizzi (1986) distinguishes two types of coordination, which appear to be structurally ambiguous: sentential coordination and phrasal coordination. The sentence (33a) illustrates this point. According to Rizzi, the structure (33b) is a case of sentential coordination: the subject position of the second conjunct, which is identical to the subject of the first one, appears as a null pronoun. In (33c) it is two VPs which are coordinated, and therefore there is no null argument in the subject position.

- (33) a. Mary arrived at 5 and left at 6.  
 b. [<sub>S</sub> Mary arrived at 5 ] and [<sub>S</sub> Øleft at 6]  
 c. Mary [ [<sub>VP</sub> arrived at 5 ] and [<sub>VP</sub> left at 6 ] ]

Rizzi relates (32b) to (33b) and assumes that the ill-formedness of the former is due to the very fact that the pronoun *on* cannot be the antecedent of the null pronoun involved in sentential coordination, whatever the nature of the null pronoun may be, as shown in (34a). There is no reason to challenge this empirically grounded analysis, and I will assume that it is on the right track. However, in order to exclude the other alternative, namely the phrasal coordination analysis of (32b), as illustrated in (34b), Rizzi is led to claim that VP conjunction is incompatible with subject clitics; in his perspective, since the subject clitic (Scl) must adjoin to a V head, it would violate the coordinate structure constraint, which requires the identity of the conjuncts: Scl [[V...] et [V...]] (=his (36b)). As for the well-formedness of (34c), it would follow from the idea that the conjoined structure does not include the adjunction site of Scl, namely the auxiliary, and therefore the coordinate structure constraint is respected.

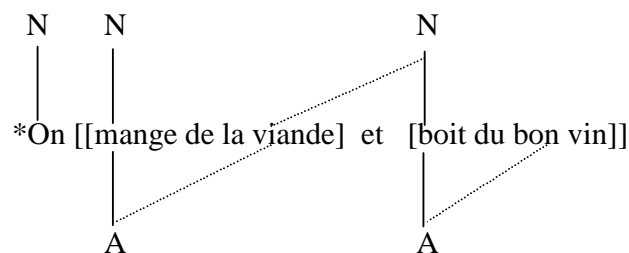
- (34) a. \*[On<sub>i</sub> mange de la viande] et [Ø<sub>i</sub> boit du bon vin].  
 b. \*On [[mange de la viande] et [boit du bon vin]].  
 c. On [a [mangé de la viande] et [bu du bon vin]].

While this analysis contains some insights with respect to coordination, it relies crucially on the assumption that pronominal clitics adjoin to tensed verbs. In my view, subject clitics do not move in simple declarative clauses, and therefore such an account is not available. In fact, the ill-formedness of (34b) can be better explained with my basic assumptions. In particular, I will show that the phrasal coordination of

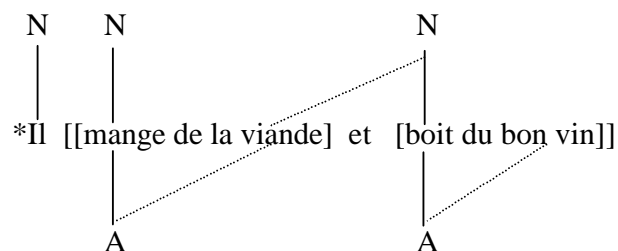
tensed verbs is ruled out precisely because such a structure lacks an empty pronominal.

In my view, since the verb in (34b) is specified for nominative Case, it must discharge this Case on an argument so as to avoid a feature clash, as discussed above. To put it another way, the nominative Case of the verb must be saturated. Now, in a phrasal coordination the tensed verb of the second conjunct lacks an argument and therefore automatically spreads its nominative Case onto the object of the first conjunct. Then a Case conflict arises, or what comes to the same thing, a mismatch between Case and grammatical relation, since the object of the first conjunct is assigned accusative Case by the higher verb. This is illustrated in (35a). Notice that this phenomenon must be independent of the nature of the subject, i.e., whether it is definite or not; for instance, the structure (32a) is ungrammatical under VP coordination as illustrated in (35b).

(35) a.

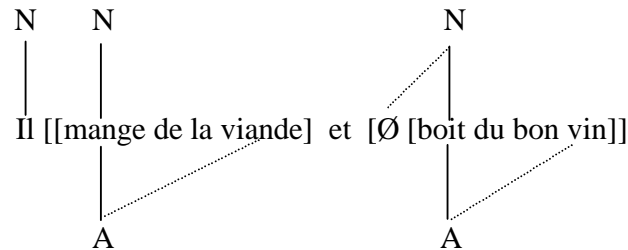


b.

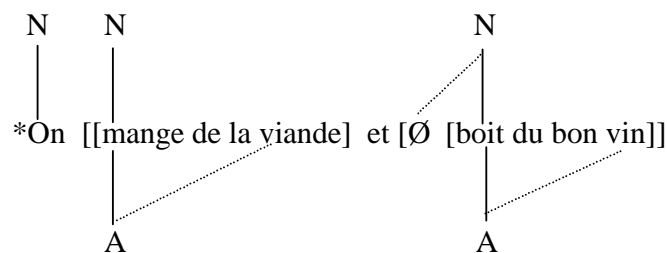


Under sentential coordination (32a) is well-formed, since the null pronoun absorbs the nominative Case of the verb, as can be seen in (36a). On the other hand, (32b) is ill-formed, as illustrated in (36b), since *on* cannot control the null pronoun in the second conjunct, as in Rizzi' s account.

(36) a.

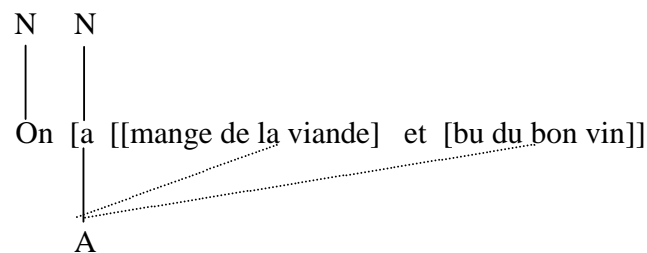


b.



Under this analysis, the well-formedness of (34c) is expected, since it involves the coordination of Caseless elements, namely participial phrases. Indeed, the auxiliary spreads accusative Case onto both conjuncts, as illustrated in (37). (Notice that this representation is not entirely accurate, as the feature of the *de*-phrase is omitted; see below.)

(37)



Summing up, *on* is taken to be specified for nominative Case, and therefore it is a subject clitic, a conclusion which is consistent with standard assumptions. It differs from other subject clitics in the paradigm in that it is specified for a feature that prevents it from controlling a referential pronominal element.

### 3.4 The syntax of *en*

Up to now, I have discussed three types of Cases in the French pronoun system, namely nominative, accusative, and oblique, each of which is associated with a semantic feature. Nominative Case corresponds to various semantic roles, including agent, experiencer, etc.; Accusative Case corresponds to theme by default (it is not specified for a semantic feature); Oblique Case is related to the semantic feature benefactive (including malefactive). The latter requires the argument to agree with the verb via the morpheme *à*.

The clitics to be discussed in this section (*en*) and in the next (*y*) are different from the previous ones with respect to thematic roles. It is argued that *en* is specified for accusative Case and for the thematic feature which is associated with the preposition *de*. Once the feature of *en* is properly determined, its syntax will follow from the general principles discussed so far.

#### 3.4.1 Feature structure

As is well known, the pronoun *en* essentially corresponds to certain phrases of the type *de*-NP in French (cf. Gross, 1968; Kayne, 1975; Milner, 1982; etc.). So the features of *en* must be the same as those of the morpheme *de* that it corresponds to. If the latter were specified for accusative Case, it would be expected to be compatible only with an accusative verb, as required by the Constraints on Merge. This is indeed the case, as shown in (38). In (38a) the verb *boire* 'to drink' takes a direct object NP, which can be pronominalized with an accusative clitic (38b). As expected, this verb may take a *de*-phrase as object, as shown in (38c). In contrast, the verb *obéir* 'to obey', which is specified as oblique-benefactive (39a,b), may not take *de*-phrase as complement, as seen in (39c).

- (38) a. Marie [boira [le vin]].  
           Marie will drink the wine.  
       b. Marie **le** boira.  
           Marie will drink it.  
       c. Marie [boira [du vin]]

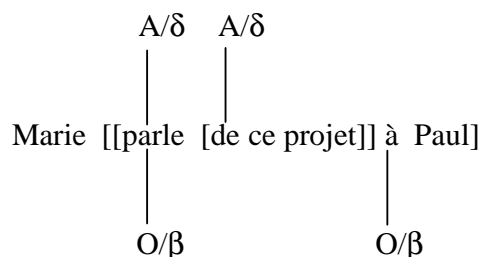


Marie will drink some wine.

- (39) a. Marie [obéit [à Paul]].  
           Marie obeys Paul.  
       b. Marie **lui** obéit.  
           Marie obeys him.  
       c. \*Marie [obéit [de Paul]]  
           Marie obeys DE Paul.

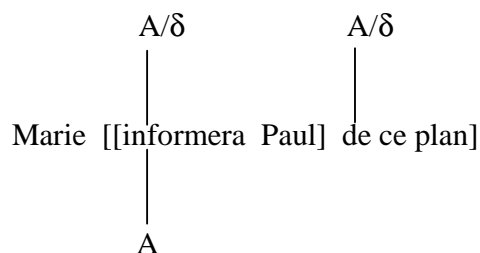
As for thematic features, *de* may have the value of a possessive, part of a whole, source, location, etc. Let us refer collectively to these values as delta ( $\delta$ ). In most cases, a verb may not be specified for these features, since the *de*-argument is not always obligatory. However, some verbs obligatorily take a *de*-phrase complement. Instead of including this information in the argument structure of the verb, as in traditional analyses, I will assume that such verbs are specified for accusative Case and for the semantic feature delta. Given the Constraints on Merge, if a verb is inherently specified for the feature  $[A, \delta]$ , its argument must agree with it (strong agreement), since a relevant morpheme is available. This is parallel to  $O, \beta$ -verbs which obligatorily take an *à*-phrase argument, as already discussed. On the other hand, a bare accusative verb (which is not specified for thematic features) may take either a *de*-phrase or an NP complement, according to the intended meaning. This is illustrated in (40) and (41) with two ditransitive verbs; the verb *parler* 'to speak' takes an  $O, \beta$ -argument and an  $A, \delta$ -argument, while *informer* 'to inform' takes a bare accusative argument and an  $A, \delta$ -argument. The Case specification of these verbs is supported by the fact that sentences (40b) and (41b), where one of the arguments (in boldface) is not conform to the Case of the verb, are ungrammatical. (Notice that in (40a), and in subsequent examples, I use an even more simplified representation, in which a single line is shown for each pair of Case/thematic features; keep in mind, however, that each feature is connected to the class feature K with its own association line.)

(40) a.



- b. \*Marie [[parle **ce projet**] à Paul].  
 Marie talks of this project to Paul.

(41) a.



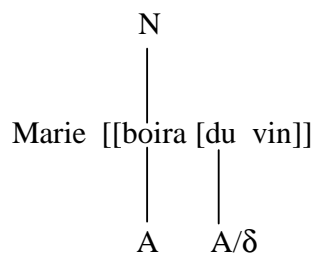
- b. \*Marie [[informera Paul] **ce plan**].  
 Marie will inform Paul of this plan.

It is clear that a verb will obligatorily take a *de*-NP complement if it is specified for accusative Case and for the thematic feature  $\delta$ . If it is not specified for thematic features, it may take either a bare NP or a *de*-NP complement; compare (42a) and (42b). In contrast, a benefactive verb can never take a *de*-NP complement, as illustrated in (42c), which contains a feature clash.<sup>8</sup>

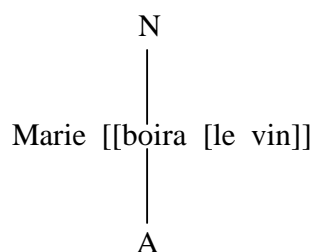
---

<sup>8</sup> If the absence of a feature is reanalysed as a null feature, as suggested in chapter 1, one may think that the null thematic feature related to the accusative Case of the verb clashes with the thematic feature of the preposition in (42a): \* $\emptyset$ vs.  $\delta$ . This is not the case, however. The assumption is that the reanalysis holds only when the verb does not share any feature with the argument, that is to say it has a class feature to which no terminal feature is attached. Put another way a clash would appear if thematic features were dependent of their own class node. Thus, in (42a), the verb being specified for accusative Case, which will be saturated by the *de*-phrase argument, its lack of thematic feature may not be reanalysed as a null thematic feature so as to provoke a clash with the thematic feature of the preposition (see the discussion of Italian and Spanish imperatives in chapter 4).

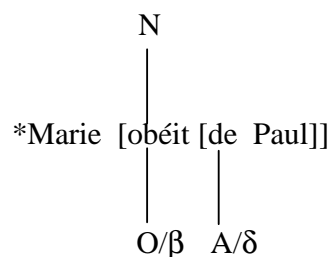
(42) a.



b.



c.



Now, since *en* may replace a *de*-phrase with the value  $\delta$ , it must be the case that it is specified for  $\text{A}/\delta$ .<sup>9</sup> Furthermore, in the normal case *en* is not compatible with a *de*-phrase containing a human NP, as illustrated by the contrast in (43) and (44). Given this fact, I will assume that *en* is specified for the feature *thing* ( $\tau$ ), a feature I

---

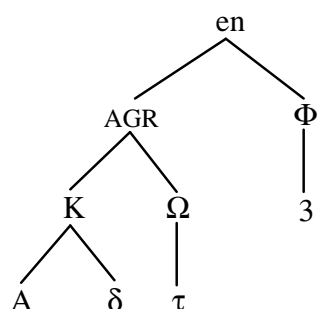
<sup>9</sup> Many *de*-phrases may not be pronominalized with *en*, for instance the following:

- (i) a. Max a grandi de deux centimètres.  
       \*Max en a grandi.  
       Max grown by two centimetres.  
       b. Jacques est de bonne humeur.  
       \*Jacques en est.  
       Jacques is in a good mood.

discussed in chapter two in the context of the Animacy Hierarchy. Therefore, the feature structure of *en* must be as given in (45).<sup>10</sup>

- (43) a. Marie parlait de son plan.  
Marie talked about her plan.  
b. Marie en parlait.  
Marie EN talked.
- (44) a. Marie parlait de Paul.  
Marie talked about Paul.  
b. ?Marie en parlait.  
Marie EN talked.

(45)



Having determined the feature structure of *en*, I turn now to its syntax. Its relative order with respect to other clitics and certain restrictions on its distribution will be shown to follow from the general principles discussed above; no additional constraint is needed.

### 3.4.2 OCP effects

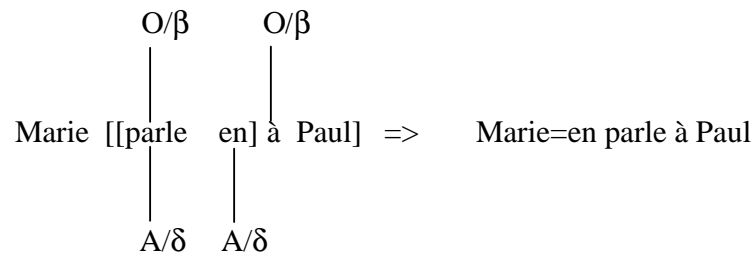
Under the proposed feature structure, the syntax of *en* is accounted for straightforwardly. Since *en* bears the same features as the preposition *de*, it may

---

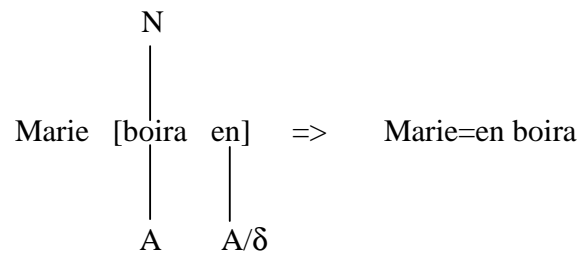
<sup>10</sup> In colloquial French, many speakers use *en* both for inanimate and human. It seems that a change is about to take place in the paradigm: the feature  $[\tau]$  seems to be neutralized for those speakers. Once completed, the change will have certain consequences for the linear order of *en* with respect to other clitics and for the expression of inalienable possession (see below in the text).

pronominalize the *de*-phrases in the above structures. On the simplest assumptions, it must be the case that it is generated in the complement domain of the verb, like other object clitics; this is illustrated by the derivations in (46), (47) and (48).

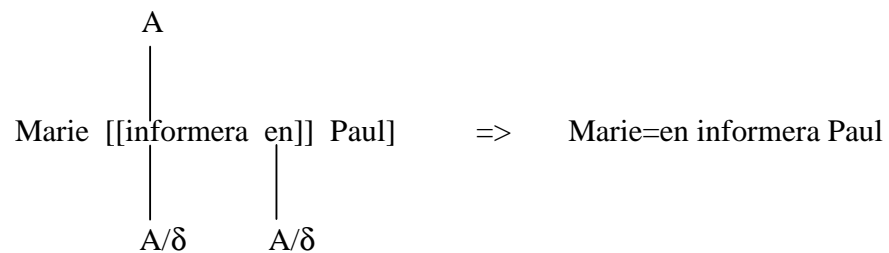
(46)



(47)



(48)



Since *en* is specified for accusative Case, it is moved by virtue of the OCP, landing at the first element outside the complement domain of the verb, namely the subject in these examples. (In the representation of ditransitive verbs, the nominative Case is omitted for simplicity.) In the next section I consider the ordering of *en* with respect to other clitics.

### 3.4.3 Animacy effects

Since *en* is specified for thing ([τ]), it is expected to behave differently from the clitics discussed in the preceding chapter. In particular, it should be ordered after an animate clitic, in accordance with the Animacy Hierarchy (AH). This is borne out: as can be seen in (49) and (50), *en* must follow the dative clitic; otherwise an ill-formed structure results.

- (49) a. Marie m'**en** parlait.  
       b. \*Marie **en** me parlait.  
           Marie EN talked to me.
- (50) a. Marie lui **en** donnera.  
       b. \*Marie **en** lui donnera.  
           Marie EN will give some to him.

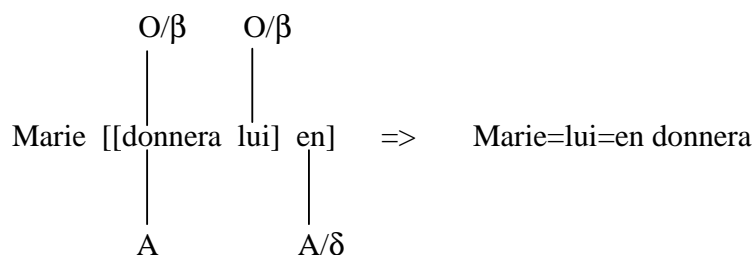
Recall that in a cluster of third person clitics, the accusative clitic (unspecified for animacy) precedes the oblique one, which is specified for animacy (51); but when the oblique clitic is first or second person, it precedes the accusative one, suggesting that the AH is irrelevant, as illustrated in (52).

- (51) a. Marie le lui donnera.  
       b. \*Marie lui le donnera.  
           Marie will give it to him.
- (52) a. Marie me le donnera.  
           Marie will give it to me.  
       b. \*Marie le me donnera.

Indeed, I have argued that in both cases the accusative clitic is computed first; when two third person clitics are involved, the one which is closer to the target moves first because both clitics have an equal number of features. However, if the lower clitic is doubly specified, it triggers a violation of the OCP in both Case tiers, and by virtue of the Worst Evil First Constraint it moves before the other clitic.

Consider the derivation of (50). Given that both clitics are specified for animacy, the one which bears the feature [animate] must be computed first so as to respect the AH ( $\alpha > \tau$ ), as shown in (53). Since the AH is a semantic-like constraint, it takes precedence over all other constraints, if there are any, moving the dative clitic first, and then the accusative clitic.

(53)

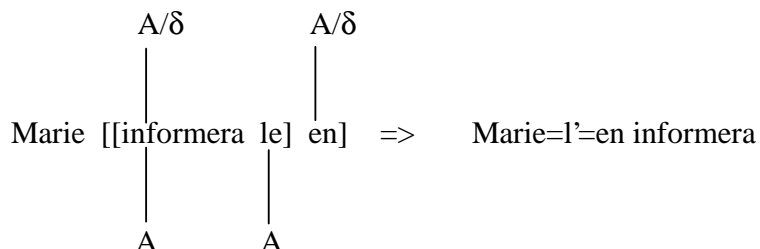


When *en* is in a cluster with another accusative clitic like *le*, which is not specified for animacy, the Animacy Hierarchy cannot apply as discussed in chapter 2. In this case, the accusative clitic *le* is computed first, since the unmarked order (in French) is DO>IO (see discussion in chapter 2).<sup>11</sup> This is illustrated in (54). (Notice that the WEFC does not hold here, so that *en* cannot move first, even though it has two types of features, namely Case and semantic features. Recall that the crucial assumption is that each of these clitics has only one K node and, therefore, they each have one offending node.)

---

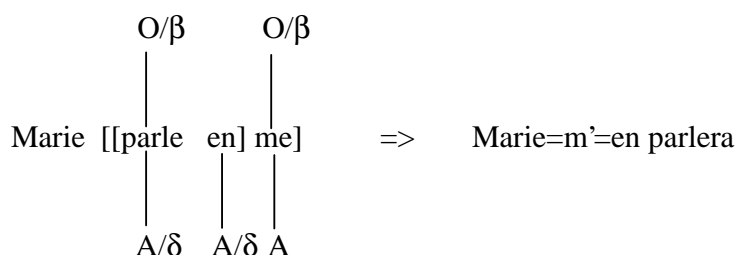
<sup>11</sup> This implies that in the absence of further constraints, the initial order of clitic arguments parallels that of NP arguments. In (53), if the arguments were NPs, the order would be NP.ACC > NP.OBL, which contrasts with the order of the corresponding clitics, namely CL.OBL > CL.ACC, because of the AH. In (54), where the AH is irrelevant, the order would be NP.ACC > NP.ACC/ $\delta$ , which parallels the order of the clitics, CL.ACC > CL.ACC/ $\delta$ . Whether or not the AH is satisfied earlier in the initial structure is not relevant, since the animate clitic has to move first anyway to satisfy both the OCP and the AH in the derived structure.

(54)



Notice that when the higher clitic is first or second person, as exemplified in (55), *en* cannot be Case-licensed. In effect the animate clitic, which is doubly specified for Case, absorbs both Cases of the verb, leaving no room for the Case-licensing of *en*. I suggest that this structure is repaired by a delinking process, at least in formal French. Under OCP, the clitic *me* moves automatically outside the complement domain of the verb. In order to allow the second clitic to be licensed, feature manipulation takes place, delinking the Case features of *me* up to structure preservation. As discussed in the previous chapter, the resulting form is *moi*, which yields the contracted form *m'*, assuming an ordinary phonological rule in French which delinks the two segments *oi* [wa] from their skeletal slots. This is illustrated in (56). (As I will show in chapter 4, a similar process takes place in Italian clitic clusters, except that Italian has no deletion rule in this context.)

(55)



- (56) a. Marie=moi=en parlera.  
 b. Marie=m' =en parlera(output)

Note that neither the rule of elision of *moi* nor the delinking process are *ad hoc* mechanisms. Both are independently needed for standard French. I have already discussed the rule of delinking in the context of imperative verbs (chapter 2). As for



the elision rule, it is very common in French phonology; for instance the clitic *le* is spelled out as *l* whenever it can be syllabified with an adjacent word, either its host (landing site) or the verb selecting it. In (57a), the clitic *le* is syllabified with the following auxiliary, while in (57b) it is syllabified with the preceding subject. The clitic *la* becomes *l* when it can be syllabified with the following verb, cf. (57c,d). One may note that a genuine strong pronoun, i.e. a strong pronoun which does not arise from a delinking process, cannot be reduced in any environment, cf. (57e,f).

- (57) a. Nous l' avons vu. [nu la v<sup>h</sup> vy]  
 b. Nous le savons. [nul sa v<sup>h</sup>]  
 c. Nous \*la/l' aimons.  
 d. Nous \*l' /la connaissons.  
 e. Toi et moi / \*t'e t moi.  
 f. Moi aussi / \*m'a ussi.

Normative grammarians suggest that strong first and second person pronouns must be elided only when they precede the pronoun *en*. The motivation for this elision is provided by the fact that, in imperative verbs, standard French contrasts with colloquial French: cf. *donnez-m' en* (formal) vs. *donnez moi-z-en!* (colloquial French). In my view, the elision rule must apply irrespective of the verb type (imperative or not) in standard French, since in a sentence like *Marie m' en parlera* (see (56)), *m'* stands for *moi*, not for *me*, as seen above. The difference between standard French and colloquial French in the context of imperative verbs can be accounted for straightforwardly once one assumes that in colloquial French the elision rule does not apply when the derived strong pronoun is a complement of an imperative verb; instead the independently needed segment *z* (liaison) is inserted.<sup>12</sup> This is illustrated in (58).

---

<sup>12</sup> One may speculate why standard French is different from colloquial French with respect to the elision rule. It may be that in dialectal French the elision rule of *moi* is ignored because it looks like a circular operation (without being a string-vacuous one). Indeed, the elision rule may apply independently to two different morphemes, *me* and *moi*, yielding the same output, *m'*. Thus, in the input structure *Marie donnera me en*, the elision rule may automatically apply to *me*, yielding *m'*. The delinking process then applies to *m'* yielding *moi*, to which another elision rule is further applied yielding *m'*; this is schematized in (ia). In dialectal French, the insertion of *z* instead of the elision of *moi* allows this apparent circularity to be avoided in imperative structures; see (ib). Perhaps, in non-imperative structures, the early elision of *me* and its syllabification with *en* trigger the reanalysis of

- (58) a. [[Donnez me] en] (initial structure)  
give me of-it  
b. [[Donnez moi] en] (delinking)  
b. Donnez-moi-z-en! (z insertion in colloquial French)  
c. Donnez m' en! (elision of *moi* in formal French)

Summing up, in the cluster *me/te+en* the verb cannot directly license *en*; as a result, a delinking rule is triggered. This rule is not obvious because the resulting strong form *moi* is further elided, becoming identical to the clitic form *me* (where the schwa is elided). The delinking rule is not specific to French grammar; in particular it is current in Italian and Spanish (spurious *se*), which will be discussed in the next chapter.

### 3.4.4 Subnominal *en*

Consider now the case where the *de*-phrase is a noun complement, as illustrated in (59). In this context, pronominalization with *en* is not always successful, as shown in (60). Following Postal (1994), I will refer to this *en* as *subnominal en*.

- (59) a. Marie connaît bien l'auteur [de cet article].  
Marie knows the author of this article.  
b. L'auteur [de cet article] connaît bien Marie.  
The author of this article knows well Marie.  
c. Marie écrira au directeur [de l'école].  
Marie will-write to-the principal of the school.
- (60) a. Marie **en** connaît bien l'auteur.  
Marie EN knows well the author.  
b. \*L'auteur**en** connaît bien Marie.  
The author EN knows well Marie.

---

the cluster *m' en* as a single opaque morpheme, as illustrated in (ic). The fusion of two clitics in a single one also occurs in Italian (see chapter 4) and in Catalan (see Bonet, 1991).

- (i) a. *me>m' >moi>m'*  
b. *me>m' >moi>moi+z*  
c. *me + en > m' en*

- c. \*Marie **en** écrira au directeur.

Marie EN will-write to-the principal.

In order to account for the behaviour of subnominal *en*, it is necessary to investigate the nature of the *de*-phrase when it is a noun complement. Consider the complex NP in (61), where the brackets reflect the basic constituency analysis (the label of each node, as usual, is irrelevant). Obviously both NPs (*l' auteur* and *cet article*) refer to distinct entities in the real world, and hence each bears a referential index. There is also a dependency relationship between them; the *de*-phrase seems to define the head noun in the manner of a restrictive clause. Let us assume then that they are in a predicative relationship in the sense of Williams (1980) (see also Tremblay (1991)). That is, the *de*-phrase is the predicate and the NP head, the subject. The coindexation of the *de*-phrase with the head noun is intended to show this relation.

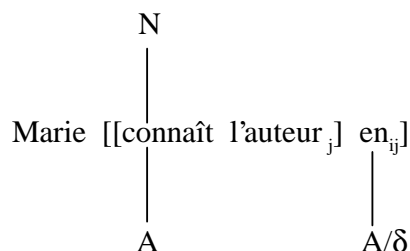
- (61) [L' auteur<sub>i</sub>[de [cet article]<sub>i</sub>]]<sub>j</sub> ]

The author of this article.

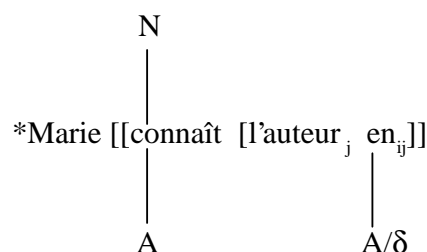
When the *de*-phrase is replaced by the relevant clitic, it is natural to assume that the indexing is not lost. In other words, the clitic *en* must be doubly indexed, since it refers to a real world entity and to an element in the structure (by predication). Thus, it turns out to have the properties of both a pronoun and an anaphor.

With this in mind, consider the derivation of (60a), where *en* is a subcomponent of a complex NP dominated by the head noun *l' auteur*. Let us take the initial structure of (60a) to be (62a), where *en* is generated as the complement of the verb. The alternative structure, illustrated in (62b), where *en* is generated as a noun complement cannot be maintained, since an NP may not select a Case-specified pronoun. Indeed, while the complex NP in (59a) may form an utterance on its own, it is never the case for the structure \**l' auteur en*. Put another way, if there is no verb specified for Case, there can be no clitic either, consistently with my proposals.

(62) a.

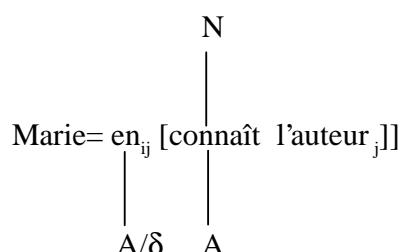


b.



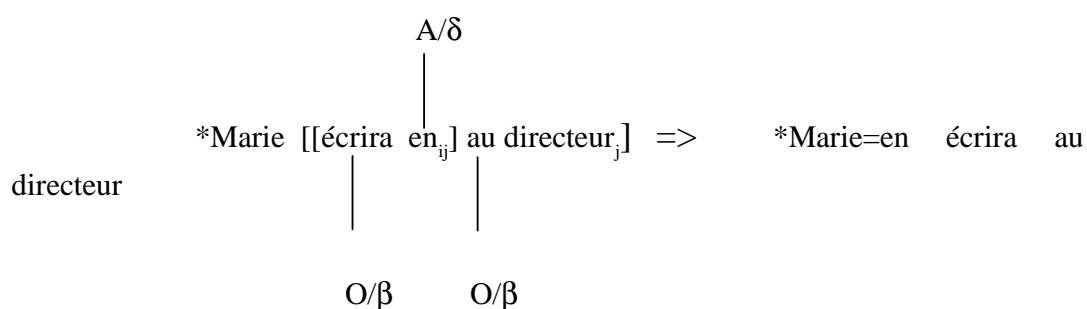
Now, as can be seen, the simple transitive verb *connaître* comes to license two internal arguments. How is this possible? This structure is in fact parallel with the expletive middle discussed above, and it satisfies the licensing constraint (cf. (1)). Indeed, the NP *l' auteur* need not be Case-marked and hence is licensed only by the argument structure of the verb. The clitic need not occupy a slot in the argument structure of the verb, and is licensed by coindexation with the NP and by the accusative Case of the verb. As for the subject *Marie*, it has no competitor, and therefore is redundantly licensed: it occupies a slot in the argument structure of the verb and is Case-marked by the verb. Notice that *en* normally moves under OCP, yielding the structure in (63).

(63)



Let us turn now to the structures where subnominal *en* yields ill-formedness, namely (60b) and (60c). In (60b), the antecedent of the clitic is the subject of the clause, while in (60c) the antecedent is the indirect object. I will show that these structures are ill-formed for different reasons, all of which are related to the licensing constraint. Consider (60c), which is more obvious. Its ill-formedness is readily understood; since *en* is specified for accusative Case, it cannot be licensed by a benefactive verb, consistently with the (d)-clause of the Constraints on Merge. Indeed, as can be seen in (64), the features of the clitic clash with those of the verb. It appears then that the ill-formedness of this structure is not due to the impossibility of extracting a clitic from a PP, as Kayne (1975) and others argue.

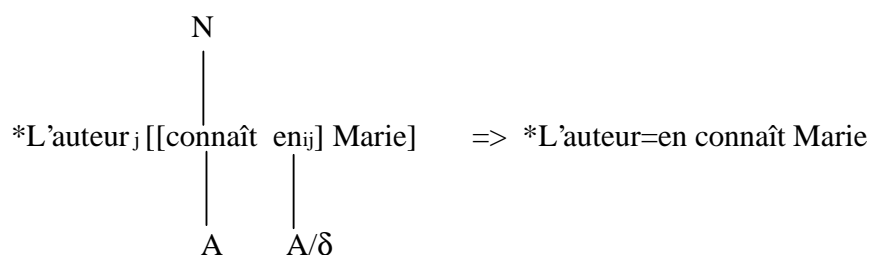
(64)



The ill-formedness of (60b) is more subtle. At first glance, all of the arguments are licensed: the clitic is licensed by Case and coindexation with the subject; the NP *Marie* is licensed by occupying a slot in the argument structure of the verb, while the NP *l' auteur* is redundantly licensed by Case and the lexical structure of the verb. The complex NP *l' auteur de cet article* fills a single slot in the argument structure of the verb. Therefore, the subnominal *en* which arises from this NP may compete for the

same argument slot as head noun (*l' auteur*), namely the subject slot, and should be licensed by the Case assigned to the latter (nominative). However, the picture is quite different. As illustrated in (65) the clitic is specified for accusative Case and therefore is compete for the same slot as the object *Marie*, which is not its antecedent.

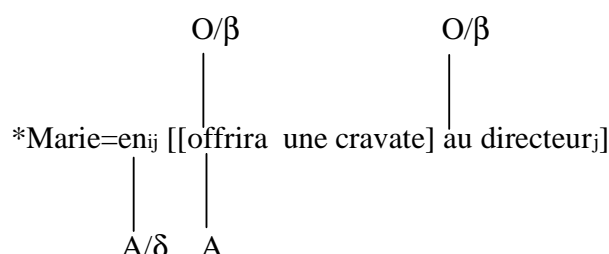
(64)



Additional facts of French confirm this analysis. When subnominal *en* arises from an NP indirect object of a ditransitive verb, see (66a), a well-formed structure cannot be derived, as shown in (66b). Here again the problem is that subnominal *en* is not properly licensed, i.e. it is not licensed by the Case corresponding to the grammatical function of the NP it is coindexed with. In effect, since it is coindexed with the IO *directeur*, it should be licensed by the Case which is assigned to its antecedent, namely the O, $\beta$ -Case. However, this is not so; it is always licensed by the accusative Case, i.e. the Case of the DO, which it has no referential relationship with. This is illustrated in (67).

- (66) a. Marie [[offrira une cravate] au [directeur<sub>j</sub> [de l' école<sub>j</sub>]] ]  
 Marie will offer a tie to the principal of the school.
- b. \*Marie en<sub>ij</sub> [[offrira une cravate] au directeur<sub>j</sub>].  
 Marie EN will offer a tie to the principal.

(67)



However, as expected, if in a ditransitive verb subnominal *en* arises from the NP direct object, see (68a), a well-formed structure is derived, parallel to (60a), as seen in (68b). This structure is perfect because the Case of *en* matches the grammatical function of its antecedent.

- (68) a. Marie offrira [une caisse<sub>j</sub> [de chocolats<sub>ij</sub>]] au directeur de l'école.  
 Marie will offer a box of chocolates to the principal.  
 b. Marie *en*<sub>ij</sub> offrira une caisse<sub>j</sub> au directeur de l'école.  
 Marie EN will offer a box to the principal.

Furthermore, it has been noted that *en* may not be separated from its antecedent by another NP; for instance, subnominal *en* cannot arise from sentence (69a), as shown by the ungrammaticality of (69b). I argue that (69b) is ill-formed because there is a mismatch between the Case that licenses the clitic and the grammatical relation of the antecedent, *la préface*. As just discussed, a clitic in such structures must take the Case that the verb would assign to its antecedent, but not the Case that is assigned to another element. In other words, if the antecedent of the subnominal clitic is a direct object, the clitic must be licensed by the Case that is usually assigned to the latter. Now, in (69b) the clitic is licensed by the Case of the direct object *la critique*, with which it is not coindexed, while its antecedent is normally Case-marked by the *de*-phrase. This results in a mismatch between the grammatical function of the head and the Case of the clitic, hence the ungrammaticality of the structure.

- (69) a. Marie [a lu [la critique [de [la préface [de ce livre]]]]]  
 Marie read the critique of the preface of this book.

- b. \*Marie **en** [a lu [la critique [de la préface]]]

Marie EN read the critique of the preface.

One further piece of evidence is provided by the intriguing facts in (70) and (71), which are discussed in Ruwet (1972) (glosses added). In (70a) and (71a), when the *de*-phrase in brackets is replaced by *en*, an ill-formed structure is derived if *en* is in the main clause, see (70b) and (71b), and a relatively acceptable structure if it appears in the infinitive clause, see (70c) and (71c).

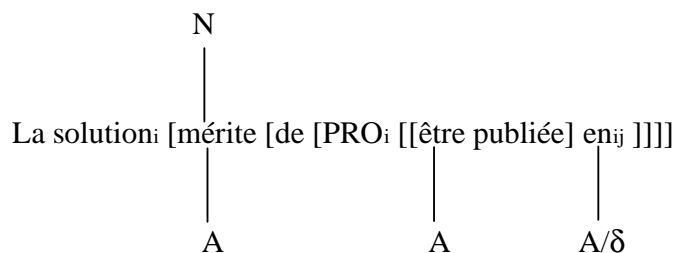
- (70) a. La solution [de ce problème] mérite d' être publiée.  
The solution of this problem merits being published.
- b. \*La solution **en** mérite d' être publiée.  
The solution EN merits being published.
- c. La solution mérite d'**en** être publiée.  
The solution of this problem merits EN being published.  
' The solution of this problem merits being published.'
- (71) a. L'h istoire [de la révolution] exige d' être écrite.  
The history of the revolution demands to be written.
- b. \*L'h istoi**re****en** exige d' être écrite.  
The history EN demands to be written.
- c. L'h istoire exige d'**en** être écrite.  
The history demands to EN be written.  
' The history of the revolution demands to be written.'

The fact that subnominal *en* is not allowed in the matrix clause follows from this discussion. Indeed, as in the structure in (65), *en* is not properly licensed: it absorbs the accusative Case of the verb, while its antecedent is assigned nominative Case by the verb. In addition, what must be accounted for is the possibility of *en* in the infinitive clause. Given that the infinitive clause is controlled by the subject of the matrix clause, subnominal *en* may be licensed by the null subject of the infinitive, PRO, which is not Case-marked and most likely has no grammatical function, since passive verbs do not license a logical subject. To put it another way, since the infinitive verb does not assign nominative Case to PRO, there can be no Case-

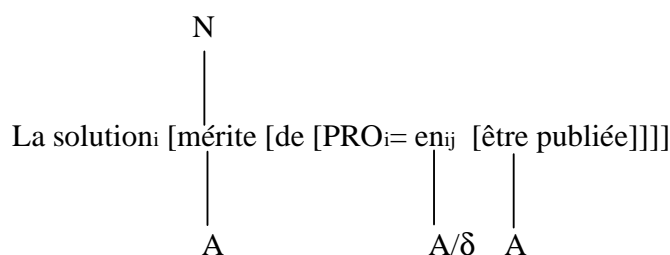


grammatical relation mismatch between *en* and PRO, hence a relatively acceptable structure. The derivation of (71c) is given in (72).

(72) a.



b.



In this section I have presented an account of the syntax of *en*, showing that it obeys the same basic principles as other clitics, namely the OCP and the Clitic Licensing Constraint. When the clitic *en* replaces the *de*-phrase complement of an argument of the verb (subnominal clitic), it is licensed by Case and by coindexation with the head noun. I have shown that the coindexation is successful only if the NP antecedent is an argument bearing a grammatical relation which match the Case of the clitic. In the next section, I will discuss the syntax of the clitic *y*.

### 3.5 The syntax of $y$

### 3.5.1 Feature structure

The clitic *y* replaces an *à*-phrase, generally non-human, expressing mostly (but not only) a locative object. In fact, as a locative clitic it may be derived from any locative preposition and may cliticize onto various verbs which do not have the locative feature. The basic facts are given below.

- (73) a. Je vais à Paris.  
I am going to Paris.  
b. J'y vais.  
I Y am going.
- (74) a. Je pense à ce problème.  
I think to this problem.  
b. J'y pense.  
I Y think.
- (75) a. Jean mange dans ce restaurant.  
Jean eats in this restaurant.  
b. Jean y mange.  
Jean Y eats.
- (76) a. Je vois un chat dans cette chambre.  
I see a cat in this room.  
b. J'y vois un chat.  
I Y see a cat.
- (77) a. Paul nuira à Marie sur la plage.  
Paul will-disturb to Marie at the beach.  
b. Paul y nuira à Marie.  
Paul Y will-disturb to Marie.

In (73) and (74) *y* must be an argument of the verb since the *à*-phrase is obligatory. Indeed, if the *à*-phrase is missing (73a) is deviant (cf. *?je vais*), while (74a) takes a slightly different meaning. On the other hand, in (75) and (76) (examples due to Gross, 1968) the *dans*-phrase is not obligatory at all, and therefore, *y* may not be considered to be an argument of the verb. In (77a) also, the PP (the *sur*-phrase) is not an obligatory argument, nor is the clitic in (77b).

One may note that when *y* is not an obligatory argument as in (75)-(77), it may cliticize either onto an accusative verb (76b) or an oblique-benefactive verb (77b). Since in the present theory clitics are obligatorily licensed by a valid Case of the

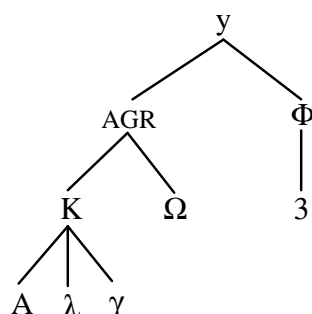
verb, it must be the case that *y* is specified for both accusative and oblique Cases. If so, the question is whether there is a single element which is doubly specified for Case, or two different elements, one specified as accusative and the other as oblique. Let us assume the second alternative, according to which there are two different *y* morphemes.<sup>13</sup>

The clitic *y* must have the same semantic features as the PP it replaces. In (75)-(77), the PP source of *y* clearly expresses the location of an action or event. One can thus assume that *y* is specified for the semantic feature [locative], which will be represented by the symbol  $\lambda$  (lambda). Notice, however, that the locative feature may not be sufficient; in examples (73) and (74) it is clear that the PP is not locative. Rather, it appears to express a goal (somewhat more abstract in (74)), i.e., the direction of a motion. I will thus take *y* to be associated with another semantic role, namely *goal*, represented by the symbol  $\gamma$  (gamma). The feature structure of each morpheme *y* is given in (78).

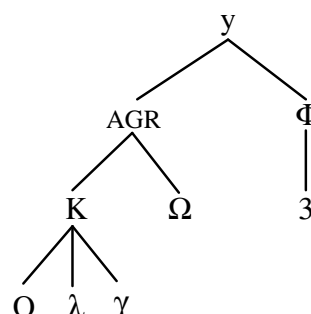
---

<sup>13</sup> Evidence for this assumption is hard to find here because the clitic *y* is not frequently involved in an argumental clitic cluster. This is due to the fact that verbs which productively take two arguments, one of which is a locative, are quite rare in French. One of the best examples of such a verb is provided by the verb *mener* 'to take to'. In the imperative, prescriptive grammarians allow the structure *menez-m' y* and reject *menez-moi-z-y* 'take me to there'. Nevertheless, they recognize that the first structure is a barbarism in current speech, and instead suggest the use of an alternative structure, *menez-y-moi* (see Thomas, 1956). Now, recall that in imperative verbs, first and second person pronouns are not used in their clitic forms, but in their strong forms. I have shown that this is due to the fact that the unsaturated nominative Case of the imperative verb comes to clash with the extra Case of the clitic, triggering a delinking process up to structure preservation. If the clitic *y* were doubly specified for Case, it would have an extra Case which conflicts as well with the nominative Case of the imperative verb. As presumably this clitic does not have a strong form, it would be difficult to repair the structure by a delinking process. Thus, this fact supports our assumption that there are two distinct elements, each specified for one Case, instead of a single element which is doubly specified for Case.

(78) a.



b.



Given (78) the syntax of  $y$  will be accounted for in a principled way, without additional constraints.

### 3.5.2 Subnominal $y$

The use of  $y$ , which is illustrated in (76) and (77), will be referred to as subnominal. Like the *de*-phrase discussed above, the PP in the structures (76a) and (77a) above may be analysed as a predicate of the NP as shown in (79).

(79) a. [un chat<sub>i</sub> [dans la chambre<sub>j</sub>]<sub>i</sub>]b. [Marie<sub>i</sub> [sur la plage<sub>j</sub>]<sub>i</sub>]

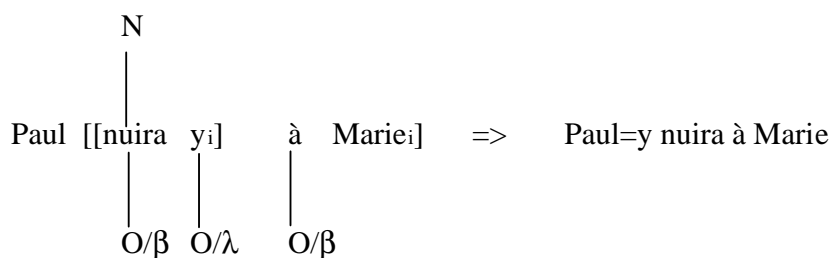
In the light of this, one can posit for (76b) the initial structure shown in (80). This structure is acceptable since every argument is properly licensed. Indeed, the subject clitic is licensed by the nominative Case and by filling a relevant slot in the argument structure of the verb; the NP *un chat* is licensed by the argument structure of the verb; the subnominal clitic is licensed by the accusative Case of the verb and by coindexation with the NP. Since the clitic is specified for accusative Case, it moves under the OCP outside the complement domain of the verb, yielding the correct output. (The irrelevant goal feature of  $y$  is omitted for simplicity.)

(80)



Similarly, the derivation of (77b) proceeds from the structure given in (81); the NP *Marie* is licensed by the lexical structure of the verb, the preposition *à* being an agreement marker. The clitic is licensed by the oblique Case of the verb and by coindexation with the NP, as discussed above. It should be noted that there is no feature clash in this structure, although the thematic features of the clitic are not checked by the verb. Indeed, the clitic absorbs (checks) the oblique Case of the verb, while the preposition discharges its oblique Case onto the NP; the benefactive feature of the verb is absorbed by the benefactive feature of the preposition. The locative (and the goal) features of the clitic are not checked, but they do not clash with any other feature of the verb, and therefore the resulting structure is well-formed.

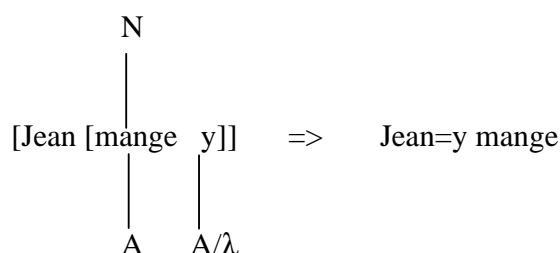
(81)



Notice that when the locative clitic appears in an intransitive structure, it is normally Case-licensed, since the intransitive verb is specified for accusative Case (recall that in my view, there is not necessarily a one-to-one correspondence between Cases and the number of arguments). What is less obvious is whether the clitic is further licensed by something other than Case, as required by (1). The clitic cannot be licensed by filling a slot in the argument structure of the verb, since the verb is intransitive. However, since the verb can license a *dans*-phrase, it surely may license

the clitic corresponding to the latter. Although I will not discuss in detail the licensing of non-arguments, I will point out that the *dans*-phrase in (75a) is arguably coindexed with the VP. Thus, the structure for (75a) would be: [<sub>VP</sub> Jean mange]<sub>i</sub> [<sub>PP</sub> dans ce restaurant]<sub>i</sub>. If this view is correct, the clitic in (82) is licensed by coindexation with the VP.<sup>14</sup>

(82)



### 3.5.3 *Y* as a subcategorized argument

Let us turn now to examples (73) and (74), repeated below as (83) and (84). I claim that the *à*-phrase is an argument of the verb, expressing a location in (83) and a goal in (84). As noted above, this is supported by the fact that if the *à*-phrase is missing, the resulting structure is ungrammatical in the case of (83), and conveys a slightly different meaning in the case of (84).

- (83) a. Je vais à Paris.  
I am going to Paris.  
b. J'y vais.  
I Y am going.
- (84) a. Je pense à ce problème.  
I think at this problem.  
b. J'y pense.  
I Y think.

---

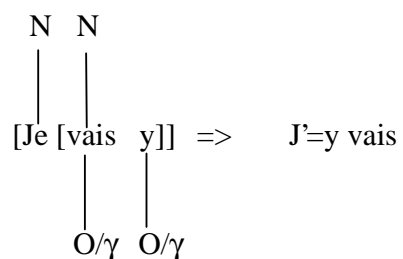
<sup>14</sup> Similarly, an ethical dative is arguably licensed by coindexation with the structure in which it appears and by the latent Case of the verb (see below).

If the  $\grave{a}$ -phrase is an obligatory argument in these examples, it must be treated on a par with the benefactive  $\grave{a}$ -phrase. That is, these instances of the preposition  $\grave{a}$  must be taken to be agreement markers, like the benefactive marker  $\grave{a}$ . The verb *aller* being specified for goal, the argument must agree with it via the goal marker  $\grave{a}$ , consistent with the (a)-clause of the Constraints on Merge. Similarly, since *penser* is a goal verb, its argument must be introduced by this goal marker. Thus, the argument structures of such verbs as *aller* and *penser* must be as given in (85), where the appearance of the preposition is required by the feature of the verb.

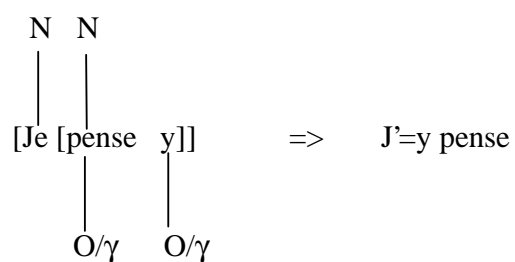
- (85) a.      [X    [aller    [ $\grave{a}$     Y]]  
                      |                    |  
                      O/ $\gamma$             O/ $\gamma$
- b.      [X    [penser    [ $\grave{a}$     Y]]  
                      |                    |  
                      O/ $\gamma$             O/ $\gamma$

Given the argument structure and the feature specification of these verbs, the derivation of (83b) and (84b) is straightforward. Each verb selects the relevant morpheme  $y$  as shown in (86) and (87). (In each example, the irrelevant locative feature of the morpheme  $\grave{a}$  is omitted for simplicity.)

(86)



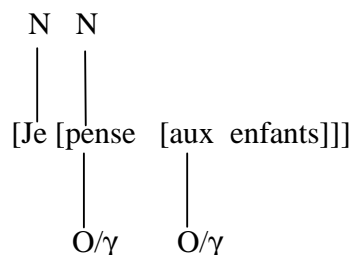
(87)



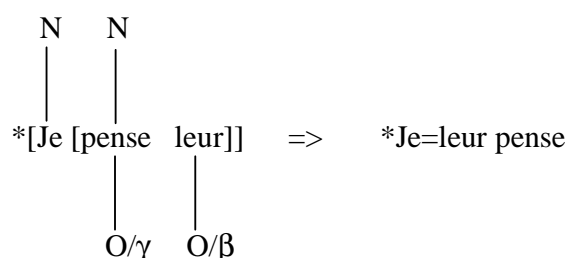
This analysis predicts that verb *penser* cannot take a benefactive clitic, even in the case where its complement is human. If indeed the goal *à*-phrase in the well-formed structure (88) is replaced by a benefactive clitic, an ill-formed structure will result, as illustrated in (89). The ungrammaticality is due to the fact that the goal feature of the verb clashes with the benefactive feature of the clitic.



(88)



(86)



Verb *penser* 'to think' is well-known in the literature precisely because it is incompatible with a benefactive clitic. Current analyses within the GB framework deal with this problem along the following lines (see for instance Borer, 1984). The morpheme *à* is claimed to be either a preposition or a pure Case-marker (cf. Vergnaud, 1974). When it Case-marks an NP, the latter may be replaced by the clitic *lui*; when it is used as a preposition, forming a PP, it is compatible with the clitic *y*, which is assumed to be an adjunct. Under this view, what is needed to account for the above structure is the further claim that *penser* subcategorizes for a PP. This view thus would correctly predict that (89) is ungrammatical, since *penser* cannot take an NP complement.

However, this analysis can be rejected on empirical grounds. Indeed, the central argument in favor of this analysis is the distinction between the Case-marker  $\grave{a}$  and the preposition  $\grave{a}$ . In conjoined structures, proponents of this analysis state that the Case-marker must be repeated before each conjunct, while the preposition must not. However, the facts do not confirm this assumption. As illustrated in (90), the morpheme  $\grave{a}$  must appear before each conjunct, like the benefactive marker (91).

(90) a. Je vais **à** Paris et **à** Lyon.

- b. \*Je vais **à** Paris et Lyon.  
I go to Paris and (to) Lyon.
  - c. Jean pense **à** ce projet et **à** sa voiture.
  - d. \*Jean pense **à** ce projet et sa voiture.  
Jean think to this project and (to) his car.
- (91)
- a. Marie écrit **à** sa mère et **à** sa soeur.
  - b. \*Marie écrit **à** sa mère et sa soeur.  
Marie writes to her mother and (to) her sister.

It turns out that the argument traditionally adduced in favor of the PP character of the complement of *penser* is weak, since it is not confirmed by the facts. Rather the evidence shows that the complement of *penser* is parallel to the complement of a benefactive verb such as *écrire* 'to write'. The verbs *penser* and *écrire* have different Case specifications, and therefore take different agreement markers. The picture is somewhat obscured because the agreement markers are homophonous: the *à* of *penser* is a goal marker ([O,γ]), while the *à* of *écrire* is a benefactive marker ([O,β]).

Summing up, I have suggested that there are two different morphemes *y*, one specified for accusative Case and the other for oblique Case. With respect to thematic roles, each morpheme is specified both as goal and locative. I have shown that *y* may cliticize either onto goal verbs or onto any verb taking a locative complement or a locative adjunct, irrespective of Case. Now, the use of *y* is not always successful; in particular it may not cooccur with another clitic when it is not an obligatory argument of the verb (cf. \**Il lui y nuira*). Such a case (in fact, not peculiar to *y*) is predicted by the Argument Licensing Constraint (see (1)) and will be dealt with in the next section.

### 3.6 On double clitics

In my view, a pronominal element is a clitic if it is specified for Case. Since a clitic must be licensed by a valid verb Case, a simple transitive verb, which is specified for nominative and accusative, may at most take one clitic per Case. Furthermore, since a clitic may occupy a slot in the argument structure of a verb, it

follows that it need not cooccur with the NP it refers to, at least as long as all other constraints are satisfied (but see the discussion of clitic doubling in Spanish in chapter 4).

However, as seen above, a clitic can cooccur with an NP, both competing for a single slot in the argument structure of a verb. In such a case, the clitic is licensed by the Case of the verb and by coindexation with the NP competitor. This is possible, because in the proposed view an NP need not be Case-licensed, while the clitic need not fill a slot in the argument structure of a verb. It is important to note that this does not entail that any clitic can cooccur with any NP. Indeed, I have shown that a clitic can be licensed by an NP only if (a) the case of the clitic is compatible with the grammatical relation of the NP and (b) the clitic is in an anaphoric relation with the NP. This section will provide additional evidence for (a) and (b). I will show that the analysis presented so far, in particular the Argument Licensing Constraint, rules out the appearance of two clitics even in contexts where there are two 'subnominal' phrases which can be cliticized separately. Also, I show immediately that only a clitic which has an anaphoric property can be licensed by an antecedent occupying a slot in the argument structure of a verb.

Recall that I have argued that only three clitics can be licensed by coindexation, namely *se*, *en*, and *y*. The last two are assumed to be both pronouns and anaphors, a property that allows them to be licensed by coindexation. In order to show that this is the case, one can argue that other clitics cannot be coindexed with an NP in the relevant context. Consider the third person clitic *le*. It may function as a neuter clitic and may replace a subordinate clause. For instance, the subordinate clause in (92a) may be replaced by the neuter clitic *le*, as shown in (92b). Since the clitic *le* is not an anaphor, it is expected that it may not be licensed by the head of the relative clause it replaces, even though there exists between them a predication-like relationship. This is borne out, as shown in (93). In (93a), the relative clause is coindexed with the head noun, as is the clitic in (93b); however, the latter is ungrammatical. By comparison, the *de*-phrase in (94a) is coindexed with the NP; when it is replaced by the clitic *en*, a well-formed structure is derived.

- (92) a. Jean sait que Paul a vu Marie.

- Jean knows that Paul has seen Marie.
- b. Jean le sait.  
Jean knows it.
- (93) a. Je connais la fille<sub>i</sub> [qui aime Paul]<sub>i</sub>  
I know the girl who likes Paul  
b. \*Je le<sub>i</sub> connais la fille<sub>i</sub>  
I know him the girl.
- (94) a. Je connais l'auteur<sub>i</sub> [de ce livre]<sub>i</sub>  
b. J' ~~e~~n connais l' auteur<sub>i</sub>

Given the ill-formedness of (93b), one may conclude that *le* is not an anaphor like *en*, *y*, and *se*, and therefore cannot be licensed by an NP antecedent. Below I provide some evidence of the importance of Case.

### 3.6.1 On the incompatibility of two subnominal clitics

Gross (1968) observes that both the NP and the *dans*-phrase in (95a) may be a source for subnominal *en* and *y*, as illustrated in (95b,c). However, these subnominal clitics cannot cooccur, as illustrated in (96).<sup>15</sup>

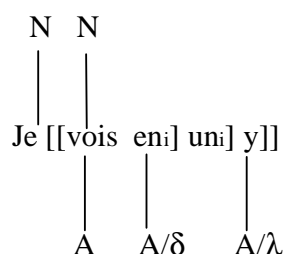
- (95) a. Je vois un chat dans la chambre.  
I see one cat in the room.  
b. J'**en** vois un dans la chambre.  
I EN see one in the room.  
c. J'**y** vois un chat.  
I Y see one cat.
- (96) \*J'**y en** vois un. (whatever the order of the clitics)  
I Y EN see one.

---

<sup>15</sup> It should be noted that *en* and *y* may cooccur only in the existential use of the verb *avoir* 'to have'; thus, alongside *il y a* 'there is', *il y en a* 'there is some' is possible. I shall not discuss this fact here.

Notice that the simple transitive verb is specified for a single accusative Case and a latent oblique-benefactive Case (see below). Since in my view a clitic must be licensed by a valid Case of the verb, it follows that both subnominal clitics cannot appear at the same time in the structure. The initial representation of (96) is given in (97), where the clitic *en* is coindexed with the numeral *un*. As can be seen, the first clitic absorbs the single accusative Case of the verb, and therefore the second clitic, *y*, cannot be licensed. Therefore, the structure is ill-formed, as expected.

(97)

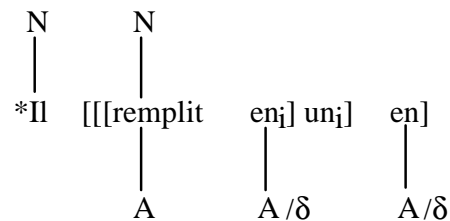


Gross (1968) further observes that in (98a), the NP complement of the numeral *un* 'on e' and the *de*-phrase may be replaced in turn by subnominal *en*, as shown in (98b,c). As in the previous example, both clitics cannot cooccur at the same time (98d).

- (98)
- a. Il remplit un verre de ce vin.  
He fills a glass with this wine.
  - b. Il **en** remplit un verre.  
He EN fills a glass.
  - c. Il **en** remplit un de ce vin.  
He EN fills one with this wine.
  - d. \*Il **en en** remplit un.  
He EN EN fills up one.

Since *remplir* 'to fill' is a simple transitive verb having only one accusative Case, it is expected that it cannot license two accusative clitics. Indeed, as shown in (99), the accusative Case of the second clitic cannot be checked by the verb, due to the presence of the first clitic, hence the ill-formedness of the structure.

(99)

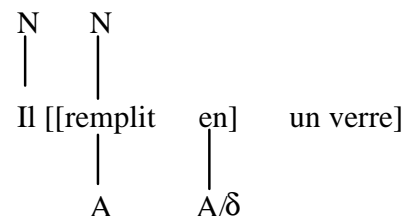


Consider the derivation of (98b,c). In order to see the origin of each subnominal *en*, the constituent structure of (98a) is given in (100), where the relevant phrases are underlined. When the *de*-phrase is replaced by a clitic, the derivation must be as shown in (101), where *en* is interpreted as a subnominal complement of *un verre*. As can be seen, both compete for the same argument slot; the NP occupies the slot in the argument structure and licenses the clitic by reference; the latter is further licensed by the accusative Case of the verb, as required by (1).

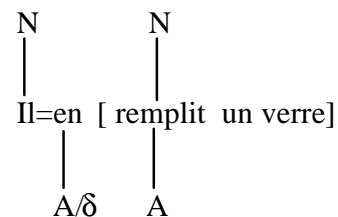
(100)

$$[Il \ [remplit \ [un \ \underline{verre} \ \underline{de \ ce \ vin}]]]]$$

(101) a.



b.



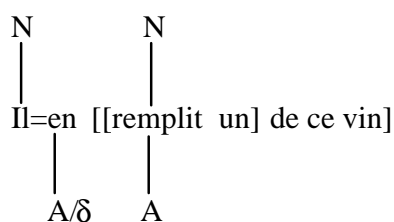
When the bare NP is replaced by subnominal *en*, the structure in (102) is derived.<sup>16</sup> In this structure, *en* is interpreted as the complement of the numeral *un*.

---

<sup>16</sup> Note that the numeral phrase behaves as if it contained an underlying preposition *de*. Thus, the sentence *il remplit un verre*, where *un* is a numeral is analysed as *Il remplit un [de] verre*. (See a discussion of this in Kayne 1975.)

The latter is licensed by the lexical structure of the verb, the Case of the verb being absorbed by *en*.

(102)



To conclude, (1) correctly predicts that in French, clitics and NPs must be in complementary distribution in most cases. If a clitic cannot be coindexed properly with an argument of a verb, it must occupy a slot in the argument structure of the verb. In either case, it must be Case-licensed by the verb. Only the reflexive *se* and the subnominals *en* and *y* may cooccur with NPs. I have argued that this is possible because they all are anaphors, which must be coindexed with some element in the structure. Notice that in Spanish, a clitic may cooccur with the NP it refers to, suggesting that there is no complementary distribution. I will show in the next chapter that clitic doubling in this language is used as a strategy to overcome a violation of the Constraints on Merge. This use of clitics will be shown to be also consistent with the Clitic Licensing Condition, as discussed above.

### 3.7. Latent Case

The finding that a verb licenses a clitic only if it has a valid Case is seemingly challenged by the so-called nonlexical dative, which is described in Leclère (1976). This includes extended datives, ethical datives, of which affected datives are a subclass, and inalienable possession. Focussing on extended datives, I will argue that all French verbs are specified for a latent oblique-benefactive Case, which is activated in a particular context.

As Leclère observes, most French verbs may take a complement of the type  $\hat{a}$ -NP, where NP is obligatorily human. This is to be distinguished from the case where the  $\hat{a}$ -phrase is an obligatory argument of the verb. The examples in (103) illustrate this fact (see Leclère, 1976; glosses added).

- (103) a. Paul a fabriqué une table (à Marie).  
Paul made a table (for Marie).  
b. Paul a recousu ce bouton (à Marie).  
Paul has sewed this button back on (for Marie).  
c. Paul a ouvert cette porte (à Marie).  
Paul opened this door (for Marie).

The optional *à*-phrase argument in (103) may be replaced by the relevant clitic, as exemplified in (104).

- (104) a. Paul **lui** a fabriqué une table, (à Marie).  
Paul made a table for her.  
b. Paul **lui** a recousu ce bouton, (à Marie).  
Paul sewed back on this button for her.  
c. Paul **lui** a ouvert cette porte, (à Marie).  
Paul opened this door for her.

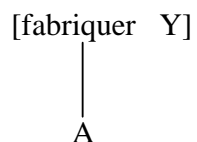
Since these verbs usually take a single direct object, the question arises as to how the dative clitic is licensed.<sup>17</sup> I would like to suggest that oblique-benefactive Case appears latently in French (and Romance) verbs. In accordance with common practice in nonlinear phonology, I assume that a latent feature is a floating one, that is, a feature which is not linked with a morpheme via an association line (see chapter 1). On this view, a latent oblique-benefactive Case is activated only in the relevant context, namely whenever an extra argument is added. Thus, given the latent Case hypothesis, a simple transitive verb such as *fabriquer* 'to make' initially has the representation in (105a), which will yield the representation in (105b), once the argument Z is computed (the external argument is set aside). Notice that in my view the grammar need not specify that the extra argument must be an *à*-phrase; as in the case of simple benefactive verbs, the agreement marker *à* is computed with the extra argument in order to satisfy the Constraints on Merge.

---

<sup>17</sup> For the semantics of these constructions, see Leclère (1976).



(105) a. (O)

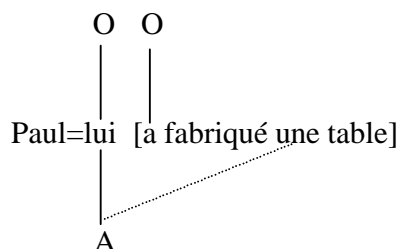
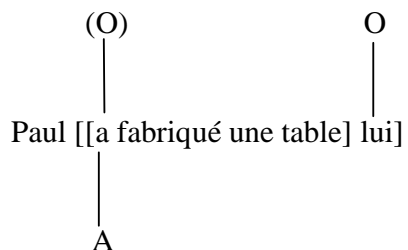


b. (O) O

[[fabriquer Y] à Z]  
|  
A

Under the latent Case hypothesis, the derivation of (104a) is straightforward: the latent Case of the verb is transferred to the auxiliary, as discussed in chapter 2, allowing the clitic to be licensed, as shown in (106a); movement takes place under the OCP as usual, yielding (106b).

(106) a.



### 3.7.1 Inalienable possession

As is well-known, in many languages inalienable possession displays a special syntax. When the possessee is inalienable, only a construction with a benefactive clitic can be used; compare (107a) and (107b). On the other hand, when the possessee is not inalienable, only a possessive can be used; compare (108a) and (108b). I claim that the oblique-benefactive clitic which is used in inalienable constructions represents the possessor and originates from a noun complement in the form of a *de*-phrase (see also Langacker, 1968, and Tremblay, 1991). This is supported by the fact that the clitic in (107a) may be replaced by a *de*-phrase in standard French (109a), but not by an *à*-phrase, which is characteristic of colloquial French (109b).<sup>18</sup>

- (107) a. Marie lui a cassé le bras.  
           Marie him has broken the arm.  
           ' Marie broke his arm.'
- b. \*Marie a cassé son bras.  
           Marie broke his arm.
- (108) a. Marie a cassé sa montre.  
           Marie broke his watch.
- b. \*Marie lui a cassé la montre.  
           Marie him broke the watch.
- (109) a. Marie a cassé [le bras [de Paul]]       (standard)  
           Marie has broken the arm of Paul.  
           ' Marie broke Paul' s arm.'
- b. Marie a cassé [le bras [à Paul]]       (popular)

Now, given that the clitic which corresponds to a *de*-phrase is *en*, as we know, the fact that in standard French (107a) is related to (109a), but not to (109b), must be accounted for. I suggest that (standard) French has a special morpheme for expressing inalienable possession, which is homophonous with the morpheme *de*

---

<sup>18</sup> I will not attempt to account for the fact that inalienable constructions are incompatible with the possessive as in (107b); but see Tremblay (1991) for an account in the GB framework.

[A/δ] discussed above. I assume that this instance of *de* ('inalienable*de*') is specified for the features [O,β]. Thus, whenever the possessee is inalienable, the Blocking Principle will force the use of the relevant instance of *de*. For instance, in (110a,b) there are two different occurrences of the morpheme *de*; in (110a) it is inalienable *de*, while in (110b) it is A/δ-*de*.

- (110) a. [Le bras [de<O,β> Paul]]  
 b. [La montre [de<A,δ> Paul]]

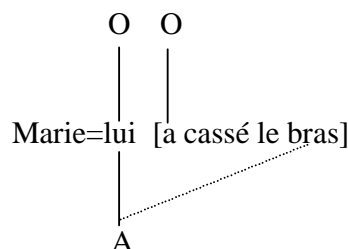
Under this hypothesis, the syntax of inalienable possession can be accounted for straightforwardly, as we will see below. Consider the feature structure of the relevant elements in (109a), repeated as (111). Since inalienable *de* is specified for [O/β], the latent Case of the verb is activated. As a result, the simple transitive verb is reanalysed as a ditransitive verb, i.e. the *de*-phrase, originally a subnominal complement, becomes the indirect complement of the verb.

- (111)
- |       |             |         |            |  |
|-------|-------------|---------|------------|--|
|       | (O/β)       |         | O/β        |  |
|       |             |         |            |  |
| Marie | [[ a cassé] | le bras | [de Paul]] |  |
|       |             |         |            |  |
|       | A           |         |            |  |

As an argument, the *de*-phrase may be replaced by a clitic. Obviously, since *de* is specified as oblique-benefactive, the only clitic it is compatible with is an oblique-benefactive clitic. This is shown in (112a); OCP triggers the movement of the clitic, yielding the output (112b).

- (112) a.
- |       |                    |      |   |  |
|-------|--------------------|------|---|--|
|       | (O)                |      | O |  |
|       |                    |      |   |  |
| Marie | [[a cassé le bras] | lui] |   |  |
|       |                    |      |   |  |
|       | A                  |      |   |  |

b.



Let us now turn to evidence for the existence of the inalienable morpheme *de*. We know that subnominal *en* usually replaces an inanimate *de*-phrase, irrespective of the inalienable character of the NP head. Thus in (113a) *en* arises from an inanimate subnominal phrase, *du livre*, though the NP *page* is inalienable to the book. When the *de*-phrase is animate, *en* cannot be used, as expected, since it is specified for [thing] (113b). Interestingly, (113b) cannot be rescued by the use of the benefactive clitic, as shown in (113c), which is otherwise acceptable under the extended dative reading. Expectedly, the use of the benefactive clitic cannot be successful in (113c), since the possessee is not inalienable. Indeniably, if there were not an inalienable *de*, (113c) would be grammatical. Since this is not the case, the analysis presented above is borne out.

- |       |    |   |                        |
|-------|----|---|------------------------|
| (113) | a. | Marie <b>en</b> a déchiré la première page. | (la page du livre)     |
|       |    | Marie EN has torn the first page.           | (the page of the book) |
|       | b. | *Marie <b>en</b> a déchiré le livre.        | (le livre de Paul)     |
|       |    | Marie EN has torn the book.                 | (Paul' s book)         |
|       | c. | *Marie <b>lui</b> a déchiré le livre.       | (good extended dative) |
|       |    | Marie him has torn the book.                |                        |

Notice that in this analysis *de* is taken to be an inalienable morpheme, but not a morpheme which is specified for some inalienable feature. In other words, the inalienable feature is the defining property of the morpheme and therefore constitutes the root node. Inalienable constructions also exist in Spanish, as discussed in the next chapter. However, there are reasons to believe that this language displays an inalienable feature in the tree structure.

To sum up, I have argued that French verbs are specified for a latent Oblique-benefactive Case. The latent Case is responsible for the licensing of extended datives, which are not required by the lexical structure of the verb. Inalienable possession is shown to be licensed by the latent Case; the inalienable possession marker, which is an instance of the preposition *de*, is specified as benefactive, and therefore is compatible with the benefactive clitic *lui*.

### 3.8 Summary

In this chapter, I have discussed the syntax of the clitics *se*, *on*, *en*, and *y*. Each of them is shown to obey the basic principles introduced in the preceding chapters. *Se* is shown to be specified for both oblique-benefactive Case and accusative Case; *on* is specified for nominative Case and refers to an indefinite referent. The clitic *en* is shown to be specified for accusative Case and for a set of semantic features, which I refer to as delta. As for the clitic *y*, I have argued that it represents two different morphemes: one is specified for Accusative Case, while the other is specified for Oblique Case; both instances of *y* are specified for the semantic features locative ( $\lambda$ ) and goal ( $\gamma$ ).

These four clitics are shown to have special features which distinguish them from those discussed in the preceding chapter. Because of the indefinite feature, *on* cannot control another pronominal. Because of its reflexive property, the clitic *se* is used in various types of constructions, including middle constructions, for which I have presented a lexical-based analysis. The clitics *en* and *y* were argued to have both pronominal and anaphoric properties. As pronouns, they are licensed by the verb's Case and by the lexical structure of the verb. As anaphors they are licensed by the verb's Case and by coindexation with an NP argument in the syntactic structure, a property which is shared with reflexive *se*.

In my view an NP may be licensed either by Case or by the lexical structure of the verb, although in the most common case it is redundantly licensed in both fashions. A clitic is licensed obligatorily by Case and either by the lexical structure of the verb or by coindexation with a relevant NP. When an NP is used with a clitic, as in the case of subnominals *en* and *y*, the NP and the clitic must be coindexed, and the

Case of the latter must be conformed to the grammatical relation borne by NP, otherwise the structure is ill-formed. Further, I have argued that in French, transitive verbs are specified for a latent oblique-benefactive Case, which is activated in extended dative constructions. This latent Case allows the benefactive clitic which is used in inalienable constructions to be licensed by the verb.

In the next chapter, I will consider the pronoun systems of various languages, including Spanish and Italian, showing that they provide evidence for the analysis developed so far.

## CHAPTER IV

### CROSS-LINGUISTIC EVIDENCE AND EXTENSIONS

#### 4.1 Introduction

The nonlinear approach I argue for succeeds in accounting for various aspects of clitics in a simple way. In this chapter, I present a cross-linguistic analysis of cliticization. My main interest will be the pronoun systems of English, Haitian Creole, Spanish and Italian. Based on the analysis developed so far, I will show why there are no (special) clitics in English and Haitian Creole. However, the same cannot be said of Spanish and Italian, since special clitics abound in both languages, just as in French. Indeed, I will show that my analysis predicts clearly that in Spanish and Italian most personal pronouns must be special clitics.

Furthermore, I will show that one of the most celebrated features of Spanish clitics, namely the so-called spurious *se*, is not an idiosyncratic phenomenon, but rather an expected consequence of the constraints on licensing of clitics. Specifically, it will be shown that the formulation of a spurious *se* rule is not needed. Instead, it will be shown that a spurious *se* is obtained automatically by a delinking process, because the second clitic cannot be licenced otherwise. Moreover, opaque clitics in Italian will be shown to follow from this constraint on licensing.

Another interesting feature of Spanish grammar to be considered is the phenomenon of clitic doubling. This phenomenon will be discussed in both standard and dialectal Spanish. I will argue that clitic doubling is mainly a repair strategy that allows a feature clash to be avoided during the building of the structure. It will be shown that the clitic is used when an argument requires two agreement morphemes which bear incompatible features.

In Spanish and Italian, clitics normally follow infinitive verbs (encliticization), and may climb to the left of a matrix verb. On the basis of Spanish facts, it is argued that infinitive verbs and clitics are in a superiority effect relationship. I will proceed to provide an account of superiority effect phenomena in terms of features, claiming that Spanish clitics, as well as infinitive verbs, are specified for a kind of feature which is usually found in *wh*-operators.

Finally, certain aspects of the syntax of *wh*-operators and negation will be discussed, focussing on French facts. It will be shown that *wh*-operators are specified for Case, and therefore are expected to move under the OCP out of the complement domain of the verb. They differ from pronominal clitics in that they cannot be head-adjoined, a property they owe to a feature I will refer to as  $\pi$ . In order to further support the analysis of clitic climbing in Spanish, superiority effects will be discussed in English; it is shown that these phenomena can be handled with the same analysis developed to account for French operators and Spanish clitic climbing phenomena. Certain aspects of negation in French will be discussed as well. Given that in standard French, negation is generally expressed via a two-word morpheme, it is suggested that one word bears the feature  $\pi$ , and the other bears nominative Case. While the Case-specified morpheme (the word *pas*) moves to avoid the well-known phonological constraint on no crossing lines, the  $\pi$ -specified word (*ne*) can never move for some reason to be discussed.

It should be noted that only a single constraint is introduced in this chapter, namely the non-commutativity of operators, which accounts for superiority effects. In addition, the hypothesis that the feature  $\pi$  makes the morpheme bearing it incompatible with head-adjunction is suggested. All we need to account for the cross-linguistic variations just mentioned is the morpheme and feature inventories of each language.

## 4.2 English personal pronouns

English personal pronouns belong to two groups: the first group includes pronouns which can be used as subjects; and the second group includes pronouns which can only be used as objects of a verb (direct object) or a preposition (indirect



object). There is syncretism in second person pronouns, singular and plural, as well as in the third person neuter (3c); this is illustrated in the following (descriptive) Table.

**Table 4.1** English personal pronouns.

<b>Persons</b>	<b>Subject</b>	<b>object</b>
1.	I	me
2.	you	you
3a.	he	him
3b.	she	her
3c.	it	it
1.	we	us
2.	you	you
3.	they	them

Current analyses in GB theory assume that English subject pronouns (i.e. the first group) are specified for nominative Case, while object pronouns (the second group) are specified for objective Case. If this claim is true, it constitutes a strong counterexample for the OCP analysis, as these pronouns never appear before the verb, unlike their French counterparts. I will in fact show that English object pronouns are unspecified for Case.

#### **4.2.1 Object pronouns**

In order to show that English object pronouns are unspecified for Case, one can compare them with NPs or with French strong pronouns, which are unspecified for Case, as we know. Recall that the position of an element in the structure depends upon its Case specification; if an element bears no Case, it is likely to be compatible with more than one syntactic position. Thus, a Caseless element, such as an NP, may either precede or follow the verb selecting it; in the first situation it has the function of subject, while in the second it has the function of object, as illustrated in (1).

- (1) a. Paul likes Mary.  
b. Mary likes Paul.

If English object pronouns are unspecified for Case, they are expected to have the same behaviour as NPs. In particular, they may appear either in subject or object positions. This prediction is borne out, as shown in (2).

- (2) a. Mary likes him.  
b. Mary and him bought a house.  
c. Mary gave a gift to him.  
d. Mary believes him to be happy.

These facts are consistent with the Constraints on Merge. In (2a) the verb is specified for accusative Case while the pronoun is not specified for this Case (weak agreement); in (2b) the pronoun is unspecified for Case, while the relevant Case of the verb is nominative (again weak agreement). In (2c) the pronoun is the object of the preposition. Whatever the Case and the semantic feature of the latter, it is compatible with the unspecified pronoun (weak agreement). In (2d) the pronoun is the subject of the infinitive verb, which is unspecified for nominative Case; therefore they agree with each other by default (default agreement).

Furthermore, like NPs and strong pronouns in French, English object pronouns may be used alone, for example as an answer to a question, as illustrated in (3).

- (3) a. Who wants to go there? Me/\*I.  
b. Who saw the burglar? Him/\*he.

Like French strong pronouns, only object pronouns in modern English may appear in a coordinated structure, either as subject or object of a verb, as shown in (4).<sup>1</sup>

- (4) a. Mary and me/\*I will go to Paris.  
b. Paul met Mary and him/\*he in Paris.  
c. Mary and him bought a house. (also (2b))

---

<sup>1</sup> In (4a), the form *I* is correct (and required) in formal English, a fact that I will not discuss here.

These facts permit us to conclude that English object pronouns are not specified for Case. They cannot be considered to be clitics, and therefore they should not move under the OCP outside the complement domain of the verb selecting them.

#### 4.2.2 Subject pronouns

Unlike object pronouns, subject pronouns have a more limited distribution in English. They can only appear in subject position. Indeed, in all the examples discussed above (except for (2b); see below), if the object pronoun is replaced by a subject pronoun, this will result in an ill-formed sentence.

If one assumes that English subject pronouns are specified for nominative Case, along with traditional analyses, the fact that their distribution is more limited may be readily explained by the Blocking Principle (BP), which forces the morpheme which is specified for the relevant Case to be computed in priority. We may note immediately that an apparent counterexample to this principle comes from the sentence in (2b), where an unspecified pronoun is used as subject of the verb. In fact, the pronoun is first merged with the NP *Mary* in order to form the coordinated phrase. By the COM, the third person pronoun must have compatible features with the NP, which prevents it from taking the Case-specified form *he*.

However, a challenging problem for the analysis according to which English subject pronouns are specified for nominative Case comes from the fact that these pronouns may be separated from the verb by a parenthetical clause, as exemplified in (5). We have seen that Case fusion automatically occurs, which excludes the presence of an intervening time-interval (gemination effect). To the extent that the fusion process is correct, the feature specification of subject pronouns in English may not be quite parallel to their French counterparts.

- (5) He, it appears, likes pop music.

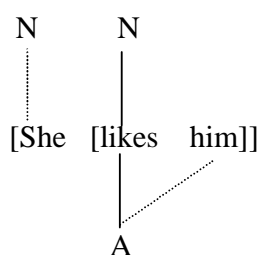
Apparently, there is a dilemma here: if one assumes that subject pronouns are specified for Case, one must reject the fusion analysis on the basis of examples like (5). On the other hand, if one assumes that subject pronouns are unspecified for

Case, like object pronouns, one cannot explain why their distribution is more limited than object pronouns and, under such an analysis, there would be no principled way to exclude a pronoun such as *he* as object of a verb or a preposition.

I would like to suggest that English subject pronouns are specified for a latent nominative Case, i.e. a floating Case, as discussed in chapter 3. An element with a floating Case must have more limited distribution than an unspecified one. In particular, the Blocking Principle operates normally, excluding the unspecified form of the paradigm as subject. However, since the Case is not anchored to the morpheme with an association line, it is reasonable to assume that the fusion process (if any) may not trigger a gemination effect.

Thus, consider the derivation of a simple sentence such as *she likes him*. The verb is merged with the object, here the unspecified third person pronoun, yielding [likes him]. Since the verb is specified for accusative Case, the use of a third person pronoun specified for nominative Case is ruled out by the COM. The computation continues, merging the structure already constructed with the subject, yielding [she [likes him]]. At this last stage, N being the relevant Case of the verb, only a nominative pronoun may be selected by virtue of the BP. This is illustrated in (6): the verb assigns accusative Case to the object pronoun by spreading, while the floating nominative Case of the subject pronoun absorbs the nominative Case of the verb and both ultimately undergo fusion.

(6)



Since the subject pronoun has a floating Case, the adjunction of a structure with several timing intervals between it and the verb is expected to be possible.

To conclude, English object pronouns are unspecified for Case and as such can undergo neither movement nor fusion. English subject pronouns are not real clitics, even though they are specified for nominative Case, a peculiarity due to the floating nature of their Case. Thus the syntax of English personal pronouns, in particular the object pronouns, supports the OCP. An example of a Case-free pronoun system is provided by Haitian Creole. The next section is devoted to this.

### 4.3 Haitian Creole

We have seen that the paradigm of French personal pronouns displays three types of Case, namely nominative, accusative and oblique, while in English there is only one floating nominative Case. In the three-Case system of French, one may find up to four morphemes for the third person: one for each Case, the other being unspecified for Case. In English, there are only two morphemes, one is specified for (a floating)

Case while the other is not. In Haitian Creole (HC), there are no Case-specified pronouns, and there is a single morpheme per person, and it covers all grammatical functions. This paradigm is given in Table 4.2. (First and second plural pronouns are syncretic.)

**Table 4.2** Personal pronouns in Haitian Creole.<sup>2</sup>

<b>Persons</b>	<b>Subject/ Object</b>	<b>Reduced forms</b>	<b>Translation</b>
1.	mwen	m	I, me
2.	ou	w	you
3.	li	l	s/he, it, him, her
1., 2.	nou	n	you, we, us
3.	yo	y	they, them

It has been noted (cf. Cadely, 1994) that the reduced forms are clitics to the extent that they cannot stand on their own. Although this observation is correct, this is not a sufficient reason to consider an element a clitic. The reduced forms are not in

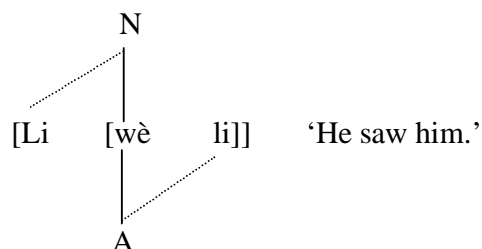
<sup>2</sup> These morphemes are also used as possessives: *liv li* 'his book'.

complementary distribution with the full forms; they are, indeed, simple phonological variants of the full forms, parallel to the clitics ' *s* (as in *is*, *has*) in English. Furthermore, their appearance depends upon whether they can be syllabified with an adjacent word. Consider the examples in (7) (Reduced forms are in bold face).

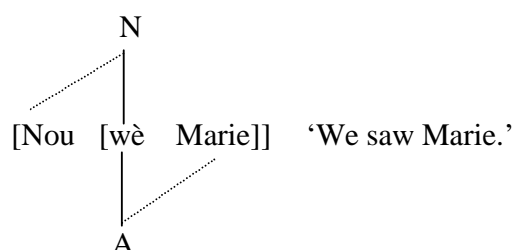
- (7)
- a. Marie wè li. / Marie wè-**l**.  
Marie saw him/her.
  - b. Marie bat li. / \*Marie bat **l**.  
Marie beats him/her/it.
  - c. Li wè Marie. / \*Lwè Marie.  
He/she saw Marie.
  - d. Nou achte li. / **N**-achte-**l**.  
We bought it.
  - e. Yo Konn Marie. / \*Y konn Marie.  
They know Mary.
  - f. Yo achte kay la. / **y**-achte kay la.  
They bought the house.
  - g. Ou ale! / **W**-ale!  
You left!

As can be seen from the examples in (7), the use of the reduced form is dependent on the shape of the verb, among other things. For instance in (7b) the reduced form is not possible with a verb which ends with a consonant. However, the interesting fact about this paradigm is that a given form may be either subject or object. The simplest assumption here is that these pronouns do not bear any Case feature, just like ordinary NPs. Under this assumption, they must be assigned Case by the verb, as shown below.

(8) a.



b.



Clearly personal pronouns in HC are Caseless. The present theory correctly predicts that in this language personal pronouns may not be special clitics. In particular, they may not undergo any process triggered by the OCP, either fusion or movement. I turn now to positive evidence for my analysis, showing that when in a given language pronouns are specified for Case, they are generally special clitics, as in French. Spanish and Italian are examples of such languages.

#### 4.4 Spanish

##### 4.4.1 Morpheme inventory and features

Spanish personal pronouns are given in Table 4.3, as they appear in many teaching grammars. Such grammars distinguish subject pronouns, direct object pronouns (DO), indirect object pronouns (IO), reflexive pronouns and pronouns which are objects of a preposition. There are also polite forms, which are set aside for simplicity.

**Table 4.3** Spanish personal pronouns.

<b>Persons</b>	<b>Subject</b>	<b>DO</b>	<b>IO</b>	<b>Reflex.</b>	<b>Prep. Obj.</b>
1.	yo	me	me	me	mí
2.	tú	te	te	te	ti
3. masc.	él	lo, le	le	se	él, sí (refl.)
3. fem.	ella	la	le	se	ella, sí
3. neut.	ello	lo	le	se	ello, sí
1.	nosotros	nos	nos	nos	nosotros
2.	vosotros	os	os	os	vosotros
3. masc.	ellos	los	les	se	ellos
3. fem.	ellas	las	les	se	ellas

The Spanish pronoun system presents many similarities with that of French. However, it contains certain particular facts which are not found in French. First, Spanish being a pro-drop language, subject pronouns are not overt in the normal case (I will take the nonovert subject to be a nonovert pronoun, *pro*, a standard assumption in the literature); when they are used, they indicate emphasis. Second, subject pronouns and the pronouns which are object of a preposition are formally identical, except the first and second person singular pronouns. Both will be referred to as strong pronouns. (Note also that the reflexive *se* becomes *sí* after a preposition, a fact that I set aside here.)

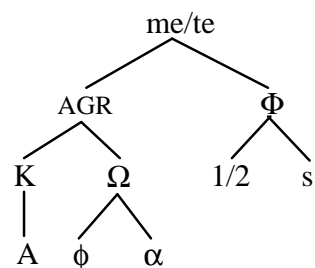
#### 4.4.1.1 Feature structure of the paradigm

I will assume that Spanish personal pronouns have the feature structure given in (9), where  $\phi$  stands for feature specificity (see, for instance, Suñer, 1988) and  $\alpha$  for animate. The weak pronouns are assumed to exist in two different morphemes, in contrast to French and Italian, which have a single morpheme doubly specified for Case (Italian pronouns are discussed below). Plural pronouns are not shown here; but

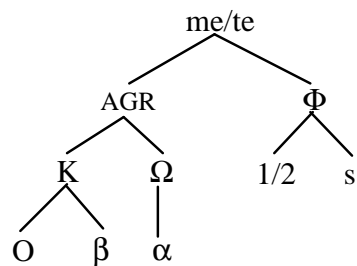


since they pattern like singular pronouns, one may assume that they are specified for the same features as their singular counterparts.<sup>3</sup>

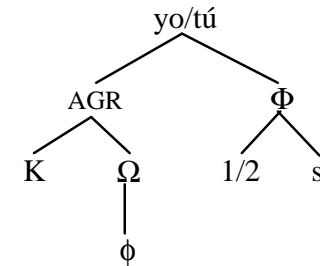
(9)a.



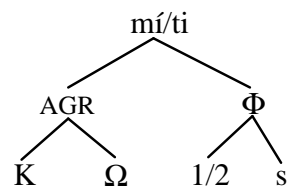
b.



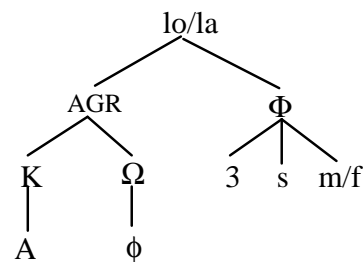
c.



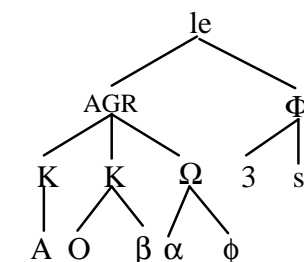
d.



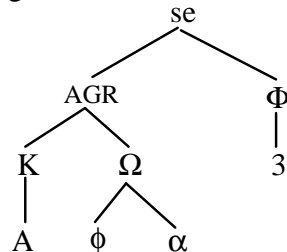
e.



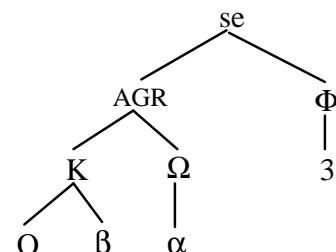
f.



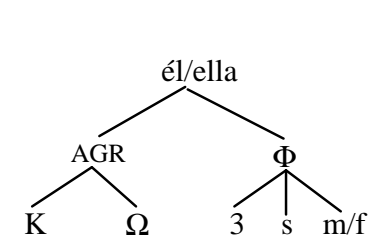
g.



h.



i.



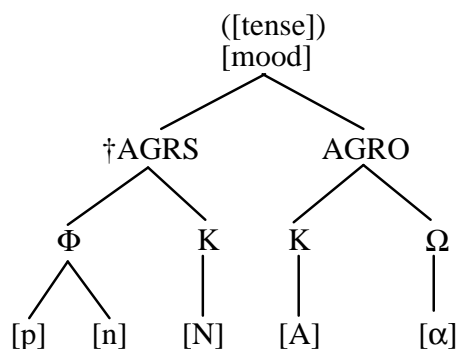
The correctness of these feature structures will become clear later on, as I will show that a number of striking facts in Spanish grammar follow straightforwardly, given the general principles discussed in the preceding chapters. Before doing so, something must be said about the feature structure of Spanish verbs, which is somewhat different from the feature structure of French verbs.

<sup>3</sup> It is claimed below that Spanish weak pronouns are operator-elements, i.e. they are specified for features which are found in *wh*-words and negation, namely  $\pi$  and  $\sigma$ . These features are omitted in the representation at this point.

#### 4.4.1.2 Feature structure of Spanish verbs

One of the most obvious (and well-known) properties of Spanish grammar is the fact that animate arguments must be introduced by the morpheme *a*, which I take to be an animacy marker, consistent with standard assumptions.<sup>4</sup> Since this morpheme is obligatory whenever the internal argument (generally the direct object) is animate, I suppose that all Spanish transitive verbs are specified for the feature [animate]. This feature is a dependent of the omega node, as diagrammed in (10) for a simple tensed transitive verb (with NP arguments).

(10)



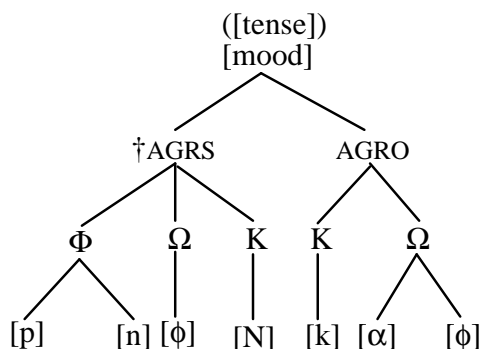
A less obvious feature of Spanish grammar is specificity ( $\phi$ ). This feature is relevant only with pronominal arguments, at least in most dialects (see below). For a reason that will be discussed below, I will assume that with respect to specificity Spanish verbs distinguish between pronominal and non-pronominal arguments. That is, the AGR nodes are sensitive to whether the argument is an NP or a pronoun. If the argument is a pronoun, it must be specific. The specificity feature, which will be motivated under the discussion of clitic doubling, is also a dependent of the omega node, as illustrated in (11) for a DO verb, the internal arguments of which are pronouns.

---

<sup>4</sup> For instance, compare the sentences in (i), where the direct object is animate and inanimate respectively.

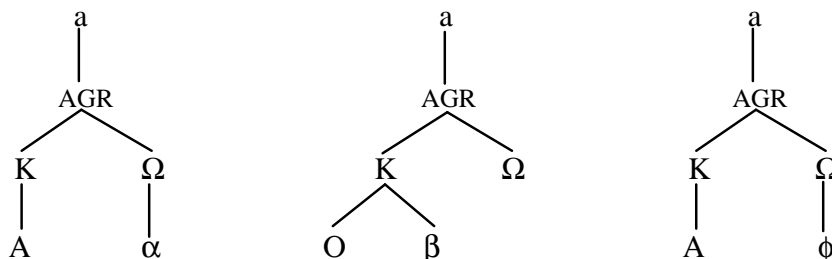
- (i) a. El chico mira a su madre.  
The child looks at his mother.  
b. El chico mira la nao.  
The child looks at the boat.

(11)



As mentioned above, an animate direct object must agree in animacy with the verb which selects it; this is realized via the morpheme *a*. In fact, I will assume that there are many instances of this morpheme; besides animacy, there are other instances of *a* which mark benefactive, inalienable possession, and specificity (at least in certain dialects) respectively.<sup>5</sup> The feature structures of animate *a*, benefactive *a*, and specific *a*, are shown in (12).

(12)



In the light of the above assumptions on feature structure, I turn to consider the phenomenon of spurious *se* in the next section.

---

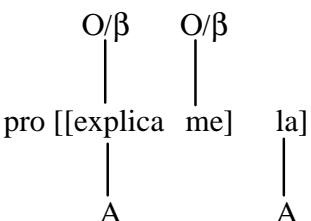
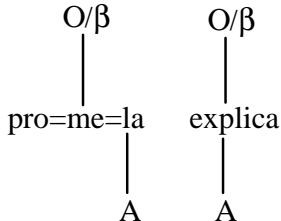
<sup>5</sup> Suñer (1988) discusses the ambivalence of Spanish *a*, which has a tripartite use: DO marker, preposition and Case assigner or manifestator with IOs. Given her framework, she is much concerned with the categorial status of the indirect object in a clitic chain, that is whether it is an NP or a PP. However, it is clear that even for GB researchers, there are many *a* morphemes in Spanish (see further references in her paper).

#### 4.4.2 Spurious *se*

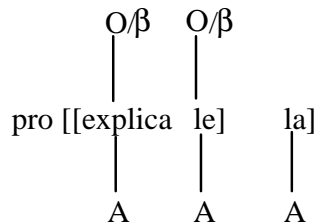
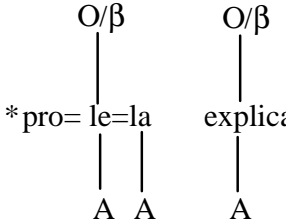
Spurious *se* will be shown to follow from a delinking process, which is triggered by the failure of the second clitic to be licensed. Bonet' s (1991) formulation of the spurious-*se* rule (see below) uses the same tool, namely the delinking process, but her analysis fails to explain why this operation must take place.

Consider a ditransitive verb where both internal arguments are clitics, as illustrated in (13). This sentence may have been derived from the initial structure in (14a); recall that IO clitics (which are animate) are computed first because the verb is specified for animacy, unlike in French (see chapter 2). As can be seen, there are two O/ $\beta$  features in the inner domain of the verb, and two accusative Cases in the larger domain, yielding a violation of the OCP in both domains. Under my assumptions, it is expected that both pronouns move outside the complement domain of the verb; the target of movement is obviously the null subject *pro*. Since the IO pronoun is more embedded (i.e. closer to the target), it moves first and then the accusative pronoun is adjoined to it, yielding the desired linear order, as shown in (14b). The derivation would be the same if the first person clitic were replaced by a second person clitic (cf. *Te la explica* 'h e/she explains it to you') .

- (13) Me la explica.  
(He/she) explains it to me.

- (14) a.  b. 

However, when the IO pronoun is third person, i.e. *le*, which is doubly specified for Case, the derivation yields an ill-formed structure, as exemplified in (15). The output is rather the one given in (16), where the IO is replaced by the pronoun *se*; it is this phenomenon which is referred to as spurious *se*, after Perlmutter (1971).

- (15) a.  b. 
- S/he explains it to him/her.

- (16) **Se** la explica.  
S/he explains it to him/her.

In my view, the ill-formedness of (15b) is expected, since the DO clitic is not licensed. In effect, the IO clitic saturates both Cases of the verb, preventing the DO clitic from checking its Case with the verb, as can be seen in (15a). Notice that there is no reason to assume that the derivation of (16) does not proceed from (15a), since the pronoun *le* is the normal non-reflexive IO. Indeed, if the DO is an NP, an acceptable structure is derived with *le*, because NPs need not be Case-licensed: *Juan \*se/le explica el juego* 'Juan explains him the game'. Thus, given my assumptions, it is natural to assume that there is a further step in the derivation in (15), where the offending feature of the IO clitic is delinked and deleted by stray erasure. Since the clitic *le* is intended to be the indirect object, the offending (i.e. the extra) feature is its accusative Case. The delinking of both the accusative Case of *le* and the relevant  $\Phi$ -features, yields an instance of *se*, (see (9)). This is illustrated by the derivation in (17).<sup>6</sup>

---

<sup>6</sup> Of course Case is not the only feature which is involved in the delinking process. Phi-features of *le*, which are not made explicit in the representation because they are irrelevant to the discussion, must be delinked, since *se* does not have this type of feature. There is indeed evidence that this type of feature matters. In some American dialects of Spanish, the number feature may be relinked to the accusative clitic, as discussed in Bonet (1995). For instance, the sentence in (i) in Iberian Spanish displays a spurious *se* which originates from the plural pronoun (*les*), while the direct object clitic is singular. In some American dialects, the number feature of the IO clitic is associated with the DO clitic, yielding the sentence in (ii), where the accusative clitic agrees with the plural *a*-phrase. Bonet reports that the same alternation is encountered with the gender feature also.

- (i) El libro, **a ellos**, quién **se** lo prestó? (underlyingly: el libro, **a ellos**, quién **les** lo presto.)  
The book to them who se 3.acc. lent  
' Who lent the book to them?  
(ii) El libro, **a ellos**, quién se **los** prestó?



she does not explain why it takes place. Noting this problem, she points out (p.633, fn. 30) that this rule may be related to the OCP, which dissimilates two adjacent morphemes. One may note that the problem remains unchanged; indeed, the analysis should say why the OCP does not apply in French to dissimilate the well-formed sequence *le lui*.

We may note that the explanation by the OCP is further developed by Heap (1996) in the OT framework (see also Grimshaw, 1997). Heap argues that spurious *se* results from the interaction of two constraints: Parse (i.e., spell out all input morphological features), and OCP: \*[-P] [-P], where -P means third person. The ranking of OCP above PARSE yields the spurious *se*. Although appealing, this account presents the same difficulties as the previous ones. In order to account for the absence of a similar process in French grammar, ad hoc ranking of constraints is necessary. Furthermore, one may note that, according to Heap, the relevant feature for the OCP here is third person, and spurious *se* can be seen as a rule of person dissimilation. Now, as discussed above, spurious *se* is third person also (whether underlyingly or via a redundancy rule). Thus, clearly the conversion of *le lo* into *se lo* cannot be a dissimilation of person.

Summing up, our account of *se* is analogous to Bonet's; however, it is different in one important respect: it provides the deep motivation for the spurious *se*, explaining why it should happen, and not only how it happens.

#### 4.4.3 On the I-II/me-lui cluster

While in modern French, the cluster I-II/II-I (*te me, me te*) is not grammatical for perhaps most, if not all, speakers, in Spanish certain speakers accept the cluster II-I (in this order) and others reject it, as noted in Perlmutter (1971) and Bonet (1991). This is illustrated in (20), which is ambiguous. We may note that no Spanish speaker accepts the reverse order, namely I-II, (21), nor the cluster *me le* or *te le*, (as non-ethical datives) as illustrated in (22).

- (20) (\*)Te me recomendaron.  
' they recommended me to you.'

' they recommended you to me.'

(21) \*Me te recomendaron.

(22) a. \*Me le recomendaron.

' They recommended me to him.'

b. \*Te le recomendaron.

' They recommended you to him.'

As regards these facts, the following questions should be answered by any analysis: (a) why the acceptability of (20) is not the same for all Spanish speakers; (b) why all Spanish speakers reject (21) and (22); (c) why French speakers differ from Spanish speakers with respect to II-I cluster such as (20). I will attempt to provide an account of this phenomenon in a highly principled way, using the constraints discussed above, namely the Animacy Hierarchy and the OCP.

It may be the case that the variation among speakers about the acceptability of (20) is due to a variation in feature specification of the pronouns involved. Suppose that for one group of speakers, first and second person pronouns, as well as third person *le*, are specified for animacy ([animate]).<sup>7</sup> Then the ill-formedness of (20) (for those speakers) is accounted for, since the Animacy Hierarchy rejects any structure where the verb has two object arguments which are specified for an identical animacy feature ( $*\alpha > \alpha$ ). Sentence (21) and those in (22) are also ruled out for the same reason, since third person *le* is specified as animate, as we know.

Suppose now that for a second group of speakers first person pronouns are unspecified for animacy. Hence, for such speakers the AH is not violated in (20) ( $\alpha > 0$ ), which is therefore grammatical. Since for those speakers the third person clitic *le* is specified [animate], sentence (22b) is ruled out for the same reason as the other speakers.

---

<sup>7</sup> In Bonet's terms, these speakers have a strong version of the I-II<sub>le-lui</sub> constraint (see chapter 2, footnote 27).



Now the question is why the clusters shown in (21) and (22a) are ungrammatical for the second group of speakers, those who have first person pronouns unspecified for animacy. Apparently, ungrammaticality is not expected, since the AH is not violated. Consider (21). The clitic combination is right, as evidenced by the grammaticality of (20) for those speakers. Thus the problem may be due to the order of the clitics. If so, the relevant question is why a first person clitic may not precede a second person clitic in Spanish (\*1>2).<sup>8</sup>

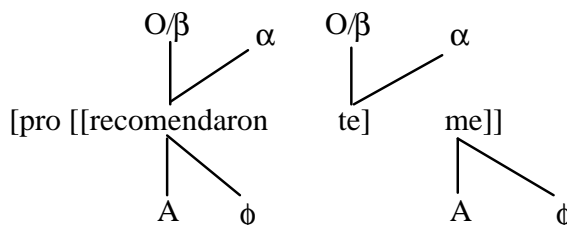
As mentioned above, in Spanish IO clitics are computed before DO clitics. Thus, in the I/II cluster (order irrelevant) if the second person clitic is intended to be the IO, it must be initially generated before the other clitic, as shown in (23). But if *me* is intended to be the IO, it must be computed first, as illustrated in (24). In any case, the clitic *te* must move first, yielding (25). In effect, in (23) *te* is more embedded and has as many features (interacting with the verb) as the first person clitic, namely K (O/β) and Ω (α) vs K (A) and Ω (φ) (recall the discussion of the cluster *le-lui* in French). In (24) it has more features offending the OCP than the first person clitic *me*, namely K (A) and Ω (φ/α) vs. K (O/β) (see the discussion of the cluster *me-le* in French and the WEFC). If this analysis is correct, it is expected that the resulting structure in (25) will be ambiguous, since the second person clitic moves first, irrespective of its grammatical function. This is borne out, as shown by the double glosses of (20) (=25), a fact noted in Bonet (1991) and Perlmutter (1971).

---

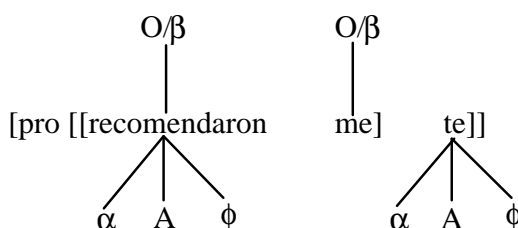
<sup>8</sup> As mentioned in chapter 2, Grimshaw (1997) states a person hierarchy constraint, which she dubs the Universal Markedness Hierarchy (UMH), (i). As can be seen, this constraint refers both to person features and grammatical functions. In the OT framework, which she adopts, it is expected that this constraint be violated, i.e. outranked by other constraints in a given language. In my view, what seems to be a person hierarchy follows from the AH, which in turn depends upon whether the involved elements are specified or not for the features [animate] and [thing]. Thus, under my view, there is no UMH; the order of the clitics must follow from independent principles and the feature inventory of a given language. Indeed, in Italian the order 1>2 is possible, but not the order 2>1 (Di Sciullo, p.c.); cf (ii). The difference between Spanish and Italian could be accounted for once one assumes that in the latter language second person pronouns are unspecified for animacy, all other things being equal. This aspect of Italian clitics will not be further discussed here.

- (i) Universal Markedness Hierarchy  
       \*2 > \*1 > \*3  
       \*DAT > ACC
- (ii) Mi ti recomando.  
       \*T<sub>i</sub> mi recomando.  
       ' he recommended me to you.'

(23)



(24)



(25)

pro=te=me recomendaron.  
 ' they recommended me to you.'  
 ' they recommended you to me.'

As for the ill-formedness of (22a), repeated below as (26a), it must be due to the incompatibility of the clitics *me* and *le*, since the reverse order of the clitics yields ill-formedness as well, (26b).

- (26) a. \*Me le recomendaron.  
 b. \*Le me recomendaron.

Therefore, this cluster is expected to be ruled out by a different constraint. In order to derive (26a) and (26b), either (27a) or (27b) may be posited as the initial structure, depending upon whether the indirect object is intended to be *me* or *le*. Now, when the DO is intended to be third person (27a), an alternative structure is available, which is given in (28). In (28), the clitic *lo* is not specified for  $[\alpha]$ , nor for  $[O, \beta]$ ; indeed, it is specified for  $[A]$  and  $[\phi]$ , and therefore, is a perfect DO. A perfect DO is less costly in the derivation than the double Case clitic *le*, because a smaller number of constraints intervene. Note that in the derivation of (27a), the OCP would force both clitics to move, and the WEFC would force the clitic *le* to move first. In (28) however, only the OCP is relevant, since both clitics are equal with respect to the

number of relevant features. Since the grammar is governed by economy, (28) must be favored over (27a), which is therefore ungrammatical for all speakers.

- (27) a. pro [[recomendaron me] le]  
b. pro [[recomendaron le] me]

- (28) pro [[recomendaron me] lo]  
pro=me=lo recomendaron.

As for (27b), where *le* is intended to be the indirect object, its rejection by all speakers is due to the saturation of the verb by the first clitic (which is specified for all features of the verb), as we know (see (9)). Thus the second clitic cannot be licensed. In my view, this structure must be repaired in order to obtain the meaning of the cluster *le-me*, where *le* is the IO. For both groups of speakers, the repair strategy consists of generating the intended IO as a pronominal *a*-phrase, as shown in (29). If the direct object appears as an *a*-phrase, ill-formedness is obtained, see (30).

- (29) a. pro [[recomendaron me] a él]  
b. pro=me recomendaron a él.  
they recommended me to him.

- (30) a. pro [[recomendaron le] a mi]  
b. \*pro=le recomendaron a mi.  
they recommended me to him.

With regard to this fact, two questions arise: (a) given that for one group of speakers, the AH is not involved in the cluster *le me*, why is it that feature delinking is not used by these speakers? (b) why isn't the DO replaced by a strong pronoun, instead of the IO, i.e. why is (30) ungrammatical? With respect to the second question, we may note that according to Bonet, the third person clitic must be spelled out as a strong pronoun, for in her system, [person] is the only relevant feature in this context. In the present theory, an account of Bonet's descriptive statement is possible. In fact, I will argue that either clitic may in principle be replaced by a strong pronoun, which is the simplest hypothesis. Subsequent ill-formedness, if any, must follow from other independent reasons. For the speakers who have the first person

clitic specified as [animate], two constraints are involved and violated in (27b): the AH and the CLC. If only the AH ( $*\alpha > \alpha$ ) is involved in a clitic cluster, any clitic can be replaced by an *a*-phrase, since ill-formedness comes from their specification for the same animacy feature (see below for a discussion of the effect of  $*\alpha > \alpha$  in I-II contexts). But if a clitic prevents another clitic from being licensed because it is specified for more Cases than needed, its extra Case must be delinked, as we know (cf. the discussion of spurious *se*). That is, more generally, the repair strategy must affect the offending element. Therefore, in the cluster *le-me*, only the offending clitic, namely *le*, must be replaced by an *a*-phrase, because for all speakers it prevents the second clitic from being licensed.

At this point, it would be interesting to compare the strategy which is used in French to circumvent the *me-lui* constraint. As discussed in chapter 2, in French the third person dative is replaced by an *à*-phrase, as in Spanish; compare (31b) and (31a,c). This apparently lends support to Bonet's solution, as this constraint violation must obligatorily be repaired by replacing the third person clitic by a strong pronoun. Indeed, like *le* in Spanish, the first person clitic is specified for all features of the verb and prevents the second clitic from being licensed. Therefore, one would expect under my analysis that the element with the offending features be affected, such that (31c) would be grammatical. In fact, I maintain that the analysis according to which any clitic can be spelled out as a strong pronoun in repairing the *me-lui* constraint is not contradicted by this fact. There is indeed an important difference between French and Spanish, which lies in the feature specification of the clitics involved. In Spanish *le me* the intended IO (i.e., *le*) is the saturator of the verb, since it is doubly-specified for Case. Unlike in Spanish, in French *me lui* the clitic which is doubly specified for Case is the intended DO (i.e., *me*). Thus, if the double-Case clitic is replaced by a pronominal *à*-phrase, the sentence in (31c) is obtained; this sentence is ill-formed, because it contains two indirect objects: the pronominal *à*-phrase (a DO is never introduced by *à* in French) and the clitic *lui*, which is not specified for accusative Case and therefore can never be the direct object. Such an analysis based on constraint interaction is more restrictive and explanatory. To repair the *me-lui* effect, either clitic may be 'spelled out' as a strong pronoun. But a clitic may be spelled out as a strong pronoun only if it satisfies other intervening constraints.

- (31) a. \*Marie me lui présentera.  
       b. Marie me présentera à lui.  
           Marie will present me to him.  
       c. \*Marie présentera lui à moi.  
           Marie will present him to me.

This analysis predicts that if in a clitic cluster only the Animacy Hierarchy is involved, any clitic may be spelled out as a strong pronoun. The repair strategy for the I-II constraint (for the speakers who have the first person clitic specified for animacy) allows any argument to be 'spelled out' as a strong form. This is true regardless of the grammatical function and the person of the clitics, as shown in (32).

- (32) a. **Me** recomendaron **a tí**.  
           they recommended me to you.  
           they recommended you to me.  
       b. **Te** recomendaron **a mi**.  
           they recommended me to you.  
           they recommended you to me.

Notice that this analysis permits us to account for the ambiguity of (32). The ambiguity of these structures is due to the fact that strong pronouns, like NPs, must be introduced by the animacy morpheme, which is homophonous with the benefactive marker. Indeed, the *a*-phrase may be interpreted as either the DO or the IO, as shown in (33), the initial structure for (32b). In (33a), the *a*-phrase is the direct object (A, $\alpha$ ), while in (33b) it is the indirect object (O, $\beta$ ). In any case, the clitic moves outside the complement domain of the verb. Since the grammatical functions of the arguments change while their phonological shapes and their positions remain invariable, the structures are ambiguous.

- (33) a.
- $$\begin{array}{c}
 \text{O}/\beta/\alpha \quad \text{O}/\beta/\alpha \\
 | \quad | \\
 [\text{pro} \text{ [[recomendaron} \quad \text{te]} \quad \text{a} \quad \text{mi}]] \\
 | \quad | \\
 \text{A}/\alpha \quad \text{A}/\alpha
 \end{array}$$
- b.
- $$\begin{array}{c}
 \text{O}/\beta/\alpha \quad \text{O}/\beta \\
 | \quad | \\
 [\text{pro} \text{ [[recomendaron} \quad \text{te]} \quad \text{a} \quad \text{mi}]] \\
 | \quad | \\
 \text{A}/\alpha \quad \text{A}/\alpha
 \end{array}$$
- pro=te recomendaron a mi.

Similarly, the initial structure for (32a) must be as given in (34); the *a*-phrase is the direct object in (34a) and the indirect object in (34b). As in the previous structures, ambiguity results because the morpheme *a*, similarly to the first and second person pronouns (IO and DO), is phonologically ambiguous. In addition, the positions of the pronouns do not vary in the structure.

- (34) a.
- $$\begin{array}{c}
 \text{O}/\beta/\alpha \quad \text{O}/\beta/\alpha \\
 | \quad | \\
 [\text{pro} \text{ [[recomendaron} \quad \text{me]} \quad \text{a} \quad \text{ti}]] \\
 | \quad | \\
 \text{A}/\alpha \quad \text{A}/\alpha
 \end{array}$$
- b.
- $$\begin{array}{c}
 \text{O}/\beta/\alpha \quad \text{O}/\beta \\
 | \quad | \\
 [\text{pro} \text{ [[recomendaron} \quad \text{me]} \quad \text{a} \quad \text{ti}]] \\
 | \quad | \\
 \text{A}/\alpha \quad \text{A}/\alpha
 \end{array}$$
- pro=me recomendaron a ti.

Let us now turn to the question of why feature delinking is not used in (27b) by the speakers who have first person clitic unspecified for [animate]. For these speakers, there are no animacy effects ( $\alpha > 0$ ) and therefore a single constraint is involved, namely the CLC. One would expect this constraint violation to be repaired by a delinking process, parallel to the *le lo* cluster. But this yields an unwanted result, see (35). It is unclear, at first glance, why this structure is ill-formed. However, if one looks at the whole paradigm, one can tentatively put forward an explanation along the following lines. The cluster *se lo*, which is used by all speakers, is ambiguous, as it has a reflexive and a nonreflexive reading. Since the use of a pronominal *a*-phrase is obligatory for another group of speakers, it might be that the speakers who do not have the first person pronoun specified as [animate] also use it instead of the normal delinking process in order to get the ambiguity of the spurious *se* confined to a cluster of two third person clitics.

- (35) a. pro [[recomendaron se] me]  
 b. \*pro=se=me recomendaron.

Summarizing, I have accounted for the *me-lui*/I-II effects, showing that it follows from the Animacy Hierarchy. I have argued that any structure containing two internal arguments dominated by the verb and specified as [animate] is ruled out. I have shown why the repair strategy must be the use of an *a*-phrase instead of one of the two clitics and not feature manipulation by a delinking process. In addition, I have provided the reason why certain speakers accept the I-II cluster and why this cluster is ambiguous.

#### 4.4.4 Spanish enclitics

As is well known, in Spanish weak pronouns are enclitics to imperative, infinitive, and gerundive verbs. In this section I will discuss the case of imperative and infinitive verbs, as well as the phenomenon of clitic climbing, which may occur with infinitive verbs as well as with gerundive verbs, which I put aside.

#### 4.4.4.1 Imperative verbs

The fact that clitics follow imperative verbs in Spanish should not be surprising, since this phenomenon occurs also in French. I have argued that the OCP is violated in French imperative verbs because of the lack of a landing site at the left of the verb. The same reason probably hold for Spanish as well (and most languages, everything being equal), and I do not discuss it further.

However, as we will see shortly, Spanish imperatives are significantly different from French imperatives. They will be discussed because they may support or invalidate the account of clitic alternation in French (*me* → *moi*). Recall that in French imperative verbs, first and second person clitics give rise to a delinking process: the nominative Case of the imperative verb, which is not assigned to an argument, produces a mismatch with the extra Case of the object clitic. The extra Case of the latter is then delinked, yielding a spurious strong form, as shown in (36).

- (36) a.    \*Embrassez-**me**!  
           Embrassez-**moi**!  
           Kiss me!

In Spanish, however, first and second person clitics normally appear as complements of imperative verbs, but not strong pronouns, as can be seen in (37). In my view, this difference is expected since in Spanish, unlike in French, first and second person clitics are specified for a single Case, and thus the nominative Case of the imperative verb, if any, may not clash with an extra Case of the clitic. (In Spanish orthography, enclitics are not autonomous words.)

- (37) a.    \*Besami!  
           b.    Besame!  
           Kiss me!

The Case-mismatch analysis seemingly faces a problem in Spanish. We know that Spanish *le* is doubly specified for Case. Thus one would expect that in the context of imperative verbs, *le* should become a spurious *se* by a delinking process,



parallel to first and second person clitics in French. But this is incorrect: for instance *mirale!* 'I look at him' is a perfect imperative sentence, but not *\*mirase!*.

I maintain that the Case-mismatch analysis is in fact correct, arguing rather that Spanish imperative verbs, unlike their French counterparts, do not have nominative Case. In order to show this, consider the imperative forms of a regular verb such as *cantar* 'to sing'; see (38a). (Spanish imperative verbs have five persons, due to the polite singular and plural.) In contrast to French, where all imperative verb forms are found in the conjugation of another mood, essentially the present indicative, in Spanish there is at least one form which is obviously inherent (i.e. a form which is not shared with another mood) to imperative verbs, the form *cantad*. Since imperative verbs, as a general rule, do not have a subject, it must be the case that a genuine imperative form lacks the subject Case, namely nominative. Let us assume that this is indeed so. Even though the other forms apparently are used in other moods, one may assume that the paradigm is regularized with respect to the genuine form, so that all imperative forms are unspecified for nominative Case.<sup>9</sup>

(38)	a.	Positive imperative	b.	Negative imperative
		Canta		No cantes
		Cante		No cante
		Cantemos		No cantemos
		Cantad		No cantéis
		Canten		No canten

Negated imperatives have no genuine forms, as can be seen in (38b); indeed all the paradigm is shared with the present subjunctive. My claim is that this can be related to the hypothesis that imperative verbs lack nominative Case. Assuming that the negative morpheme is specified for nominative Case (see the analysis of French negation below), it must be licensed by a verb with nominative Case, and not by

---

<sup>9</sup> With respect to imperative morphology, see Silva-Villar (1998) and references therein. This researcher points out that in *voseo* varieties of Spanish (Argentina, Honduras, etc.) imperative verbs have a special morphology which is different from any other verbal forms. Thus the fact that in some Spanish dialects the imperative conjugation is overtly different from other conjugations supports the claim made above.

genuine imperative verbs which are not specified for nominative Case; hence the use of the subjunctive forms. I will return to this issue.

Thus, if Spanish positive imperative verbs are not specified for nominative Case, unlike their French counterparts, there can be no Case-mismatch, and therefore an imperative structure with a double Case clitic should be grammatical. This is illustrated in (39) with a double Case clitic and a single Case clitic.

(39)



Summarizing, Spanish positive imperative verbs are not specified for nominative Case. Therefore they do not trigger any delinking process, contrary to French where a feature clash appears with first and second person clitics.

#### 4.4.4.2 Infinitive verbs

As is well-known, major Romance languages (e.g., Spanish and Italian) differ from French in that they require that clitics follow infinitive verbs; cf. *Pablo quiere verte* 'Pablo wants to-see-you'. In my view, this means that the clitic remains in situ, implying either that the OCP is violated for some reason or that the OCP is irrelevant in that context, also for some reason which must be found out.

The latter alternative is possible only if one assumes that infinitive verbs are unspecified for Case. However, this cannot be correct. If infinitive verbs in Spanish were Caseless, it would be difficult to account for the fact that they undergo fusion with object pronouns, so that both are tightly bound together. Further, infinitive verbs would select an unspecified (strong) pronoun as object, consistent with the Constraints on Merge; for instance one would have a structure like *\*verél* 'to see it' instead of the correct *verlo*. Since on the one hand specified pronouns cliticize normally on infinitive verbs, while on the other hand strong pronouns may not be

objects of infinitive verbs, one may reject the hypothesis that infinitive verbs are unspecified for Case.

Alternatively one may assume that the OCP is violated in this context. Then the question is to explain this state of affairs, providing the deep reasons which block the movement of object clitics in infinitive verbs. I suggest the hypothesis that in Spanish, infinitive verbs and clitics are operator-like elements.<sup>10</sup> Under this hypothesis, an infinitive verb and its clitic complement can be considered to be in a superiority effect relationship. Since traditional accounts of superiority effects rely on the Empty Category Principle and LF (cf. Chomsky, 1973, 1981; Aoun, Hornstein and Sportiche, 1981; Aoun and Li, 1993; Hornstein, 1996; Hornstein and Weinberg, 1995; Huang, 1995), which is not crucial in the present analysis, I am led to present an account of this phenomenon which is more consistent with my assumptions.

I will attempt to account for superiority effects along the following lines. Suppose that operators, including clitics and infinitive verbs (in Spanish), are rigidly ordered within the clause such that their initial order may not change throughout the derivation. Let us refer to this constraint as the Non-Commutativity (of Operators) (40).<sup>11,12</sup>

(40) Non-Commutativity Constraint (NCC)

---

<sup>10</sup> This hypothesis is not new. Di Sciullo (1982) argues that clitics in Romance are operators. In GB an element  $\alpha$  is an operator if it occupies an A-bar position and binds a variable. Since traces are not used here, this hypothesis cannot be tested with the same arguments which are developed in the GB framework. However, it does not rule out the GB analysis, which I assume to hold at the abstract LF level.

<sup>11</sup> Note that the hypothesis that infinitive verbs in Spanish (and presumably in other major Romance languages, except perhaps French) are operator-like elements is not unusual or odd. Recall that Pollock (1989) claims that a verb like *do* in English must move because it is an operator which must bind a variable, which is provided by its own trace.

<sup>12</sup> Guerssel (no date) observes that in Berber, clitics are attached to the verb as enclitics. But when there is a negation morpheme or a *wh*-word in the structure, they attach instead to this element. Since the negation morpheme or the operator attracts the clitic, it must be the case that the clitic shares a feature with those elements.

Similarly in Portuguese, personal pronouns follow the verbs in root clauses and in questions containing no *wh*-words. But in subordinate clauses (with complementizer), negated clauses, and *wh*-questions, the clitic interacts with those elements, and precedes the verb (see Rouveret, 1993, for some data). Again, this suggests that clitics and operators are likely to have some feature in common.

Given a domain *S* and several operators *OP*<sub>1</sub>, *OP*<sub>2</sub>, etc. within *S*, the relative position of the operators must be preserved throughout the derivation.

Given (40), for the time being the simple structures in (41a) can be accounted for in the following way: the operator *what* remains in situ because it may not pass over the higher operator. Indeed, the structure (41b), where movement takes place, is ill-formed. (See a full discussion of superiority effects in English in section 6 below.)

- (41) a. Who bought what?  
b. \*What did who buy?

Under the hypothesis that infinitive verbs and clitics are operators (in Spanish), it is correctly predicted that the movement of the clitic outside the complement domain of the infinitive verb induces a superiority effect, since the clitic will come to be reordered before the infinitive verb in the embedded clause.

Now this yields a constraint conflict: the OCP requires the clitic to move out of the complement domain of the verb, while the NCC prevents the clitic from moving. Since constraints are not ranked, the OCP violation cannot be avoided at the expense of the NCC. The initial structure cannot be safely modified and therefore an OCP violation will appear in the output. Thus in (42a), the initial order of the infinitive verb and the clitic may not change by movement of the clitic to *PRO*, since they are operators (operators are marked with an exclamation mark for the sake of clarity). Indeed, as can be seen structure (42b) in which movement of the clitic occurs is ungrammatical. It is interesting to compare further (42a) and (42b). Both contain a constraint violation: the OCP in the former and the NCC in the latter. (42b) is not equal or better than (42a) because the movement operation increases the complexity of the structure without eliminating or decreasing the number of violated constraints. In contrast, in French, where clitics are not operators, the movement under the OCP must normally take place within VP (=S), as shown in (42c), otherwise an ill-formed structure is derived, (42d). Notice that in Spanish, unlike in French, the clitic is allowed to climb to the matrix clause, a point I will discuss shortly.

- (42) a. [Pablo [quiere [<sub>!VP</sub> PRO [<sub>!ver</sub> **!te**]]]]  
 b. \*[Pablo [quiere [<sub>!VP</sub> PRO=**!te** <sub>!ver</sub>]]]  
 c. [Paul [veut [<sub>VP</sub> PRO=**te** voir]]]  
 d. \*[Paul [veut [<sub>VP</sub> PRO voir **te**]]]

Summing up, basically post-infinitival clitics in Spanish may not move because they induce a superiority effect. I have taken superiority effects to be a scope relationship between two operators, such that their relative positions cannot change within the clause (NCC). In the next subsection, I will provide evidence for this analysis, based on the phenomenon of clitic climbing. To the extent that clitic climbing is consistent with this view of superiority effects and is accounted for straightforwardly, this analysis is strongly supported.

#### 4.4.4.3 Clitic climbing

As is well-known, in Spanish and other standard Romance languages, clitics are allowed to climb over infinitive verbs, a process which is not possible in French; cf. *\*Paul te veut voir* vs. *Pablo te quiere ver* 'P. wants to see you'<sup>13</sup> The generalization is that clitic climbing is likely to occur in languages in which post-infinitival clitics (or enclitics) are allowed. Indeed in French, clitics must move out of the complement domain of infinitive verbs (i.e., they must be proclitic), but may not climb to matrix clauses, while in Spanish, clitics may not move within an embedded infinitival S (i.e. they are always enclitic), but may do so out of it.<sup>14</sup>

<sup>13</sup> On clitic climbing in the (GB) literature, see for instance Bok-Bennema and Kampers-Manhe (1994), Kayne (1989, 1991), Rooryck (1994), Terzi (1996a), etc.

<sup>14</sup> In the literature, it is often pointed out that French allows clitic climbing only in causative structures. In my opinion, the long movement of clitics in causative structures is quite different from the type of clitic climbing discussed here. Under the hypothesis that causatives in *faire* select two internal arguments, one of which is a VP infinitival, as opposed to an S, the clitic complement of the embedded V moves normally under OCP. In most cases, it may not end up to the right of the causative verb, i.e. within the infinitival VP, because the latter being specified for accusative Case, a further OCP violation would result; see (ib). Thus, clitic climbing is obligatory in causatives, because the clitic, being the complement of both the causative verb and the infinitive verb, has to move outside the complement domain of each of these verbs. In Spanish the infinitive verb has a null subject, but the clitic cannot adjoin to it for the reason discussed in this text.

(i) a. Marie=le fait lire a Jean.  
 b. \*Marie fait pro=le lire a Jean.  
 Marie made Jean read it.

This is accounted for straightforwardly under the NCC. The NCC only holds within the infinitival clause which contains the clitic, and therefore the latter may move out of its clause under certain circumstances. I suggest that the clitic may move to the matrix clause only if it can be reanalysed as complement of the matrix verb. That is, when the clitic moves to the null subject PRO, the NCC is violated; but the structure is rescued by a reanalysis, under which the clitic comes to be in the complement domain of the matrix verb. Therefore, the OCP triggers further movement of the clitic outside the complement domain of the matrix verb.

However, if the reanalysis is blocked for some reason, clitic climbing may not take place. This occurs in negated infinitival clauses; compare (43a) and (43b). I claim that the negative operator blocks the reanalysis of the clitic as complement of the matrix verb. The fact is, the clitic first escapes the complement domain of the infinitive verb and adjoins to PRO, as just discussed. But PRO being preceded by the negative morpheme *no*, reanalysis is not possible, since this morpheme can never appear to the right of the verb it modifies. Therefore the clitic must stay in its new position within the embedded clause and the structure is ruled out by the NCC (43c).

---

The object of the embedded verb may be a reflexive clitic. In this case, the reflexive cannot move to the causative clause; compare (iia) and (iib). In my view, the violation of the OCP is avoided in the complement domain of the embedded verb but not in the complement domain of the causative verb. This is presumably due to a conflict between the OCP and the semantic interpretation of the structure. The reflexive *se* must refer to *Paul* and not to *Jean*; if it moves out of the complement domain of the matrix verb, it will be wrongly interpreted as a reflexive of the subject of the causative verb. Hence the violation of the OCP in (iia) is required for the proper interpretation of the structure, consistently with my assumptions.

- (ii) a. Jean a fait **se** raser Paul.  
 b. \*Jean s'est fait raser Paul.  
 Jean made Paul shave himself.

In causatives with *laisser*, which arguably may select an infinitival S as a complement, the clitic may not climb to the matrix clause, since there is a no further OCP violation, the causative verb discharging its Case on the subject of the infinitive verb (iii).

- (iii) a. Marie laisse Paul=le lire.  
 b. \*Marie=le laisse Paul lire.  
 Marie let Paul read it.

Note further that negation of the embedded infinitive in causatives in *faire* yields a deviant structure (ivb). But since this structure would still be deviant without a clitic (iva), there is no evidence that the deviance of (ivb) is related to the negation morpheme. The effect of negation on clitic climbing is discussed below.

- (iv) a. ?Marie fait ne pas lire ce livre à Paul.  
 b. ?Marie=le fait ne pas lire à Paul.

- (43) a. [Pablo [quiere [!S **!no** PRO [!ver **!te**]]]]  
 b. \*[Pablo=**!te** [quiere [!S **!no** PRO **!ver**]]]  
 b. \*[Pablo [quiere [!S **!no** PRO=**!te** **!ver**]]]

Notice that clitic climbing is blocked by the negative morpheme because clitics must adjoin to the first head outside the complement domain of the verb. The movement of a *wh*-operator, in contrast, cannot be prevented by negation because such elements may not be head-adjoined, as discussed below.<sup>15</sup> Notice further that if the negative element is in the matrix clause, it may not interfere with the reanalysis and clitic climbing may take place. The NCC is not violated, since the clitic adjoins to the right edge of the negative element, leaving unchanged their relative positions in the structure, as can be seen in (44).

- (44) a. [Pablo [**!no** [quiere [!S PRO [!ver **!te**]]]]]  
 b. [Pablo [**!no**=**!te** [quiere [!S PRO **!ver**]]]

Let us now turn to further effects of the NCC. In structures where an infinitive verb has two clitic complements, both clitics must climb together, that is the cluster cannot be split so that one clitic is left in situ while the other climbs to a higher clause; compare (45) and (46) (after Roldán, 1974: 134).

- (45) a. Tengo que ir a pedír**selo**.  
 b. Tengo que í**selo** a pedir.  
 c. **Se lo** tengo que ir a pedir.  
 ' I have to go ask him for it.'
- (46) a. \*Tengo que ir**le** a pedir**lo**.  
 b. \***Le** tengo que ir a pedir**lo**.  
 c. \***Lo** tengo que ir**le** a pedir.  
 ' I have to go ask him for it.'

---

<sup>15</sup> For instance, one has a structure such as *¿qué quieres **no** ver?* ' what do you want not to see.' (The inverted interrogation mark is Spanish orthographical convention for all interrogative sentences.)

It is clear that whenever the first clitic moves, it forces the second clitic to move as well. One cannot interpret this fact as indicating that both clitics move together, since they do not form a syntactic constituent. Now, since they do not move together, the question is why the movement of the first clitic induces the movement of the second clitic. Actually this fact follows from the NCC. Indeed, the initial ordering is as follows:  $[\!V > \!CL_1 > \!CL_2]$ ; see (45a). When only the first clitic moves to a higher clause, as in (46a), the structure  $[\!V > \!CL_2]$  is obtained in the lower clause. Since operators are not commutative,  $[\!V > \!CL_1]$  is not equivalent to  $[\!V > \!CL_2]$  and therefore the structure is ill-formed. Notice that the lower clitic cannot move, while the higher clitic remains in situ (cf. *\*lo tengo que ir a pedirle*), because clitic movement from the complement domain of the verb depends upon a locality condition by virtue of which the first clitic moves first, unless the second is the worst evil (cf. WEFC). In Spanish, the worst evil is never the second clitic (which is the direct object). Therefore the first clitic must climb first.

Additional facts of Spanish support this analysis. Consider the paradigm in (47) (from Roldán, 1974: 134). Initially, each infinitive verb hosts its own clitic argument (47a). The lower clitic may move out of its clause, adjoining to the higher clitic, yielding (47b).<sup>16</sup> (47c) is derived when the higher clitic moves to the matrix clause, as illustrated in (47d), and the lower clitic moves there too. Here again the relative positions of the clitics remain unchanged. But, interestingly, when the first clitic moves to the tensed clause, the lower clitic may not move only to the higher infinitive verb, as shown by the ungrammaticality of (47e). The reason is that operators are not commutative; the higher infinitive verb may not have direct scope over the lower clitic. As discussed above, one initially has  $[\!V_1 > \!CL_1 > \!V_2 > \!CL_2]$ , which may not become  $[\!V_1 > \!CL_2 > \!V_2]$ , consistent with the NCC.

- (47) a. Tienes que ver**lo** hacer**lo**.

---

<sup>16</sup> Note that the appearance of spurious *se* in this context is not strange under our proposal. Both *le* and *lo* can be the DO of the verb *ver* in (47a). Thus (47a) may have been spelled out also as *Tienes que ver**le** hacerlo*. One may reasonably assume that only in the latter version the lower clitic is allowed to climb to the higher infinitive verb. Indeed, in this context, the latent oblique Case of the verb may license the higher clitic, which becomes a spurious *se* by losing its accusative Case so that the verb can support the second accusative clitic.



- b. Tienes que vérselo hacer.
  - c. **Se lo** tienes que ver hacer.
  - d. **Lo** tienes que ver hacer**lo**.
  - e. \***Lo** tienes que ver**lo** hacer.
- ' You have to see him make it.'

To summarize, under the hypothesis that in Spanish clitics and infinitive verbs are operators, post-infinitival clitics (encliticization) and clitic climbing phenomena are accounted for straightforwardly under the NCC. The blocking of clitic climbing by negation is no longer a mystery. This analysis is quite simple in that it makes it possible to dispense with the various and complex assumptions frequently invoked in the literature (see the works cited above and references therein). I will return to the NCC, showing that it makes possible an account of superiority effects in English in a highly principled way.<sup>17</sup>

#### 4.4.5 Clitic doubling

##### 4.4.5.1 The problem

In all Spanish dialects, clitic doubling of internal arguments appears in the following contexts: (a) optionally whenever the internal argument is an indirect object expressing a goal (which I refer to as benefactive) (48); (b) obligatorily if it is a strong pronoun (49); and (c) obligatorily in inalienable constructions (50). (51) shows that direct objects are preferably clitic doubled in River Plate Spanish (more data will be given below).<sup>18</sup>

- (48) a. Miguelito (**le**) regaló un caramelo *a Mafalda*.  
' Miguelito gave Mafalda a (piece of) candy.'

---

<sup>17</sup> Although I focus on Spanish facts, this analysis conceivably holds also for other Romance languages, for instance Italian and Portuguese, where clitics are enclitic to the infinitive verb. But note that in Portuguese, pronominal clitics do not move out of the complement domain of tense verbs in simple root clauses, i.e. clauses which do not contain *wh*-words, quantifiers, negative morpheme and subject clitics (see Rouveret, 1993). I leave this problem for future research.

<sup>18</sup> I focus on Spanish facts, but clitic doubling is encountered in various languages, including Hebrew, Rumanian (Borer, 1984), Yagua (Everett, 1996), Berber (Guerssel, 1995).

- b. Miguelito (**les**) regaló caramelos *a unos chicos del barrio*.  
' Miguelito gave some candy to some neighborhood kids.'
  - c. Miguelito **le** regaló un caramelo.
- (49) a. **Lo** vimos *a él*.  
b. \*Vimos *a él*.  
' We saw him.'
- (50) a. **Le** duele la cabeza *a Juan*.  
b. \*Duele la cabeza a Juan.  
' John has a headache.'  
c. **Le** rompieron la pata *a la mesa*.  
d. \*Rompieron la pata *a la mesa*.  
'T ~~he~~y broke the leg of the table.'
- (51) a. **Lo** vimos *a Juan*. (River Plate Spanish)  
b. Vimos *a Juan*.  
' We saw Juan.'

In traditional analyses the problem is reduced to the following: how can the clitic cooccur with the NP it replaces in the same phrase marker? As observed in chapter one, this problem leads many researchers to abandon the movement analysis of clitics. Under one type of analysis, clitics are base-generated under the V node, and the corresponding NP appears in the normal position of arguments (Jaeggli, 1982; Borer, 1984; Suñer, 1988, 1991). However, the real question as to why a given language uses clitic doubling is not answered in those analyses, nor in Everett's (1996) phi-feature analysis.

Consider for instance Suñer's analysis of clitic doubling. It is based on her Matching Principle, which requires that a clitic and the NP it doubles agree in features. Assuming with Borer (1984) that Spanish clitics are agreement morphemes generated as part of the verb, she goes on to argue that as such they must agree with the NP by virtue of the Matching Principle. Thus, a clitic doubled construction will be well-formed if and only if the features of the clitic and those of the verb do not clash. Furthermore, clitic doubling is shown to be possible in River Plate (including

Porteño) Spanish whenever the direct object is [+specific,+animate] (see also Jaeggli, 1982). While this claim is descriptively correct, it does not explain why clitic doubling must take place. Rather the Matching Principle appears to be a well-formedness condition on clitic doubled structures. Moreover, as Everett (1996: 61) points out, ' she does not make clear why other Spanish dialects do not allow this type of agreement' , i.e. clitic doubling of direct object.

We may note that Suñer' s analysis is radically incompatible with my assumptions, and presents no independent benefits. By claiming that clitics are agreement morphemes, she is led to conclude, given the GB framework, that they cannot be Case absorbers nor thematic role absorbers, since they are not in A-positions. Furthermore, the problems of clitic order, clitic movement, etc., cannot be accounted for under such assumptions, as discussed in the preceding chapters.

Another approach to the problem of clitic doubling is proposed by Everett (1996). In this work, clitics are taken to be bundles of phi-features in AGR, which are adjoined in the syntax to a maximal projection. Everett assumes a phrase-structure tree for Spanish which contains three AGRPS above the VP, besides CP and TP. AGRP<sub>1</sub> is the subject agreement marker, which is assumed to be incompatible with clitics. AGRP<sub>2</sub> and AGRP<sub>3</sub>, on the other hand, can only appear as clitics.

This analysis relies on the central assumption that ' referents necessarily affected by an action are identified within the extended projection of the verb' (p.63), affected referents being: agent, theme, speaker or hearer, possessor, and goal. In the case of a goal indirect object verb, it is assumed that the verb obligatorily assigns a goal theta-role. A clitic thus will appear in AGRP, since for each theta-role assigned by the verb, there must be a corresponding AGRP (under this view, clitics are bundles of phi-features in AGRP). In order to account for the optionality of clitic doubling of the goal complement, Everett claims that ' the need for AGR to be overt will increase in direct proportion to the configurational distance of the grammatical function from the phrasal head' (p.64). Since the goal clitic, which is in AGRP<sub>3</sub>, is very close to the verb, it may then be optionally overt.

Besides the assumptions on phrase markers and the nature of clitics which are radically different from mine, this analysis seems to have been designed to account for one fact about clitic, namely clitic doubling. Indeed, when one assumes that clitics are bundles of phi-features which are base-generated in AGR, the problem of clitic order cannot be solved. With respect to this problem, Everett follows Bonet (1991), who in turn relies on Perlmutter's (1971) claim that clitic clusters in Romance are framed in a template in which their relative positions as well as constraints on cooccurrence are specified, both on a language-particular basis. As discussed above, the template is at best a descriptive statement of the problem. It may not be taken as an explanation. Moreover, this analysis relies on the claim that each AGR corresponds to a type of clitic: AGRP<sub>2</sub> may only be an ethical dative, and AGRP<sub>3</sub> may be available for ethical, possessor, or indirect object clitics, a proposal which is likely to be incorrect, as he himself acknowledges (p. 62). We may note that under such a view there are no independent means to check the configurational distance of the grammatical function to the phrasal head, since the relative position of each AGR is stipulated on a case-by-case basis.

The same type of reasoning is found in Everett's account of inalienable constructions, pronoun doubling and direct object doubling. For instance, to account for the fact that inalienable constructions obligatorily require clitic doubling, he claims that 'any referent necessarily affected by the action of the verb be theta-marked within the extended projection of the verb.' (p. 65). Thus, since the possessor in inalienable constructions is affected by any action which affects the possessee, it must be theta-marked. And since it is theta-marked, an AGR, i.e., a clitic, must be generated in the extended projection of the verb.

In the next section I will attempt to account for the problem of clitic doubling: in section 4.5.2, I consider the problem of clitic doubling which is common to all Spanish dialects, referred to as Common Spanish for expository purpose; in section 4.5.3 certain peculiarities observed in River Plate Spanish dialects, including Porteño Spanish (the dialect of Buenos Aires, as reported in Suñer, 1988, 1991), will be accounted for.

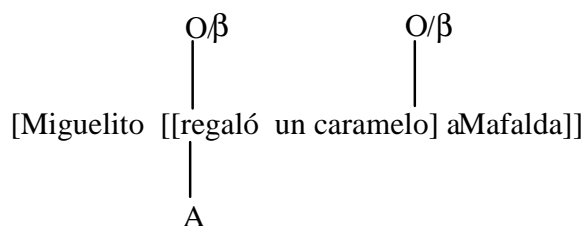
#### 4.4.5.2 Common Spanish

##### 4.4.5.2.1 Indirect object doubling

Under our proposal, it is not surprising that clitics may cooccur with NPs. Recall that pronominal clitics are licensed obligatorily by Case, i.e. they may appear in a structure only if there is a verb with the relevant Case, and either by the lexical structure of the verb or by referring to a relevant argument of the verb. NPs on the other hand, are primarily licensed by the lexical structure of the verb, though in most cases they are assigned a Case as well. Thus a clitic and an NP may cooccur in a structure only if the former is licensed by the latter (i.e. they are coindexed). Now, if a clitic refers to an NP, it must agree in features with that NP, an obvious descriptive fact. Thus, descriptively, the problem is clear. What must be explained is why it happens. I will show that clitic doubling is a repair strategy used to avoid a feature clash. Therefore, if there is no feature mismatch during the derivation, it may not take place.

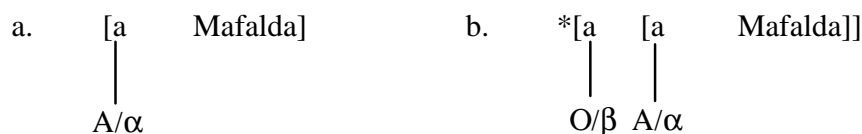
With this in mind, consider (48a). Since in Spanish, internal arguments must agree in animacy with the verb (both AGR nodes in the tree structure of a double-object verb in Spanish contains the feature [animate]), whenever the argument is animate, it is expected to merge first with the animacy marker *a*, consistent with the Constraints On Merge. In the case of an IO verb, the animacy feature may be redundant since a benefactive argument is necessarily animate. Thus with such a verb, it is expected that the animacy marker may not be absolutely required as in the case of accusative verbs. Thus, when the animacy feature is ignored (because of its redundancy), the structure in (52) is derived (spreading omitted), which is quite parallel to the double-object construction in French, as discussed in chapter 2.

(52)



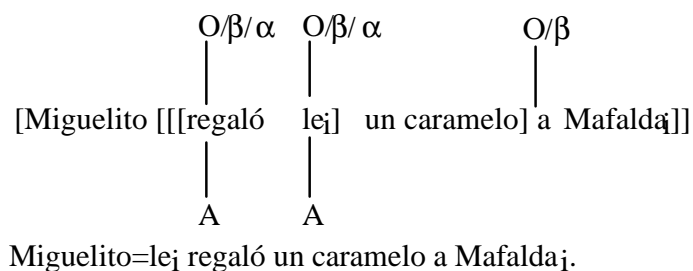
When the animacy feature of the verb is (redundantly) used, the argument must agree in benefactiveness and in animacy with the verb. This means that the argument must be first merged with the animacy marker (53a) and then the *a*-phrase thus formed must be merged with the benefactive marker (53b), or vice versa. But (53b) is not a well-formed structure, since all the features come to clash with each other, in violation of the COM.

(53)



I suggest that a clitic-doubled construction is precisely intended to circumvent the problem in (53b). Since there is a clitic which is specified for [O, $\beta$ ] and [ $\alpha$ ], one of the *a* morphemes may be replaced by the relevant clitic. Under this condition, the verb is merged with the clitic, which absorbs all the relevant features of the latter, and then with the direct object and the *a*-phrase, as illustrated in (54). Both the NP (DO) and the *a*-phrase satisfy the lexical structure of the verb, and therefore are fully licensed. The clitic also is licensed, since it saturates the Case of the verb (Case-licensing) and is coindexed with the *a*-phrase.

(54)

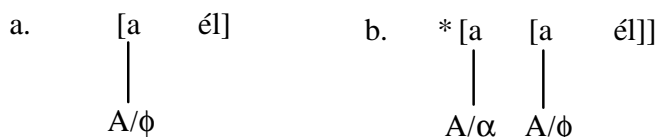


To conclude, benefactive complement doubling is intended to avoid a feature mismatch between the benefactive marker and the animacy marker. In the next section, strong pronoun doubling will be shown to follow from a mismatch between specificity and animacy.

#### 4.4.5.2.2 Strong pronoun doubling

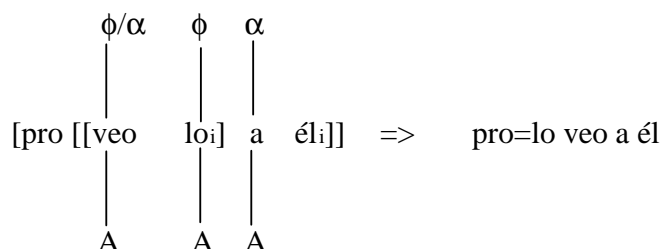
Consider now the case of strong pronouns, which must obligatorily be clitic doubled in all Spanish dialects. In Spanish the feature structure of a verb having a strong pronoun argument differs from that of a verb having an NP argument in that the former contains the specificity feature in addition to the animacy feature. Thus, a strong pronoun complement must be introduced by the animacy marker and a specificity marker, if there is any. Suppose indeed that there is an instance of the morpheme *a* which is specified for specificity (as given in (12) above). Therefore, under the operation Merge, the features [animate] and [specific] will clash with each other, although both morphemes share the feature accusative, as shown in (55).

(55)

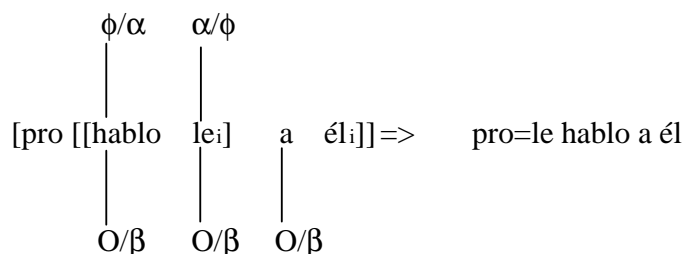


Here again, in my view a clitic doubled construction must be used to avoid the feature mismatch (55b). Since a clitic need not be licensed by the lexical structure of the verb, which is already filled by the strong pronoun, it can be used as an agreement marker, providing the feature specificity. This is shown in (56); the clitic is used as the specificity marker, while the strong pronoun is introduced by the animacy marker. Therefore the feature mismatch no longer exists in the structure. With a benefactive verb, more features are involved: [animate], [benefactive] and [specific], which the clitic *le* is specified for, as illustrated in (57) (the irrelevant accusative Case of the clitic is omitted). Since the clitic is specified for all relevant features, the strong pronoun must be introduced by any instance of *a*, so as to agree with the clitic.

(56)



(57)



Before concluding this section, I would like to motivate the feature [specific]. In first and second person singular, the strong forms are *yo* and *mí*, and *tú* and *ti* respectively. I argue that *yo* and *tú* are the only strong pronouns which are specified for specificity. These forms are obligatory when an overt pronominal subject is required, and after a restricted set of prepositions, including *menos*, *salvo*, *excepto* 'except' *incluso* 'enclosed' *según* 'according to' *entre* 'between'. All these prepositions share a common feature to the extent that they single out an element (their argument) from an indefinite set. Indeed their arguments are necessarily specific. Since the strong forms *yo* and *tú* may be the complement of these prepositions, but not *mí* and *ti*, it must be the case that they are inherently specific. The exclusion of the other forms is ensured by the Blocking Principle. The strong pronouns of all other persons do not display such an opposition, and one may assume that they are unspecified for specificity.

Notice that the feature [specific] appears on each AGR node of the verb, and therefore each strong pronoun argument must be marked for specificity via the specificity marker *a*. But, in fact, only strong pronouns which are internal arguments are introduced by the specificity marker. This is due to the fact that the specificity



marker being also specified for accusative Case ([A,φ]), a feature mismatch (nominative vs. accusative) will occur if that morpheme (the specificity marker) introduces a strong pronoun which is used as subject.<sup>19</sup>

#### 4.4.5.2.3 Possessor doubling

This type of analysis extends naturally to 'inalienable possession' constructions, where the possessor is obligatorily clitic doubled; see (50a,b) and (50c,d) repeated below as (58) and (59) respectively. According to Jaeggli (1982: 34) "the presence of the clitic is required to fix appropriately the thematic relation of the *a*-phrase. Presence of the clitic implies one particular thematic role, the role that is found in 'inalienable possession' constructions— we can call  $\theta_p$ ." I agree with Jaeggli that the particular thematic role which is involved here is the inalienable possession. However, I assume that such a thematic feature is conveyed by a special morpheme, homophonous with the animacy, specificity, and benefactive morphemes. Given this assumption, clitic doubling will be shown to arise from a feature mismatch, namely inalienable vs. benefactive.

- (58) a. Le duele la cabeza a Juan.  
' John has a headache.'  
b. Le rompieron la pata a la mesa.  
' They broke the leg of ~~the~~ table.'
- (59) a. \*Duele la cabeza a Juan.  
b. \*Rompieron la pata a la mesa.

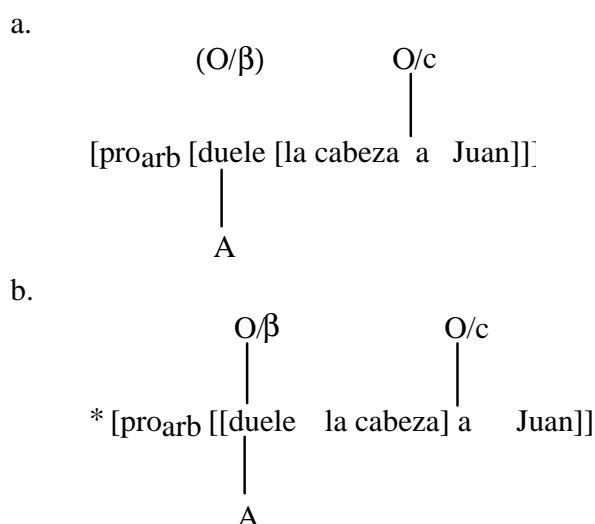
Let us take *la cabeza a Juan* ' Juan' s head' to be a complex possessive NP. In this NP, the possessee is 'included' in the possessor, that is to say, it is a subpart of

---

<sup>19</sup> It should be noted that the strong pronouns *yo* and *tú* may never appear as objects of the verb, since they cannot be merged with the animacy morpheme, given their feature specificity (cf. \**busco a tú* ' I am looking for you' ). Given the analysis presented above, one would expect a clitic doubling to take place, yielding \**te busco tú* . In fact this structure is ungrammatical. The ungrammaticality comes from the fact that the feature [specific] of the pronoun *tú* is not checked by the verb. The clitic *te* absorbs all features of the verb, including [specific], leaving no room for the strong pronoun.

him/her. Suppose that Spanish grammar marks this relation of inclusion with a special thematic feature, call it tentatively *c*, for inclusion, and assume that a further instance of the morpheme *a* is specified for [O,*c*]. The structure (59a), for instance, may have been derived as follows. The simple transitive verb *duele* is merged with the complex NP *la cabeza a Juan*; yielding the structure in (60a) (assume a null arbitrary subject for this construction). Since the *a*-phrase is specified as [O], it triggers the activation of the latent Case of the verb, namely [O, $\beta$ ] (see chapter 3). The simple transitive verb is then reanalysed as a double object verb. Under this condition, a feature mismatch appears in the structure, namely  $\beta$  vs. *c*, hence the ungrammaticality of the resulting structure (60b).

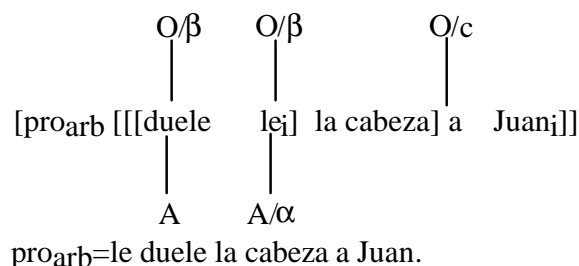
(60)



Parallel to the cases discussed above, I argue that use of a clitic allows this mismatch to be avoided ( $\beta$  vs. *c*). The clitic must agree in  $\Phi$ -features with the NP in order to be licensed by coindexation. Obviously the only clitic which may fill this function is *le*, since it is specified for [O, $\beta$ ], and for accusative Case as well. The clitic *le*, indeed, will saturate both the latent Case of the verb and its accusative Case. (The A Case of the verb must be saturated, otherwise it will clash with the O Case of the *a*-phrase.) Therefore, the features of the *a*-phrase no longer clash with the features of the verb, as shown in (61). Notice that the *a*-phrase, which induces the feature mismatch, cannot appear in the structure without the clitic buffer. However,

the clitic can stand without the *a*-phrase, since it is licensed by Case and, in the absence of the *a*-phrase, it occupies a slot in the argument structure of the verb.

(61)



The next subsection is devoted to clitic doubling in River Plate Spanish. River Plate Spanish is different from standard Spanish in that the specificity feature, instead of being confined to the personal pronouns, plagues the paradigm of NPs as well.

#### 4.4.5.3 River Plate Spanish

The basic facts of River Plate Spanish (RPS) are given below (the data are from Jaeggli, 1982; Borer, 1984; Suñer, 1988, 1991). Indirect and direct objects are preferably clitic doubled; cf. (62a) and (63a). Notice that the animacy marker is obligatory with all animate direct objects, clitic doubled or not; compare (63a,b) and (63c,d).

- (62) a. Miguelito **le** regaló un caramelo **a Mafalda**. (preferred)  
 b. Miguelito regaló un caramelo **a Mafalda**.  
 M. gave a candy to M.

- (63) a. **Lo** vimos **a Juan**. (preferred)  
 b. Vimos **a Juan**.  
 c. \*Vimos **Juan**.  
 d. \***Lo** vimos **Juan**.  
 We saw Juan.

Furthermore, in the general case, an inanimate specific direct object may be introduced by the morpheme *a* but may not be clitic doubled, as shown in (65). In

contrast, an inanimate nonspecific cannot be introduced by the morpheme *a*, nor can it be clitic doubled, as illustrated in (64).

- (64) a. Vimos una camisa.  
 b. \*Vimos a una camisa.  
 c. \***La** vimos una camisa.  
 d. \***La** vimos a una camisa.  
 We saw a shirt.

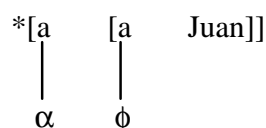
- (65) a. Vimos a la camisa.  
 b. Vimos la camisa.  
 c. \***La** vimos a la camisa.  
 d. \***La** vimos la camisa.  
 We saw the shirt.

If the analysis according to which clitic doubling is used to avoid a feature mismatch is correct, it must be the case that some feature that is not found in other dialects is involved in the structures above. The examples in (64) and (65) make it possible to track down this feature. Unlike common Spanish, an inanimate DO may be preceded by the morpheme *a*, as can be seen in (65a). Obviously, this instance of the morpheme *a* has nothing to do with the animacy marker *a*; since it only introduces specific direct objects (cf. the ill-formedness of (64b)), it must be the case that it is the specificity marker, which we discussed in the context of pronominal *a*-phrases. If so, it appears clearly that this Spanish dialect makes no distinction between pronominal and non-pronominal arguments with respect to the specificity feature. Therefore, parallel to the case of strong pronouns, it is expected that a feature conflict should appear between the features [animate] and [specific], suggesting that clitic doubling is obligatory. Indeed, it has been reported that a direct object may be clitic doubled if and only if it is animate and specific (see Suñer, 1988, and references therein).

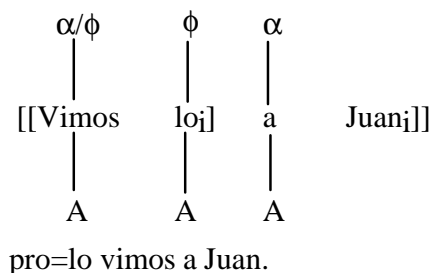
Consider the derivation of the clitic doubled construction in (63a). Since specific NPs in RPS must be marked for specificity, the specific object *Juan* must be merged with the specificity marker, namely *a*. Furthermore, since this NP is animate,

it must be merged with another morpheme *a*, namely the animacy marker. But this yields a feature clash,  $\alpha$  vs.  $\phi$ , as shown in (66). Therefore, as in the case of strong pronoun doubling, a clitic must be used in place of one of the morphemes *a*; the clitic must replace the specificity morpheme, since all clitics are specific but not necessarily animate; see (67).

(66)



(67)



It has been argued that clitic doubling is possible only if the NP is assigned Case by a preposition, given that the clitic absorbs the Case of the verb. This is usually referred to as Kayne's generalization (Jaeggli, 1982). However, Suñer (1988) points out that this generalization is a spurious one, since one can find bare NPs (not introduced by the morpheme *a*) which are clitic doubled, as illustrated in (68) (her example (14a)).

(68) Yo la tenía prevista esta muerte.

I had foreseen (it) this death.

The present account of clitic doubling does not depend upon such a generalization. Moreover, it must not do so, since in my view an NP need not be licensed by Case. Nevertheless, I must say something about this structure, for the NP is doubled in a context where apparently there is no clash between animacy and specificity. This may be accounted for straightforwardly. Since the clitic is specified for specificity, nothing prevents it from being used instead of the specificity marker,

even in the absence of the animacy marker. In other words, the clitic in this case is not intended to avoid a feature mismatch but to supply a missing feature, that is, it is used as a simple agreement marker.

Let us further consider implications that support the whole analysis of clitic doubling. One may predict that if an element is animate, but non specific, it may not be clitic doubled, since a feature conflict does not arise under this condition. This is borne out, as shown in (69): in (69a) it is the animacy morpheme which is involved, since *alguien* cannot be specific; absence of the animacy morpheme yields ill-formedness (69b). Since there is no conflict between specificity and animacy, clitic doubling is not necessary, and indeed is ungrammatical (69c,d).

- (69)     a.     Yo vi a alguien.  
              ' I saw someone.'  
              b.     \*Yo vi alguien.  
              c.     \*Yo lo vi a alguien.  
              d.     \*Yo lo vi alguien.

A further piece of evidence is provided by wh-extraction. The prediction is that a wh-word being necessarily non-specific, if it is animate there will be no clash in features, and therefore no clitic doubling is expected. All the data discussed in the literature confirm this prediction. As an exemple, consider the extraction of the *a*-phrase in (63a) repeated below as (70) (see Suñer, 1988, 1991, for more data). Recall that the clitic in (70) is used in order to avoid the feature clash which arises from the conflict between animacy and specificity. Thus, if the argument is replaced by a relevant wh-operator (i.e. an animate one), it loses its specificity, and only the animacy morpheme is required (71a). In (71a), wh-movement may further take place yielding (71b).

- (70)             Lo vimos a Juan.

- (71)     a.     ¿[Vimos [a quien]]?  
              b.     ¿A quien vimos?

Extraction of an indirect object which is clitic doubled is always possible, since in all Spanish dialects there is an overlap between [animate] and [benefactive], hence the optionality of clitic doubling. In RPS, the opposition involves three features, namely [benefactive], [specific] and [animate]. Thus, if the argument is *wh*-extracted, the feature [specific] is irrelevant, but there still remains the features [animate] and [benefactive], parallel to the standard dialect. The following example ((1) in Suñer, 1988) illustrates *wh*-extraction of a clitic-doubled indirect object.

- (72) ¿A quién **le** regalaron un auto?  
       to whom him/her-dat gave a car.  
       Whom did they give a car?

Before summarizing, as mentioned above, clitic doubling exists in many other languages, including Rumanian and Hebrew, as reported in Borer (1984). These languages present a striking similarity with Spanish in that they possess an animacy marker. Analysing subject clitic doubling in Berber, Guerssel (1995) shows that in this language the subject is necessarily definite, which suggests that my view of the data is not unusual.

#### 4.4.6 Summary

In clitic-doubled constructions, the clitic functions either as a buffer between the verb and the argument, the features of which would mismatch otherwise, or as an agreement morpheme, in order to avoid a feature mismatch between two different instances of the marker *a*. Clitics may be used in this context because they do not need to be licensed by the lexical structure of the verb. Conversely, NPs need not be licensed by Case, the lexical structure of the verb being sufficient. This analysis of clitic doubling is quite simple and highly plausible: the deep reason for clitic doubling in Spanish is provided, such that the reason for the lack of a similar phenomenon in French can be deduced straightforwardly.<sup>20</sup> In the next section, I

---

<sup>20</sup> Clitic doubling in Spanish is required with psych-verb of the third class of Belletti and Rizzi (1988), as illustrated in (i) (Montrul, 1995, ex.1). Psych-verbs have a special syntax that I cannot address in this thesis.

(i) A Juan **le** gusta la musica.

discuss the main properties of Italian clitics, which provide additional evidence for the type of analysis I have developed so far.

#### 4.5 Overview of Italian Clitics

Italian clitics provide further support for the analysis developed in this thesis. Recall that the prediction is that if a pronoun is specified for Case, it must show clitic-like behaviour. In particular, it is likely to be a special clitic, i.e., a Case-specified pronoun which moves under the OCP out of the complement domain of the verb. Since, as is well-known, Italian Clitics precede the verb (in most cases), like in French and Spanish, the prediction is borne out.

While in Spanish there is a single (and celebrated) case of opaque clitics, namely spurious *se*, in Italian the first clitic in a clitic cluster is usually opaque, that is, its morphological shape is not the same when it is used in isolation. I will show that opaque clitics in Italian follow from the same cause as Spanish spurious *se*, namely the failure of the second clitic to be Case-licensed.

##### 4.5.1 Feature specification

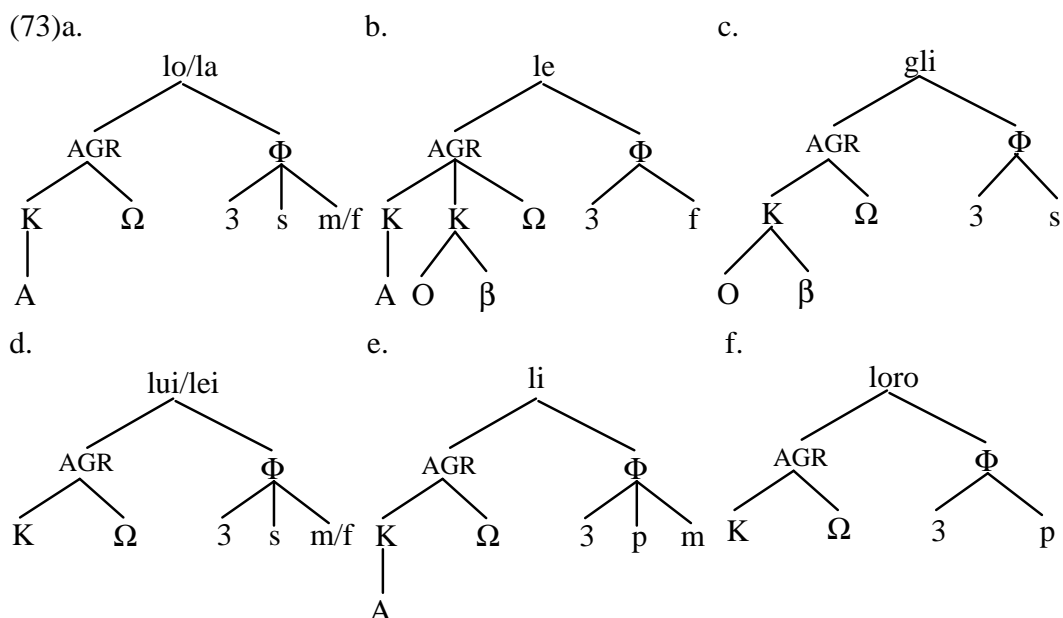
Below I present the essential of the paradigm of Italian personal pronouns (subject pronouns are set aside), as can be found in normative grammars. Only standard Italian is considered.

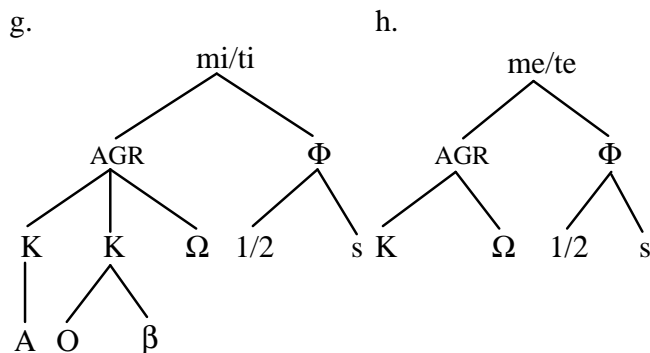


**Table 4.4** Italian personal pronouns.

	Direct object	Indirect object	Strong forms
1	mi	mi	me
2	ti	ti	te
3-masc.	lo	gli	lui
3-fem.	la	le	lei
1	ci	ci	noi
2	vi	vi	voi
3-masc.pl	li	loro	loro
3-fem.pl	le	loro	

The descriptive facts in Table 4.4 can be best captured by assuming the feature structures given in (73). First and second person plural pronouns are not shown below because their feature structures are identical to their singular counterparts. I take strong pronouns to be unspecified for Case, consistently with standard assumptions.





If the feature structures given in (73) are correct, one may predict that all third person pronouns in Italian, except *loro*, are special clitics, that is, given the OCP, they must appear in preverbal position. This prediction is borne out, as illustrated in (74).

- (74)
- a. Maria **lo** / **la** vede.  
Maria sees him/her.
  - b. Maria **li** / **le** vede.  
Maria sees them.
  - c. Maria **gli** / **le** parla.  
Maria speaks to him/her.
  - d. Maria parla **loro**.  
Marie speaks to them.
  - e. \*Maria **loro** parla.

Notice that first and second person pronouns are doubly specified for Case, parallel to their French counterparts. Therefore these pronouns are expected to be special clitics, parallel with the others previously discussed. Some examples are given in (75).

- (75) a. Maria **ti** guardava.  
Maria looked at you.  
b. Maria **mi** presta quel libro.  
Maria lends me this book.

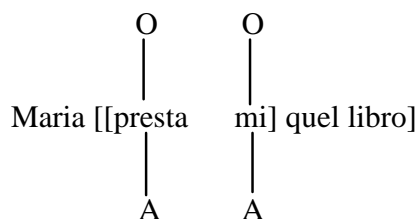
Below I will show that the double Case-specification of these pronouns is precisely what prevents them from appearing in a clitic cluster without feature manipulation.

#### 4.5.2 Opaque clitics

As mentioned above, the clitics *mi*, *ti*, *ci*, and *vi* become respectively *me*, *te*, *ce*, and *ve*, i.e. they are replaced by the corresponding (same person) strong pronoun, whenever they are followed by another clitic. A correct theory of clitics must explain how this transformation is possible and why it must take place. From the discussion of Spanish spurious *se*, we know that opaque forms result from a delinking process up to structure preservation, an operation which is required because clitics must be Case-licensed. I show that opaque clitics in Italian arise from the same reason, namely the Case-licensing requirement.

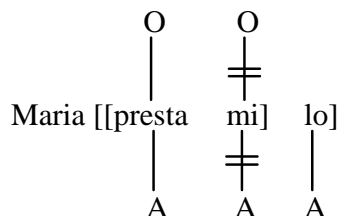
Consider (75b). It may have been derived from the structure in (76) by the movement of the clitic under the OCP. Since the clitic is doubly specified for Case, it saturates both Cases of the verb, and therefore the DO *quel libro* is not assigned Case; it is licensed, however, by the lexical structure of the verb, as we know.

(76)

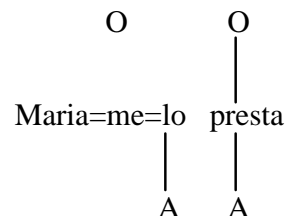


Suppose now that the DO is also a clitic, as illustrated in (77). Since the first clitic saturates both Cases of the verb, the second clitic cannot be licensed by Case. Therefore, the delinking processes automatically take place, eliminating the offending accusative Case of the first clitic, and its oblique Case as well, given the operation must yield an existing clitic (structure preservation). A spurious strong form is thus derived and the floating accusative Case is then deleted by Stray Erasure, allowing the second clitic to be licensed and then to move, also under the OCP (77b).

(77) a.

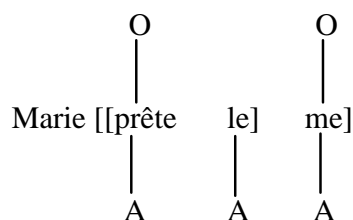


b.

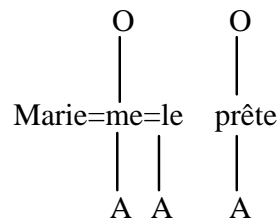


We may recall that the same feature configuration appears in French, except that the DO clitic initially precedes the IO clitic, as shown in (78). The DO is licensed before the IO clitic (double Case) saturates the accusative Case of the verb; thus the need for feature manipulation is obviated.

(78) a.



b.



Similarly, when the third person IOs *gli* and *le* are followed by *lo*, *la*, *li*, *le*, etc., instead of having a sequence of two different clitics, a single opaque form is obtained, namely *glielo*, *gliela*, *glieli*, *glielle* respectively. However, these opaque forms may not come from a delinking process. In effect, *gli* being specified for a single Case may not prevent the second clitic from being licensed by the accusative Case of the verb. Furthermore, although the cluster looks like a single morpheme, it saturates the argument structure of the verb. Indeed, a verb having such an opaque form as argument may not license an NP, cf. *Gino glielo regala* (\**un libro*/\**a Maria*) 'G. gives it to her (a book/to Maria)'. Therefore, I take the single opaque form to be merely an orthographic convention, which seems to be motivated by the appearance of the phonological alternation, namely *glielo* vs. \**glilo*. In previous work, I assumed incorrectly that this phonological alternation is simply due to the insertion of a vowel, perhaps a rule which is limited to the paradigm of personal pronouns. But Anna Cardinaletti (p.c.) points out to me that in the (orthographical) cluster *ie*, *i* is not spelled out, only *e* is. In light of this fact, I am led to accept the suggestion by Angela

Ralli (p.c.), according to which the forms *glielo*, *gliela*, etc. result from paradigmatic levelling, by analogy with the other forms *melo* and *telo*.<sup>21</sup>

Observe that when the first clitic is *le* (descriptively used as feminine, plural or singular, direct object or indirect object) one obtains the same forms just discussed, and not *le lo*, *le la*, etc. This result is expected since *le* is doubly specified for Case, and there exists another clitic in the paradigm which bears a single oblique-benefactive Case, namely *gli*. Thus the target of the delinking process is *gli* and not a Caseless strong form. It is likely that *le* is specified for [feminine] and not for number, while *gli* is specified neither for gender nor for number. Given the Blocking Principle *le* is confined to feminine use, plural or singular, while *gli* acquires the masculine and the singular features by default.<sup>22</sup> Since *gli* is unspecified for number, it can be used as a plural clitic as well when a delinking process has taken place. Perhaps the use of the strong form *loro*, which is specified as [plural], reflects an earlier stage of the paradigm, and indeed, is not used in informal speech.

#### 4.5.3 Summary

The behaviour of personal pronouns in Italian confirms the correctness of the OCP and more generally the type of analysis argued for in this thesis. French and Italian are alike in that in both languages first and second person pronouns are doubly specified for Case. However, in Italian, indirect objects are computed before direct objects (recall the discussion of the Animacy Hierarchy), yielding a delinking process whenever there is a second clitic. The delinking process converts the indirect object clitic into a spurious strong form. Opaque forms involving the third person clitic *gli* are due to paradigmatic levelling.

This section brings to an end the analysis of personal pronouns. I believe that I have attained my primary objectives, which were the following: to account in a principled way for (a) the clitic-hood of personal pronouns, and (b) morphological

---

<sup>21</sup> Thanks to Anna Cardinaletti and Angela Ralli for their suggestions.

<sup>22</sup> That is, in a paradigm where there is an opposition between masculine and feminine, it is unlikely that both features exist. Rather, given privative underspecification theory, one morpheme must be specified for either of these features, while the other is unspecified.

variations within languages, without parameterizing any constraints. In the Romance languages discussed so far, namely French, Spanish and Italian, we have a clear view of the phenomena of cliticization, including procliticization, encliticization, clitic climbing, opaque forms, etc. In the next section, this type of analysis is extended to *wh*-pronouns.

#### 4.6 Wh-operators as special clitics

I have shown that the term 'clitic' may not be taken to be a type of lexical entry, i.e., a genuine grammatical category. All the properties covered under this term are shown to follow from OCP effects, namely fusion and movement. If this view is correct, other types of categories should behave like pronominal clitics, provided that they are specified for Case features. In this section, I will consider certain *wh*-operators in French, showing that their syntax can naturally be explained in a highly principled way with the same general principles which I use so far.

Considering *wh*-words to be clitics is not a new idea. Bouchard and Hirschbuhler (1987) attempt to show that in French the *wh*-operator *que* is a clitic, parallel with personal pronouns. They show that this element behaves in all respect like other clitics. One of their central arguments is the fact that this *wh*-pronoun alternates with another element, a strong form, *quoi*. As their analysis is embedded in the GB framework, they take a clitic to be a genuine category, making therefore no attempt to explain why the *wh*-word is a clitic. I will begin by analysing *que*, showing why it is a clitic and how it differs from personal pronouns with respect to its position within the clause. Then I will extend the analysis to other (bare) *wh*-phrases.

##### 4.6.1 Interrogative *que*

The French *wh*-words *que* and *quoi* are synonymous, although they are in complementary distribution. In finite verbs, *que* obligatorily appears in sentence-initial position and triggers subject inversion, as illustrated in (79). *Quoi* on the other hand appears postverbally with finite verbs, as illustrated in (80). With nonfinite verbs, *que* must obligatorily be in sentence-initial position (81a), but surprisingly

*quoi* also can show up there too. We may note that *que* and *quoi* are used only for nonhuman (or nonanimate) arguments.

- (79) a. Que fait Marie?  
what does Marie.  
b. \*Marie fait que?  
c. \*Que Marie fait?  
What does Marie do?
- (80) a. Marie fait quoi?  
Marie does what.  
b. \*Quoi Marie fait?  
c. \*Quoi fait Marie?  
What does Marie do?
- (81) a. Que faire?  
b. Quoi faire?  
What to-do.

It appears clearly that the operators *que* and *quoi* are mostly in complementary distribution. As a minimal hypothesis, suppose that *que* is specified for accusative Case, while *quoi* is Caseless. That *que* is specified for accusative Case may be evidenced by the fact that it cannot be used in isolation, nor as the object of a preposition; in such contexts, only *quoi* is possible, as expected; see (82).<sup>23</sup> This fact is quite parallel to the opposition between strong forms and weak forms in the

---

<sup>23</sup> In (82), *parler* is an A/δ verb, and therefore its internal argument must be introduced by a δ-marker, namely *de*. Even though this instance of *de* is specified for accusative Case, as I argued in the preceding chapter, it may not introduce a (bare) accusative-specified argument. The reason is that the preposition would saturate the Case of the verb, and the Case-specified argument would not be licensed by the verb. We may note further that if this verb were to take a bare accusative element as complement, it would not be differentiated from a true bare-accusative verb. Under the obvious assumption that this piece of information cannot be lost, such a structure may not be grammatical. Thus, the only way to obtain a grammatical structure is to only allow the strong form introduced by the preposition as complement. This is illustrated in (i).

- (i) a. Vous parlez de quoi?  
b. \*Que parlez-vous?  
c. \*De que parlez-vous?

paradigm of personal pronouns, as Bouchard and Hirschbühler point out. Furthermore, since *que* is specified for accusative Case, it cannot be the subject of a verb either, whether the subject is human or nonanimate, as illustrated in (83). (I will return to (83).)

- (82) a. \*De que avez-vous parlé?  
       b. De quoi avez-vous parlé?
- (83) a. \*Que est venu?  
           What came?  
       b. \*Que se passe?  
           What's happened?

The nonappearance of *que* in the complement domain of the verb is accounted for straightforwardly under my assumptions: its specification for accusative Case forces it to move outside the complement domain of the verb by virtue of the OCP. In contrast, *quoi* must remain within the complement domain of the verb, since it is unspecified for Case (but see the discussion of (81b) below).

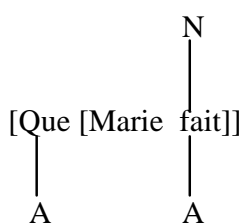
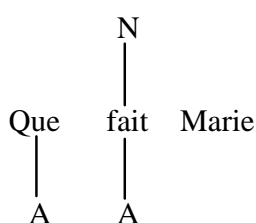
Clearly, since *que* is specified for Case, it may not remain within the complement domain of the verb. However, while pronominal clitics generally adjoin to the first element outside the forbidden domain, *que* moves to clause-initial position. The relevant question is why. In my view, the difference between wh-operators and personal pronouns (operators or not) should be attributed to a feature which is present in one of these elements and absent in the other. Suppose indeed that certain operators, including *que*, are specified for a special feature, call it  $\pi$ , such that it cannot be head-adjoined, that is, it must stay on its initial timing tier. (Recall that head-adjunction automatically puts the adjoined element on a new timing tier.) Under this hypothesis, the clause-initial position of *que* is accounted for straightforwardly: because it is Case-specified, it may not stay within the complement domain of the verb and because it is  $\pi$ -specified it may not adjoin to the first element outside this domain, unlike pronominal clitics.



I would like to suggest that  $\pi$ -specified operators, although they move to clause-initial position, are not adjoined to any elements in the structure. My claim is that such elements are removed from the phrase marker such that they form an independent single-word structure. That is, given the structure in (84a), the removal of the operator from the complement domain of the verb will yield either of the two substructures in (84b) and in (84c), namely ' [que]' and ' [Marie fait]' . The linear order of the substructures need not be stipulated; one of them, namely (84c), is excluded by an independent reason, namely the ban of string-vacuous operations. Indeed, (84c) is not differentiated from the initial structure (84a), and therefore is blocked.

- (84) a. [Marie [fait **que**]]  
 b. [**que**] [Marie fait]  
 c. \*[Marie fait] [**que**]

As example (79a) shows, *que* triggers subject inversion. This is accounted for straightforwardly given the gemination effect induced by the OCP. Since the subject intervenes between the verb and the Case-specified operator, see (85a), it must move to the right of the verb, yielding the subject inversion structure (85b).

- (85) a.  b. 

Note that the NP *Marie* is not head-adjoined to the verb in (85b). I argue that like the operator *que*, it is removed from the phrase marker, and is relocated, or juxtaposed, to the right of the verb. If the subject NP were head-adjoined to the verb in (85b), one would expect it to appear between the auxiliary and the lexical verb in compound tenses. Now, as can be seen in (86c), this is not the case; rather the NP appears to the right of the complex aux+verb (86d). In contrast, subject clitics do appear between the auxiliary and the verb, suggesting that they are head-adjoined to the auxiliary via rightward movement; compare (86a) and (86b). Therefore, one may

conclude that the NP in (85) is not head-adjoined to the verb.<sup>24</sup> The question arises of why subject NPs behave differently from subject clitics with respect to landing site. It may be that the head-adjunction of an NP, which is unspecified for Case, is not possible. If an argument is Caseless, it must absorb the relevant Case of the verb, if any. The fact is, temporal adjunction resembles a morphological process whereby the adjunct is incorporated into the head, i.e. they form a single element scattered between two timing tiers. Under the reasonable assumption that a verb cannot assign Case to a part of itself, the Case of the verb is not absorbed when the NP is head-adjoined, and therefore such a structure is excluded.

- (86) a. Qu' a-~~il~~ fait?  
           what has he done.  
           ' What did he do?'  
       b. \*Qu' a fait~~il~~?  
       c. \*Qu' **Marie** fait?  
       d. Qu' a fait Marie?  
           ' What did Mary do?'

In nontensed clause, the OCP forces the wh-operator *que* to be removed from the structure, as shown in (87).

- (87) a. [pro [ faire que]]  
           |     |  
           A     A  
       b. [ Que][PRO faire]  
           |     |  
           A     A

A potential problem for this analysis comes from the fact that *quoi*, which is unspecified for Case, also appears to the front of an infinitival clause, as exemplified in (81b), repeated as (88). At this time, I have no strong explanation for this fact.<sup>25</sup>

---

<sup>24</sup> Along with normative grammarians, I assume that the morpheme *t* in (86a) is inserted in French inversion, when it is not present underlyingly for spelling reason.

<sup>25</sup> Note that in some contexts *quoi* seems to function as a non-interrogative subject; in such a case, the structure introduces an enumeration, as in (i). If so, *quoi* is then base-generated as subject of the infinitive verb. Perhaps this use of *quoi* may influence the appearance of interrogative *quoi* in preverbal position.

(88)           Quoi faire?

Before concluding this section, we may note that to account for sentences such as (83), Bouchard and Hirschbuhler (1987), following Kayne (1976), Chomsky (1981), use a rule which changes *que* (be it a complementizer or a pronoun) into *qui* in certain contexts. This produces the desired result in the case of (83a), but not in (83b), as can be seen in (89). Bouchard and Hirschbühler then rely on some interpretive constraint to rule out (89b).

- (89)   a.     Qui est venu?  
          b.     \*Qui se passe?

In my view, the *que/qui* rule is not needed (see Desouvrey, 1995/1996, 1997). Indeed, *que* being strictly an accusative pronoun, it may not be used as a subject, any more than the personal pronoun *me*. When the subject is human, another interrogative pronoun is used, namely *qui* (see below). With the verb *se passer* 'to happen', which takes either an expletive or a non-animate subject, *que* cannot be used as a subject, since it is specified for accusative Case; cf. the ill-formedness of (83b), repeated as (90a). But if *que* is the object of the verb, while the subject is an expletive, a well-formed sentence is derived, as shown in (90b). A well-formed structure can also be obtained with a complex operator; see (90c). A detailed analysis of complex operators of this type, which are not lexical, falls far beyond the scope of this thesis; yet we will note that the instance of *qui* which appears here is a relative pronoun, and its antecedent is the operator. This is more obvious in (90d) (see the discussion of clitic doubled operators below).

- (90)   a.     \*Que se passe?  
                   what-ACC SE happens  
          b.     Que se passe-t-il?  
                   what-ACC SE happens it-NOM.  
          c.     **Qu'**est-ce qui se passe?  
          d.     C' **est****quoi** qui se passe?   (colloquial French)

- 
- (i)    Quoi faire aujourd' hui: théâtre, promenade, etc.  
       What to do today: theater, walking, etc.

Summing up, the evidence shows that *que* can be best analysed as a Case-specified element which moves under OCP. I have suggested that the fact that it appears in a clause-initial position is due to its  $\pi$  feature, which prevents it from adjoining to another element within the clause. By virtue of this feature, it must be removed from the structure, so as to remain on the original tier. It appears that the general principles that govern personal pronouns apply to interrogative *que* as well. It is predicted that this element is a clitic, like the weak personal pronouns. Therefore there is no reason to assign it a clitic feature, any more than to personal pronouns. In the next section, I show that this analysis applies to interrogative *qui* as well.

#### 4.6.2 Interrogative *qui*

While *que* is an accusative wh-pronoun replacing an argument denoting a thing, the interrogative pronoun *qui* may only replace an animate argument, which may be either the object or the subject of the clause. It displays the behaviour of both *que* and *quoi*, with a few differences in the detail. Some basic facts can be observed in (91) and (92).

- (91) a. Qui aime Marie?  
Who likes Marie?  
b. Qui vois-tu?  
c. Tu vois qui?  
Who do you see?  
d. À qui parles-tu?  
To whom do you speak?
- (92) a. Jean voit qui?  
b. \*Qui voit Jean?  
c. Qui Jean voit-il?  
d. Qui Jean voit?  
Who does Jean see?

In (91a), *qui* is the subject of the clause; in (91b,c) it is the direct object (in situ in (91c)), while in (91d) it is the indirect object (benefactive). In (92a) it is the direct

object in situ, while the subject is an NP. (92b) shows that the appearance of *qui* in a clause-initial position yields ill-formedness when the subject NP is inverted. Instead, it triggers the appearance of a personal pronoun which doubles the subject; cf. (92c). Notice that the absence of the pronoun leads to ungrammaticality; cf. (92d).

I will attempt to account for these facts in a principled way, using features that are obvious from the data. As a minimal hypothesis, suppose that interrogative *qui* is exactly parallel with interrogative *que* in that it represents two different homophonous morphemes: one is specified for accusative Case, while the other is unspecified for Case. In other words, one instance of *qui* is parallel with *que* and the other is parallel with *quoi*. All the facts given in (91) and (92) strongly support this hypothesis. Indeed (91c) and (91d) are parallel with the examples with *quoi*; both are repeated below in (93) and (94) for convenience. Moreover, the clitic doubling structure in (92b) may be constructed with *que* also, although with the latter it is more marked (or literary) (95).

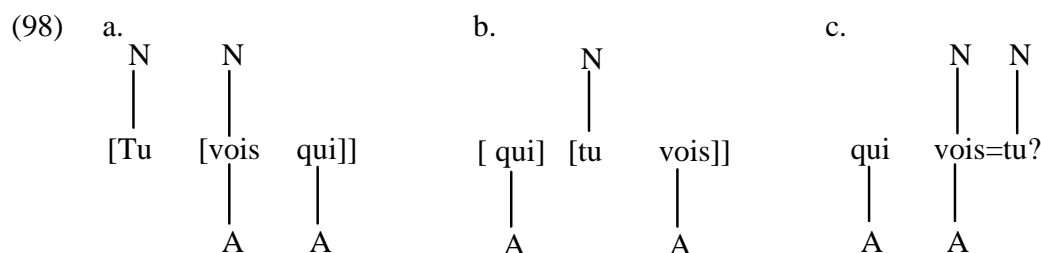
- (93) a. Qui vois-tu?  
b. Tu vois qui?  
c. De qui parlez-vous?
- (94) a. Que vois-tu?  
b. Tu vois quoi?  
c. De quoi parlez-vous?
- (95) a. Qui Jean voit-il?  
b. ?Que Jean voit-il?

Then since *qui* is parallel to *que*, the analysis presented above for *que* must account for the behaviour of *qui* as well. What is not accounted for is a structure such as (95) in which the subject is clitic doubled. Specifically, one must explain why the clitic-doubled construction is rather marginal with *que* and obligatory with *qui*, and above all why clitic doubling is required in such a structure.

I would like to suggest that the NP subject is clitic doubled in wh-questions in order to avoid the ambiguity of *qui*, which can be interpreted either as subject or



no such ambiguity, and use of clitic doubling is quite marked; cf. the unnaturalness of (95b). Furthermore, this analysis is supported by the fact that clitic doubling is possible only if the subject is an NP. Indeed, if the subject is a personal pronoun, which is specified for Case, *qui* may not be confused with the subject; this is illustrated by the derivation in (98): in (98a) the OCP forces the operator to move, yielding (98b), where a gemination effect is created; then the subject clitic (*tu*) moves rightward onto the verb, as can be seen in (98c), which is the output. Since this subject pronoun can never be used as an object, there is no ambiguity, and therefore clitic doubling is excluded. We may recall that clitics must be licensed by Case, and therefore a double clitic is not possible in (98). In contrast, the clitic doubling in (95b) is allowed because the NP need not be licensed by Case, nor the clitic by the lexical structure of the verb.



If this analysis is correct, one may predict that in languages or dialects which do not allow clitic doubling for some reason, another strategy may be used to circumvent this type of ambiguity (due to syncretism). Indeed in certain dialects, including Montreal popular French (see Lefebvre, 1982), there is no clitic doubling in interrogatives. Nevertheless, in this dialect the so-called doubly filled complementizer is required in *wh*-questions; compare the dialectal French (99) with its standard French equivalent (100).

- (99)
- a. Qui qui aime Marie?  
Who likes Mary?
  - b. Qui que Marie aime?  
Who does Mary like?

- (100)
- a. Qui aime Marie?      (same meaning as (99))
  - b. Qui Marie aime-t-elle?

As shown in Desouvrey (1995/1996, 1997), there is no such thing as a doubly-filled complementizer. Rather in the structures in (99) the *wh-qui* is the antecedent of a relative clause, a subject relative clause in (99a) and an object relative clause in (99b). The 'relativization strategy' is possible because relative pronouns are not ambiguous; indeed their phonological shapes vary according to their functions, namely *que* as direct object (accusative) and *qui* elsewhere (no Case). Thus the structures in (99) must be analysed as follows. The relative clause and its head (the *wh*-operator) are generated as two independent structures, as illustrated in (101) (see Desouvrey, 1997). In (101b) the relative pronoun *que* is removed from the structure, since it is specified for accusative Case; (101c) is thus obtained. In (101a) there are no OCP effects, *qui* being a Caseless subject, and it is adjacent to its antecedent. (Both relative clauses are ultimately adjoined to the *wh*-antecedent; see Desouvrey, 1997, for more details on the syntax of French relative clauses.)

- (101) a. [Qui<sub>i</sub>] [qui<sub>i</sub> aime Marie]  
       b. [Qui<sub>i</sub>] [Marie aime que<sub>j</sub>]  
       c. [Qui<sub>i</sub>] [[que<sub>j</sub>] [Marie aime]]

Summarizing, I have suggested that *qui* represents in fact two morphemes, a Case specified morpheme, which is similar to *que*, and an unspecified one, which is similar to *quoi*. Consequently it is expected that *que/quoi* and *qui/qui* behave exactly the same way with respect to the OCP. The fact that the NP subject must be clitic doubled when Case-specified *qui* is the object is essentially due to the ambiguity of the structure resulting from the movement of the Caseless subject rightward, which comes to be interpreted as the object of the clause. Certain dialects use another strategy which consists in relativizing the clause with respect to the interrogative pronoun. In the next subsection, I consider yes-no questions, showing that they are parallel to *wh*-questions, once one assumes that a null operator is present in the structure.

#### 4.6.3 A null *wh*-operator

I have reduced the inversion in interrogatives to a simple OCP effect, more precisely to the geminate-like properties of Case-specified elements which force any



material intervening in between them to move out. This is a costless analysis, since such a mechanism is independently needed and has been used successfully to account for other facts. I would like to extend this analysis to a type of inversion where apparently there is no *wh*-operator. To the extent that this attempt is successful, it constitutes an additional piece of evidence for the analysis presented above. Moreover, the facts to be discussed will make it possible to present a further argument against many traditional analyses which claim that the complex verb+clitic moves past the subject, creating inversion. Consider the simple facts in (102).

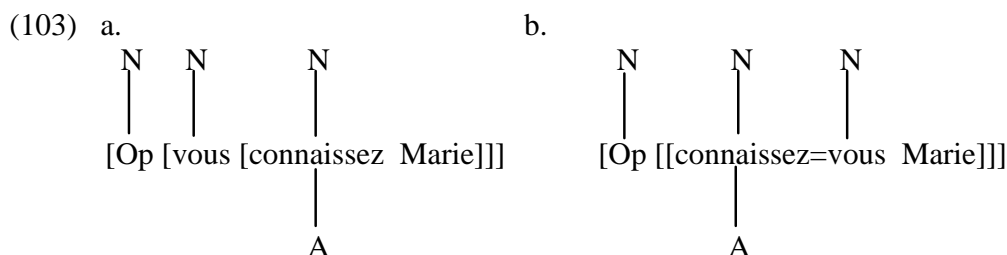
- (102) a. Connaissez-vous Marie?  
           know you Marie.  
           ' Do you know Marie?'  
       b. Paul connaît-il Marie?  
           Paul knows he Marie.  
           ' Does Paul know Marie?'  
       c. \*Connâit Paul Marie?  
           knows Paul Marie.

The sentence (102a), in which the subject clitic is inverted, exhibits a strong parallelism with the *wh*-questions discussed above; so does the structure in (102b), where the subject NP is clitic-doubled. (102c) shows that inversion of an NP subject gives rise to an ill-formed structure, again like the *wh*-questions above. It is unlikely that such a strong parallelism is not induced by the same cause. Thus, as a minimal hypothesis, suppose that French possesses a nonovert *wh*-element, Case-specified, which serves to form yes-no questions. Let us refer to it as *Op*.

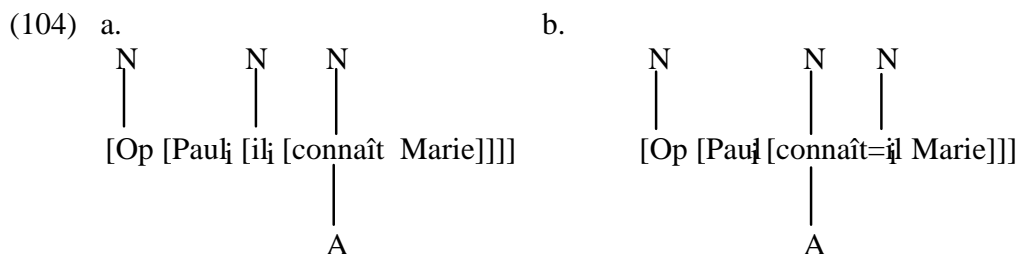
This hypothesis is highly plausible. As mentioned above, all *wh*-operators in French alternate with a complex form, namely *que* vs. *qu' est-ce que/qui*, *qui* vs. *qui est-ce qui/que*. Interestingly, the nonovert operator also alternates with an overt complex operator, namely *est-ce que*. The existence of such an operator may be interpreted as evidence for the existence of the nonovert operator.

Thus, under the hypothesis that there is a null operator, the sentences in (102) are accounted for straightforwardly, just in the same way as the *wh*-questions just

discussed. Indeed, (102a) may have been derived from the structure shown in (103a), assuming that the nonovert operator is specified for nominative Case. The intervening clitic must move to the right edge of the verb, as discussed above, yielding (103b).



Similarly (102b) may have been derived from the initial structure (104a), where the NP subject is clitic-doubled for the same reason discussed above. The NP cannot move, but the clitic can do so. As shown in (104b), it adjoins to the right edge of the verb under the gemination effect created by the pressure of the verb and the abstract operator.

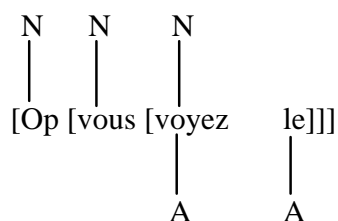


It is clear that in such structures the verb remains in situ, and the movement of the subject is triggered by the gemination effect. However in many traditional analyses of cliticization in French, it is claimed that the object clitic is adjoined to the verb and the complex moves past the subject. The empirical evidence adduced in favor of this claim comes from structures with subject inversion such as (105).

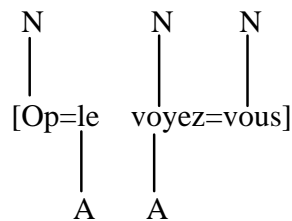
- (105) Le voyez-vous?  
 him-ACC see you-NOM  
 Do you see him?

In my view the subject and the object move to different targets. Recall that my assumption is that (non-semantic type) constraints are not ranked. Except under some locality effect, all elements which have to move do so simultaneously (so to speak), and they adjoin to the first available element, unless some feature makes them incompatible with head-adjunction. Thus under such assumptions, (105) may have been derived from the initial structure (106a). As can be seen, both arguments independently satisfy the requirement for movement: in the complement domain of the verb, the accusative clitic must move, since it induces an OCP violation; the subject clitic must move under the gemination effect, since it is between two Case-specified elements, the verb and the abstract operator. Thus, under the assumption that both elements move simultaneously, the accusative clitic targets the abstract operator while the subject clitic targets the right edge of the verb, yielding the desired structure (106b).

(106) a.

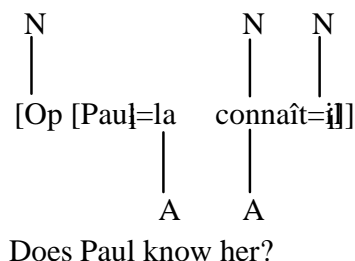


b.



In (106) the object clitic adjoins to the operator, which is the only available site outside the complement domain of the verb. However, if there is another closer element, the clitic will not move up to the operator. Rather, it will adjoin to the intervening element, consistently with my assumptions. Indeed, if in (104a) the direct object is replaced by the relevant clitic, the latter will land on the subject NP, as can be seen in (107). But if the sentence is negated, the clitic will land on the negative morpheme *ne*, as shown in (108).

(107)



- (108) Op Paul ne=la connaît-il pas?  
Does not Paul know her?

To conclude, it is shown that yes-no questions are parallel to wh-questions, once one assumes that there exists a null operator which marks yes-no questions. In the next section I return to superiority effect phenomena, focussing on English.

#### 4.6.4 Superiority effects

This section is mainly intended to provide further support for the analysis of clitic climbing presented above. Recall that, in Spanish, clitics and infinitive verbs are taken to be operators, an assumption which allows enclitics in Spanish to be analysed as a case of superiority effect. Clitic climbing has been shown to be possible because superiority effects do not hold outside the clause. The key to this analysis is the NCC, which rules out the commutation (reordering) of operators within the clause. I consider below the core facts of superiority effects in English, showing that they are consistent with the NCC.<sup>27</sup>

Consider the facts in (109), which illustrate superiority effects (from Aoun and Li, 1993: 31). The initial structure for (109a) and (109b) is given in (110a) and (111a) respectively, assuming the generation of *do* whenever the object operator is

---

<sup>27</sup> Recall that the NCC forbids the reordering (or commutation) of operators within the clause. Thus if all operators are removed from the clause, they in principle may appear in any order, unless some independent constraint is affected. According to our initial observation, this seems to be the case in Serbo-Croatian, where superiority effects are apparently violated (see Rudin, 1988; Boskovic, 1997; and references therein). In my view the relative freedom of ordering of wh-words in Serbo-Croatian may be due to movement out of the initial domain.

intended to be removed from the structure.<sup>28</sup> Assuming that English operators are Case-specified and  $\pi$ -specified, they must be removed from the structure, as opposed to movement to a head, since  $\pi$ -specified elements cannot be head-adjoined. When the higher operator is removed from (110a) and (111a), a gemination effect is created, since the subject is now between the operator and the verb, as shown in (110b) and (111b). Therefore, the subject is moved to the right edge of *do*, yielding the desired result, (110c) and (111c). The two operators being equal with respect to the number of features, the removal of the lower operator, while the higher operator remains in situ, is not possible by virtue of the locality effect discussed in chapter 2. It is indeed such an operation which gives rise to the ill-formed structures in (109).<sup>29</sup>

- (109) a. **Who** did you give **which** check?  
           \***Which** paycheck did you give **whom**?  
       b. **Which** check did you send to **whom**?  
           \***Whom** did you send **which** check to?  
           \***To whom** did you send **which** check?
- (110) a. [you [did [[give **who**] **which** check]]]  
       b. [**who**] [you [did [give **which** check]]]  
       c. [**who**] [did=you [give **which** check]]

---

<sup>28</sup> In the complex *do+lexical verb*, *do* bears the tense of the structure much like auxiliaries in French, and therefore it is specified for nominative Case. However, it is unclear whether it is specified for accusative Case as well. Notice that while the lexical verb is not a participle, it still has accusative Case (or oblique Case).

<sup>29</sup> Under the present analysis, languages are expected to vary according to whether operators are Case-specified,  $\pi$ -specified or both. If a language contains operators which are not specified for Case, no wh-movement is expected. This is the case in Chinese, where each wh-word is used either as subject or object, suggesting unspecification for Case (some basic facts can be found in Huang, 1982). If on the other hand a language contains Case-specified operators, which do not have the feature  $\pi$ , wh-movement is expected, but unlike in French and English, these operators may be head-adjoined like pronominal clitics. This seems to be the case in Polish (and some other Slavic languages). In this language, all wh-words must move; but interestingly they do not necessarily appear in clause initial position. According to Rudin (1988, footnote 4), this is due to the fact that Polish has a topic position which is pre-Spec,CP. This topic position is necessary to save the GB claim that CP is the first position of the clause, and that wh-elements may move only to CP. However, in her glosses of the examples, there is no way to distinguish this topic element from an ordinary subject argument, as illustrated in (i), her (7a).

(i) Boris **na kogo kakvo** kaza [če šte dade — —]?  
       Boris to whom what said [that will give-3s]  
       **What** did Boris say that (he) would give **to whom**?

- (111) a. [you [did [[send [**which** check]] to **whom**]]]  
 b. [**which** check] [you [did [send to **whom**]]]  
 c. [**which** check] [did=you [send to **whom**]]

A further logical possibility must be accounted for: why can both operators not move outside the complement domain of the verb? Apparently this transformation should be possible, since the NCC holds only within the initial domain, namely the clause. The movement of the operators would yield the structures in (112a) and (112b) respectively. I show that these structures are ruled out because they give rise to a series of gemination effects which cannot be repaired, creating sort of an endless loop.

- (112) a. \*[**who**] [**which** check] [you did give]  
 b. \*[**which** check] [to **whom**] [you did send]  
 c. [**who**] [did=you give] [**which** check]  
 d. \*[**which** check] [**who**] [did=you give]

Consider (112a). A gemination effect would take place triggering the movement of all the elements intervening between the auxiliary verb and the leftmost operator, namely *who*. Thus the second operator, the *which*-phrase, would move as well as the subject *you*. The latter could normally be head-adjoined to the right-edge of the auxiliary. But the *which*-phrase could be reordered either to the right of the verb or to the left of *who*. The first case is a string-vacuous movement which yields the correct structure where no movement has taken place (112c); in other words, the derivation would be circular if this movement were possible. The second alternative illustrated in (112d) obviously yields a further gemination effect, forcing the reordering of *who*; and this process would repeat ad infinitum. The same endless loop would be obtained with (112b) as well. Therefore, a well-formed structure cannot be obtained when both operators are removed from the clause.

In the examples above, both *wh*-operators are internal arguments. Consider now the simplest case of superiority effects, where the subject and the object are operators, as illustrated in (113a). Since, normally, subjects do not move under OCP,

the object operator could be removed from the structure. Now, this yields an unwanted result, irrespective of the position of *who*, which would have to move under the gemination effect, as illustrated in (113b-d).

- (113) a.       **Who** saw **what**?  
           b.       \*What did who see?  
           c.       \*What who did see?  
           d.       \*What did see who?

I show that the ill-formedness of (113b) (and (113c)) comes from the endless cycle triggered by the gemination effect. Given the obligatory appearance of *do*, the removal of the object operator yields the structure in (114a), which is composed of two substructures. This structure is consistent with the NCC since the operators are no longer in the same domain. However, (like the subject *you* in (109a)) the subject operator may not stay in its position, due to the gemination effect triggered by the object operator preceding it. Under this condition, *who* could be either reordered to the right of the verb, as in (114b), or to the left of *what*, as in (114c). In the latter case, a gemination effect would take place ad infinitum, excluding this structure. One must explain now why (114b) is ill-formed. I show that (114b) is ruled out because it has more constraint violations than the structure it is intended to repair, namely (113a). In (113a) only the constraint of movement under the OCP is violated, since the Case-specified operator remains within the complement domain of the verb. However, all specified arguments are strictly adjacent with the verb, as required by fusion/gemination effects. This is not the case in (114b), which must be derived from a structure with *do*. Since the auxiliary *do* is specified for nominative Case, while the lexical verb *see* is still specified for accusative Case, each argument in (114b) is inverted with respect to its verb so that the object *what* is separated from the lexical verb (accusative) by the auxiliary, and vice versa, yielding two unrepairable gemination effects, since the verbs cannot move.

- (114) a.       \*[what] [who did see]  
           b.       \*[what] [did see] [who]  
           c.       \*[who] [what] [did see]

The same analysis carries over to structures containing nonargumental operators (i.e. adjuncts, in more familiar parlance), which exhibit superiority effects, as illustrated in (115) (Huang, 1982). The initial structure of (115a), as well as (115b), must be as given in (116a) (the order of the substructures is irrelevant), in which the operator *why* is initially independent of the clause, forming its own substructure.<sup>30</sup> Assuming that this operator is specified for nominative Case, if *what* remains in situ as in (116a), the gemination effect induced by *why* will trigger subject inversion, yielding the output (115a), as shown in (117). However, if *what* is removed from the clause, either (116b) or (116c) is derived. Under the gemination effect, the intermediate wh-elements, *why* in (116b) and *what* in (116c), may not remain in their new positions, and will loop ad infinitum. That is, the movement of *why* in (116b) would yield (116c), and the movement of *what* in (116c) would yield (116b). Observe that if in (116b) *why* is reordered after the verb, the ill-formed structure (115b) is derived, which is parallel to (114b). That is, (115b) contains more constraint violations than the structure with one wh-element in situ: indeed, the accusative operator is no longer adjacent with the lexical verb, nor the nominative operator to the auxiliary verb.

- (115) a. Why did you buy what?  
       b. \*What did you buy why?
- (116) a. [why] [you did buy what]  
       b. [what] [why] [you did buy]  
       c. [why] [what] [you did buy]
- (117) [why] [did=you buy what]

To conclude, I am led to account for superiority effects in English in order to provide evidence for the analysis of clitic climbing. I have shown that wh-operators in situ and postinfinitival clitics in Spanish (and presumably other Romance

---

<sup>30</sup> The generation of the adjunct as an independent substructure, along the lines of Desouvrey (1995/1996), should not be surprising, since adjunct clauses may be freely ordered with respect to main clauses. For instance, compare (ia) and (ib).

- (i) a. [I bought this book] [because I know the author].  
       b. [Because I know the author] [I bought this book].



languages) follow from the NCC, which bans the commutation of operators. English wh-operators are different from Spanish clitics in that they are  $\pi$ -specified, a feature which prevents them from being head-adjoined to an element within the clause.

#### 4.6.5 Summary

I have shown that the syntax of wh-operators obeys the same general principles, which I use to account for personal pronoun clitics. Because wh-operators are Case-specified, they exhibit clitic-like behaviour. They differ from personal pronouns in that they are  $\pi$ -specified elements, hence they may not be head-adjoined. Rather, they must be removed from their structure. It has been shown that the appearance of a clitic-doubled construction is a strategy which is intended to circumvent the ambiguity of wh-operators; in Montreal French another strategy is used, namely the relativization of the clause with respect to the operator. I have discussed superiority effects in English in order to adduce evidence for the analysis of enclitics in Spanish. Given the hypothesis that  $\pi$ -specified elements cannot be head-adjoined, wh-operators in situ in English ( $\pi$ -specified) and enclitics in Spanish (not  $\pi$ -specified) are accounted for straightforwardly with two constraints: the gemination effect (OCP) and the NCC.

#### 4.7 French negation

The basic facts of negation in Standard French are given below. The main negative marker is the complex morpheme *ne pas*, in contrast to other standard Romance languages which have a simple morpheme (on negation in Romance, see Zanuttini (1997) and references therein). I take *pas* to be the actual negative morpheme, assuming that *ne* is only a  $\pi$ -marker, which is likely to be a marker of sentential negation, as discussed in Williams (1994). In Standard French the tensed verb, whether auxiliary or lexical, appears between the two parts of the morpheme, as shown in (118). Both parts of the negative morpheme must precede the lexical infinitive verb, as illustrated in (119). Infinitival auxiliary verbs may appear either between or right after both parts of the negative morpheme, as shown in (120),

nevertheless the first alternative is more natural. (Examples (119) and (120) are taken from Pollock, 1989.<sup>31</sup>)

- (118) a. Paul **ne** connaît **pas** Marie.  
           Paul NE knows not Marie.  
           ' Paul does not know Marie.'
- b. Paul **n'a pas** connu Marie.  
           Paul NE has not known Marie.  
           ' Paul did not know Marie.'
- (119) a. **Ne pas** posséder de voiture en banlieue rend la vie difficile.  
           'NE not to own a car in the suburbs makes life difficult.'
- b. **Ne pas** regarder la télévision consolide l' esprit critique.  
           'n e not to watch television strengthens one' s independence.'
- (120) a. **Ne pas** être heureux est une condition pour écrire des romans.  
           'NE not to be happy is a prerequisite for writing novels.'
- b. **N'être pas** heureux est une condition pour écrire des romans.  
           'NE not to be happy ...'
- c. **Ne pas** avoir eu d' enfance heureuse est une condition pour écrire des romans.  
           'NE not to have had a happy childhood is ...'
- d. **N'avoir pas** eu d' enfance heureuse est une condition pour écrire des romans.  
           'NE to have not had a happy ...'
- e. **N'avoir pas** de voiture en banlieue rend la vie difficile.  
           'NE to have not a car ...'

Pollock (1989) and Chomsky (1991, 1992), relying on Emonds (1978), propose an analysis of adverbs and negatives in which the verb moves to a higher functional

---

<sup>31</sup> According to Pollock, examples (120,b,d,e) ' are usually considered somewhat literary and "recherché" ' (p. 373). However, while this observation seems to be correct in the case of (120b,d) it is not so for (120e), which is unacceptable for most speakers, as Bouchard (1995) points out. This example, as well as the "recherché" (120b,d), will not be discussed here, taking them to belong to an earlier state of the language; see Labelle and Hirschbühler (1994) for a discussion of this fact.

head T (for tense). In these theories, there is a basic structure for all clauses, where phrasal nodes are ordered with respect to each other by stipulation. Indeed, it is assumed that a negative phrase (NEGP) is ordered right after the Tense node (TP). According to Pollock, the negative morpheme in French is composed of a head (*ne*) and a specifier (*pas*). Pollock and Chomsky, as well as many others, use a complex apparatus and a series of assumptions which I will not discuss in detail here since they are unusable in the context of the analysis I am proposing. (See Iatridou, 1990, and Williams, 1994, for a skeptical look on those approaches, as well as Bouchard, 1995, for a comparison of these approaches, as well as his own theory of negation).

In my view, there are no functional heads and no predetermined basic structures for the clause. All we need to know in order to compute the negative morpheme is its inherent features. First, we may note that the negative morpheme in Standard French is clearly a complex morpheme which has two subparts *ne* and *pas*. As seen above, in one context the verb intervenes between each part of the morpheme, and in the other it follows both parts of the negative morpheme. Consistently with this observation, one may assume that underlyingly both morphemes are adjacent, and that for some reason they split in the course of the derivation. From this perspective, for a simple sentence such as (116a), one may posit either the initial structures (121a) or (121b). If (121a) is correct, then one must find the main reason that prevents *pas* from appearing to the right of the verb in the output; on the other hand, if (121b) is correct, one must explain why *ne* always precedes the verb in the output.

- (121) a. [Paul [**ne pas** [connaît Marie]]]  
 b. [Paul [[connaît [**ne pas**]]Marie]]

Clearly, both the initial structures (121a) and (121b) cannot be correct. There is indeed one fact that favors (121a) over (121b): both parts of the negative morpheme precede the infinitive verb, see (119), and one never finds a context where the contrary is possible (*\*posséder ne pas de voiture*). Thus the most natural hypothesis is that both parts of the negative morpheme are generated to the left of the verb, as in (121a) and that they split in the course of the derivation, for a reason which I will attempt to provide.

Since it is clear that the negative morpheme is initially generated to the left of the verb, one may take it to be specified for nominative Case. In my view the Case feature of an element governs its initial position with respect to the verb. Indeed, accusative-specified elements are merged first with the verb and appear to its right (in French and similar languages). Nominative-specified elements, in contrast, appear to the left of the verb. Now, given that the negative morpheme has two parts, we must know whether both morphemes are Case specified. Obviously, if both *ne* and *pas* were specified for Case, one would expect them to behave alike. But clearly this is not so, as can be seen from the data above. Moreover, given fusion, they would be inseparable, since they are daughters of the same node. Thus, one can assume that both *ne* and *pas* cannot be specified for nominative Case; specifically, I take *pas* to be specified for nominative Case. We may note that *ne* may be omitted in current speech, and it may appear in nonnegative contexts, without *pas*. However, its position invariably is to the left of the verb. I will thus assume that it is specified for the feature  $\pi$ , which makes it incompatible with head-adjunction.<sup>32</sup>

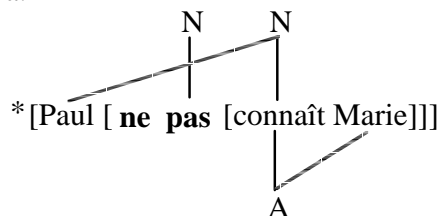
Under this assumption, from the initial structure (121a) one can straightforwardly derive the output structure by movement of *pas* to the right of the verb. The main reason for movement is obviously either the no crossing-lines constraint or the gemination effect, as shown in (122a,b) and (122c,d) respectively. (Observe that as suggested above, the negative morpheme may not saturate the verb, since by temporal adjunction it comes to be a subpart of it.) In compound tenses, the negative morpheme adjoins to the first element to its right, namely the auxiliary, as illustrated in (123) (=118b). The first part of the negative morpheme, *ne*, may not move even when it intervenes between two Case-specified elements (122c,d), since it is  $\pi$ -specified.<sup>33</sup>

---

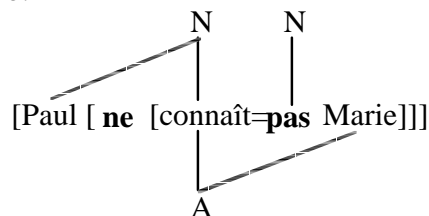
<sup>32</sup> The English negative word *not* seems to be Case-specified (nominative) and  $\pi$ -specified. This seems to be the case for Italian and Spanish also. If so, these languages differ from Standard French in that in the latter  $\pi$ -feature and Case-feature are split into two different morphemes.

<sup>33</sup> This analysis can apply straightforwardly to English. If English *not* is Case-specified and  $\pi$ -specified, as suggested in the previous footnote, it is expected that it cannot intervene between the verb and the subject, nor be head-adjoined to the right edge of the verb, as illustrated in (i). In Desouvrey (in preparation) it is suggested that the computation of *do* is intended to avoid the crossing line constraint. Indeed, the lexical verb Case-marks the object *Monica*, while the auxiliary verb assigns nominative Case to the subject; the negative morpheme remains in situ and its Case does not interact with any other features (ii).

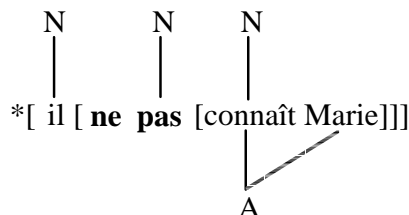
(122) a.



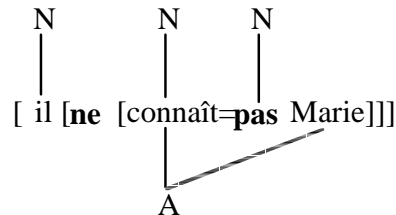
b.



c.



d.

(123) [Paul [ne [[a=**pas** connu] Marie]]]

Since the movement of the negative morpheme *pas* in (122a) is intended to avoid the prohibition on crossing-lines, it is expected that movement should not take place with verbs which do not assign nominative Case. This is borne out: in infinitive verbs movement of the negative does not occur, as shown in (124) (=119b).<sup>34</sup>

- 
- (i) a. \*Bill not kissed Monica.  
 b. \*Bill kissed=not Monica.  
 (ii) [Bill [did [not [kiss Monica]]]]

<sup>34</sup> Bouchard (p.427) points out that *pas* can follow the verb in a contrastive context, indicating constituent negation, as in (i).

- (i) Ne posséder **pas** une, mais deux ou trois voitures devrait être l' ambition de tout bon consommateur.

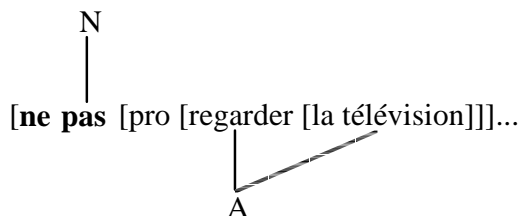
'To own not one, but two or three, cars should be the ambition of all good consumers.'

I agree with Bouchard that this is a real case of constituent negation, where the morpheme *pas* modifies *une (voiture)*. However, it seems that this type of negation cannot be analysed as being syntactically derived. Rather, it seems to be a case of negative concord between the verb and the argument. That is, *pas* and *une* form a constituent, with which the verb agrees via the morpheme *ne*. Indeed, it is well-known that in French, when the subject or the object is a negative polarity item, the verb must agree with it, as illustrated in (ii). In those examples, *ne* is not a negative marker, for it cannot cooccur with *pas*.

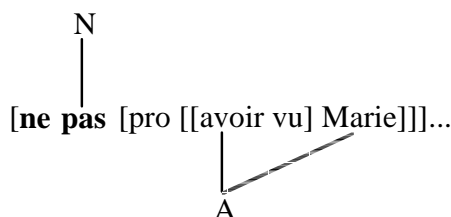
- (ii) a. Nul **n'**est (\*pas) heureux.  
 Nobody is happy.  
 b. Personne **n'**est (\*pas) venue.  
 Nobody came.

We may note, furthermore, that *pas* may negate the subject, in which case negative concord takes place normally, as illustrated in (iiia), which is parallel to (ia), except that the negated constituent is the subject. In contrast, *ne* may not negate the object, and *pas* cannot be used to express negative

(124) a.



b.



Now the negated infinitive apparently poses a problem for the present account of cliticization. The morpheme *pas*, being specified for Case, is expected to appear only with a verb specified for a relevant Case. In fact, this condition may not apply in this context. In pronominal paradigms, where there is an opposition between Case-specified morphemes versus Caseless morphemes, the Blocking Principle allows the latter to be excluded whenever the verb is specified for the relevant Case, unless a nuance of meaning is needed (focus for instance). Since there is no evidence that the morpheme *pas* alternates with an unspecified form, one may assume that its Case does not matter in a derivation such as (124).

Moreover, in these structures nominative Case is irrelevant as long as it does not clash with another feature of the verb. In this respect, it is convenient to take the absence of Case in infinitive verbs to be a null feature, as suggested in chapter one. If so, in (124) the null feature of the infinitive verb is checked with the null feature of the nonovert subject PRO, and therefore the nominative Case of the negative morpheme can be ignored, as long as it does not interact with any valid Case of the verb.

---

concord, as in (iii). (Example (iii) may be seen as evidence that the actual negation marker in French is *pas*.)

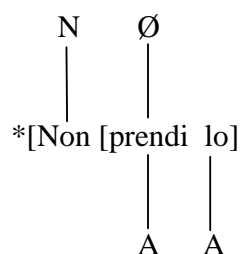
- (iii) a. [pas un [n' [est venu]]]  
 b. \*[pro [pas posséder [n' une]]]

The evidence for this view is provided by certain peculiarities of Italian and Spanish imperatives. In Italian, positive imperatives are expressed differently from negated imperatives, as shown in (125) and (126) (Haegeman, 1995: 121). That is, the proper imperative morphology, which is used in (125a), is incompatible with negation (125b), while infinitive morphology is not (126).

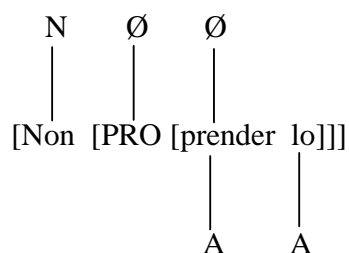
- (125) a.     Prendilo!  
              take-it  
      b.     \***Non** prendilo!  
              Not take it.
- (126) a.     **Non** lo prendere!  
              non it take  
      b.     **Non** prenderlo!  
              non take it

Haegeman reports that Zanuttini (no date) argues that there is a relation of dependency between TP and NegP so that the absence of TP entails the absence of NegP too. This claim is consistent with the GB framework, where the phrase marker contains predetermined functional positions. Such an analysis is incompatible with the present theory, however. Indeed a more natural explanation is available, once the following assumption is made: (a) in Italian, imperative verbs lack nominative Case, as in Spanish; (b) the negative morpheme *non* is specified for nominative Case, parallel to its French counterpart. Thus, when the negative morpheme is merged with the VP, the lack of features of the verb is interpreted as a null feature, which yields a clash since the verb does not include the features of the negative morpheme (N vs Ø, as shown in (127a). Given this, use of the infinitive verb is clearly understood: it appears to be a strategy to avoid the clash of feature. Moreover, the nonovert subject, PRO, which is Caseless, checks the null feature of the infinitive verb, and the nominative Case of the negative morpheme is ignored, just like in French, as shown in (127b). Note that since Italian infinitive verbs are operators, like their Spanish counterparts, the clitic may climb to the negative morpheme, yielding (127c).

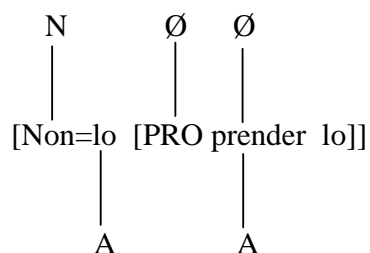
(127) a.



b.



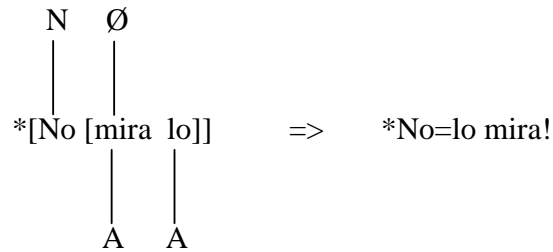
c.



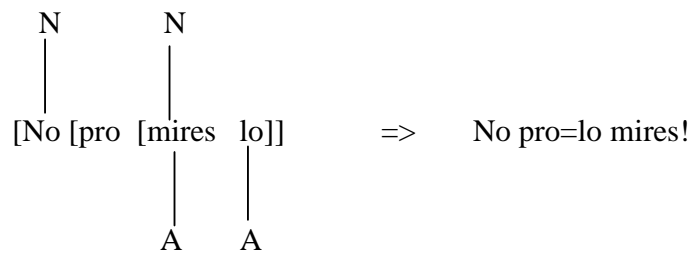
The fact that Spanish negated imperatives are constructed with subjunctive morphology follows from the same reasoning. The same feature mismatch (N vs.  $\emptyset$ ) appears in Spanish, as exemplified in (128a), since imperative verbs are not specified for nominative Case, as discussed earlier. By using the subjunctive morphology which has nominative Case and a null subject slot (*pro*), the mismatch is avoided, as seen in (128b). Both N Case ultimately enters in a fusion process (not shown in schema), which may not have further consequence on the structure, since the null argument *pro* cannot move. Thus the same problem exists in both Spanish and Italian, but the repair strategy is slightly different: Italian uses infinitive morphology in negated imperatives, while Spanish uses subjunctive morphology.



(128) a.

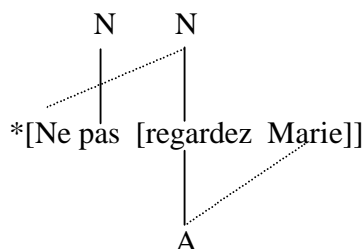


b.

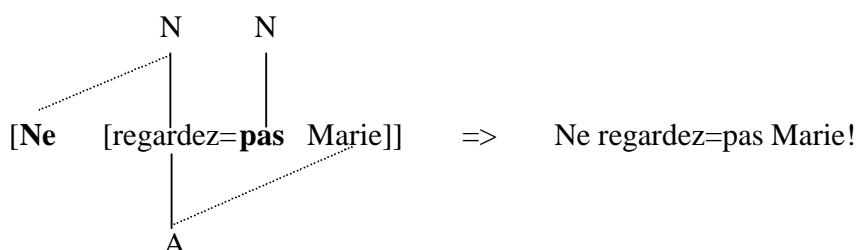


Note that in French, movement of the negative morpheme normally takes place in imperative verbs. The fact is, even though there is no subject, overt or null, in this mood, the morpheme *ne* allows the verb's Case to be discharged. If *pas* remains in situ, there is line-crossing and therefore ill-formedness obtains. The derivation of a negated imperative verb is shown in (129).

(129) a.



b.



To conclude, I have shown that the negative morpheme is generated to the left of the verb, presumably because it is specified for nominative Case. With finite verbs, the negative morpheme *pas* moves to the right edge of the verb in order to avoid the crossing-lines constraint, or to allow a fusion of Cases, if the subject is a clitic. In non-finite verbs, the negative morpheme is correctly predicted to remain in situ, since the verb is not specified for nominative Case. I suggest that the lack of features may be interpreted as a null feature, a decision which permits us to account for certain peculiarities of Italian and Spanish negated imperatives.

#### 4.8 Summary

In this chapter, I have provided cross-linguistic evidence for the analysis developed for French in the preceding chapters. I have shown that the OCP predicts that object personal pronouns in English must not appear before the verb, contrary to their Romance Counterparts. The same prediction appears to be correct for Haitian Creole, where pronominal elements do not exhibit clitic-like behaviour, because of their lack of Case specification.

My analysis carries over to other Romance languages, such as Spanish and Italian, the pronoun systems of which have been studied in some detail. Opaque clitics in these languages are accounted for in a highly principled way, namely the need for the following clitic to be Case-licensed by the verb. The tools I have used, namely feature delinking, make it possible to describe the nature of the process, and more importantly, to explain why it must take place in these languages. Furthermore, I have shown that clitic doubling occurs in Spanish grammar because there is a feature mismatch at some point of the derivation, where a verb requires the object to agree with two different features, for instance animacy and benefactiveness, animacy and specificity. I have also shown that the *me-lui* constraint follows from the Animacy Hierarchy.

Spanish and Italian differ from French in that the former require enclisis with infinitive verbs. I account for this fact by making the hypothesis that infinitive verbs and clitics are operators in Spanish and Italian. As such their relative positions must be preserved within the clause, in order to avoid superiority effects. Clitic climbing, which occurs in these languages, follows from the analysis that superiority effects hold only within the embedded domain.

The analysis presented for pronominal clitics has been extended to wh-operators. It was shown that certain wh-operators in French are Case-specified, and therefore they must move under the OCP outside the complement domain of the verb. The fact that they appear in a clause-initial position is accounted for with the hypothesis that they are specified for the feature  $\pi$ , which makes them incompatible with head adjunction, like other operators. Rather they are removed from the clause. The linear order of the moved operator and the rest of the clause follows from general principles regulating the interaction of features, namely the various effects of the OCP: movement, fusion, and gemination effects.

Finally, the analysis has been extended to negation. Focusing on French, I have argued that the actual negative morpheme in French is *pas*, *ne* being a  $\pi$ -marker. I have shown that the negative morpheme *pas* moves in finite clauses, either under a gemination effect or the ban on crossing lines. In nonfinite clauses, *pas* cannot move, because infinitive verbs are not specified for nominative Case.

## CONCLUSION

In this thesis, I have attempted to present an account of pronominal clitics in Romance, focussing mainly on French facts. While most traditional analyses consider clitics to be an ontological category, i.e. a lexical entry parallel with verbs, for example, I derive all aspects of pronominal clitics from feature specifications, mainly Case features. Under this approach it is expected that any type of grammatical element may exhibit clitic properties provided that it is specified for some relevant features.

Features are central in this thesis. The features that I have taken into account are Case features, thematic features, phi-features, etc. Features are shown to constitute a tree structure, in which the lexical entry is the root node (see also Bonet, 1991, etc.).

Methodologically, whenever a given problem is considered in the pronominal system of a language, the approach consists in carefully determining the features which the pronouns, and all relevant elements, are specified for. On the assumption that a Case-specified element has a more limited distribution than a Caseless element, it is possible to attain a high degree of accuracy in the inventory of features. My approach to features is radically different from current assumptions. Indeed, in the Minimalist Program, features are assumed to be present in all elements. Therefore they cannot be distinctive, and their interaction is unclear. In fact, in the GB-MP framework, this problem is circumvented by the claim that features can be strong or weak, a distinction which is not needed in my account.

A main feature of my analysis is the use of the formalism of nonlinear phonology augmented by a constraint-based approach to grammatical principles. The constraints used in this analysis are formulated in terms of features, which make them highly plausible and less arbitrary than the principles encountered in current theories.

I have distinguished two types of constraints: semantic-type constraints and structural-type constraints. The former take precedence over the latter in any conflicting situation; however, constraints of the same type are not ranked; they all apply in parallel in the derivation, and one cannot be sacrificed to allow another to be applied.

Nonlinear representation makes it possible to shed new light on the nature of the constraints which regulate the phenomenon of cliticization. The main constraint appears to be the Obligatory Contour Principle (OCP), which can be seen as a family of constraints, given its various effects on the syntactic structure. The OCP-type effects are given in (1).

- (1) OCP-type constraints
- a. Movement;
  - b. Fusion;
  - c. Gemination effects.

These account for the following facts: (a) a pronominal element which is specified for Case cannot stay within the complement domain of the verb, that is, it is a special clitic in the sense of Zwicky; (b) a pronominal element specified for Case must be adjacent to the verb selecting it; and (c) no lexical element may intervene between two Case-specified elements.

Another central constraint which depends crucially on autosegmental representations is the checking process, which requires absolute compatibility between the features of the verb and those of the argument. The claim is that, at the very least, each pair of features of a verb and its arguments must be compatible in order to form a tier. If the features do not match, one of them must be deleted via a delinking process, an operation which is constrained by the structure preservation principle. That is, the delinking of any nodes in the feature structure must yield an existing output and not a random element. The delinking process triggered by the failure of the checking constraint has been shown to be responsible for the appearance of the strong form, instead of the weak forms, in imperative verbs (cf. *\*regardez-me* vs. *regardez-moi* ).

Also, the delinking process has been shown to be activated by the Clitic Licensing Constraint, which requires that all clitics be licensed by a verb which is specified for a relevant and valid Case. This occurs when the first argument which is merged with a ditransitive verb saturates it completely. That is, the argument is specified for all features of the verb. If the next internal argument is an NP it is licensed by being required in the lexical structure of the verb. However, if it is a clitic, a delinking process occurs automatically, deleting the redundant feature of the first argument. The delinking process which takes place under the Clitic Licensing Constraint is responsible for the spurious *se* in Spanish, as well as opaque clitics in Italian.

Not all constraints crucially rely on nonlinear representations. Indeed, the Animacy Hierarchy, which I assume, is currently used in other frameworks, including Functional Grammar. This constraint makes it possible to account for the *me-lui* /I-II effects. These clusters are universally impossible because elements which are specified for identical animacy features (*thing* or *animate*) may not be ordered. Apparent violation of these constraints by some Spanish speakers has been shown to be due to the lack of an animacy specification in first person pronouns.

- (2) Animacy Hierarchy
- a. [animate] > [thing]
  - b.  $\emptyset$  > [animacy] >  $\emptyset$
  - c. \*[animate] > [animate]
  - d. \*[thing] > [thing]

Similarly, the constraint against the commutation of operators (Non-Commutativity Constraint) does not rely on the nonlinear representation. This constraint allows superiority effects, including encliticization to infinitive verbs and clitic climbing in Spanish, to be accounted for. This constraint is formulated in (3).

- (3) Non-Commutativity of Operators Constraint
- Given a domain D and operators  $O_1 > O_2 > O_3$ , etc., the order of the operators cannot change within D.

I have proposed a highly plausible analysis of clitic order and clitic placement. I have argued that there are no clitic positions. All arguments are merged with the verb according to the same general constraints which act at the computational stage of the structure. Clitics and lexical NPs appear to the left of the verb if they are subjects and to the right if they are objects. The verb is merged first with its objects and then with its subject. The order of the computation of clitic arguments in a ditransitive verb depends on the Animacy Hierarchy and the feature inventory. In French, the direct object clitic is generally computed first (V>DO>IO), while in Spanish and Italian it is computed after the indirect object (V>IO>DO). In French, where verbs are not specified for animacy, one encounters either the surface order IO>DO (cf. *Marie me le montre*) or the order DO>IO (cf. *Marie le lui montre*), a variation which remains a mystery in current analyses. In this thesis, it is accounted for straightforwardly; when both the DO and the IO are equal in features, the locality which is inherent to the tree structure forces the element which is more embedded (the DO) to move first, yielding the surface order DO>IO. When on the other hand the arguments are not equal in features, the one which has more features offending the OCP must move first. Thus, first and second person clitics which are doubly specified for Case override the locality effect, moving first, and yielding the order IO>DO. This constraint is referred to as the Worst Evil First Constraint (4), and it interacts with another constraint, namely the ban on string vacuous operations in the derivation of linguistic expressions.

(4) Worst Evil First Constraint.

Given a structure where two elements independently violate a constraint, the more offending element must move before the other.

I have accounted for the placement of clitics in a very simple and natural way. Subject clitics do not move and therefore appear in the position where they have been generated, namely to the left of the verb. Object clitics in contrast move outside the complement domain of the verb, targeting the first element encountered. The latter may be either the subject (nominal or clitic), an adverb (basically negation), a null pronominal, or whatever element may legitimately appear to the immediate left of the verb. Under this analysis, the fact that clitics do not move in imperative verbs is

straightforwardly explained by the lack of an adjunction site, under the assumption that imperative verbs lack the subject slot (at least in Romance).

I have proposed a new analysis of the adjunction operation, referred to as temporal adjunction, which is conceptually very simple. Temporal adjunction consists of removing one element from the initial timing tier and attaching it to the closest edge of the head, with a single association line. This yields a nonlinear structure which is further levelled (or flattened) via the Levelling Convention operating at the end of the derivation, at the interface between the grammar and the articulatory system. This results in a substantial benefit: the relative linear order of the head and the adjunct need (and may) not be stipulated by the analyst, as in current analyses; the adjunct will always appear to the right of the head. Not all elements can be head-adjoined. I argue that in languages like French and English, operators must remain on the initial tier of the structure. Instead of being head-adjoined, they are removed from their place, and relocated to another edge of the structure, normally that opposite to their initial position, so as to avoid string-vacuous movement.

Under my analysis, the phenomenon of clitic doubling which is frequent in Spanish was accounted for straightforwardly. Given that clitics need not occupy a slot in the argument structure of the verb, they can be used to avoid a clash between two elements at some stage of the computation of the structure. This occurs when the argument must agree with two features of the verb; for instance an NP is merged with an agreement marker, forming a phrase, which must be merged in turn with a further agreement marker; consistently with the Constraints on Merge, given in (5), such a computation is ruled out.

- (5) Constraints on Merge (COM)
- A head X and an element Y can be merged if either (a), (b), or (c) holds:
- a. X and Y bear the same features.
  - b. X bears some feature, but Y does not bear any feature.
  - c. Neither X nor Y bear any feature.
- Merger is impossible if:
- d. X and Y each bear different features.



It appears that my analysis makes it possible to reconcile in a highly principled way two divergent properties of clitics: on the one hand they may be an argument, filling a slot in the argument structure of a verb, and on the other they may be used as agreement morphemes, cooccurring with the argument they refer to. This outcome is ensured by the Argument Licensing Constraint.

(6)           Argument Licensing Constraint

a.       Clitic licensing

A clitic is licensed (i) by a verb with a valid Case and (ii) either by the lexical structure of the verb or by coindexation with an argument of the verb.

b.       NP licensing

An NP is licensed by the lexical structure of a verb.

This thesis covers important facts in Romance cliticization. I have shown that other elements, including *wh*-words and negation, may be analysed with the same general principles, which are based essentially upon feature and morpheme inventories. This type of analysis can account for all other grammatical phenomena with the same accuracy and without parameterizing or ranking the constraints in individual languages. To the extent that this analysis is more simple and covers more facts in a highly principled way, the burden of proof now rests on the shoulders of proponents of more complex accounts.

## REFERENCES

- Ackema, Peter and Neele Schoorlemmer (1995). Middles and nonmovement. *Linguistic Inquiry*, 26, 173-197.
- Adams, Marianne (1987). From old French to the theory of pro drop. *Natural Language and Linguistic Theory*, 5, 1-32.
- Anderson, Stephen (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.
- Anderson, Stephen (1996). How to put your clitics in their place, or why the best account of second-position phenomena may be something like the optimal one. *The Linguistic Review*, 13, 165-191.
- Aoun, Joseph, and Audrey Li (1993). *Syntax of scope*. Cambridge, MA: MIT Press.
- Aoun, Joseph, N. Hornstein, and D. Sportiche (1981). Some aspects of wide scope quantification. *Journal of Linguistic Research*, 1, 3, 69-95.
- Archangeli, Diana (1988). Aspects of underspecification theory. *Phonology*, 5, 183-208.
- Aronoff, Mark (1976). *Word formation in generative grammar*. Cambridge: MIT Press.
- Auger, Julie (1994). On the nature of subject clitics in Picard. In Mazzola, Michael (ed.), *Issues and theory in Romance Linguistics: Selected papers from the linguistic symposium on Romance Languages XXIII*. Washington, D.C.: Georgetown University Press. 159-179.
- Auger, Julie (1995). Les clitiques pronominaux en français informel: une approche morphologique. *Revue québécoise de linguistique*, 24, 21-59.
- Authier, J.-Marc and Lisa Reed (1996). On the Canadian French middle. *Linguistic Inquiry*, 27, 513-523.
- Baker, Charles L. (1991). The syntax of English *not*: The limits of core grammar. *Linguistic Inquiry*, 22, 3, 387-429.
- Baker, Marc C. (1997). Thematic roles and syntactic structure. In Haegeman, Liliane (ed.), *Elements of grammar: handbook of generative syntax*, 73-137. Dordrecht: Kluwer Academic Publishers.

- Barbosa, P. (1996). In defense of right-adjunction for head movement. In DiSciullo, A.-M. (ed.), *Configurations: essays on structure and interpretation*, 161-183. Somerville: Cascadilla Press.
- Beletti, Adriana and Luigi Rizzi (1988). Psych-verbs and theta-theory. *Natural Language and Linguistic Theory*, 6, 291-352.
- Bessler, P., S. Cummins, D. Heap, T. Nadasdi, and Y. Roberge (1992). Cliticization morphologiques ou catégories fonctionnelles. In Dyck, Carrie, Jila Gomeshi and Tom Wilson (eds.), *Proceedings of the 1992 Annual Conference of the Canadian Linguistic Association*, 28-42.
- Blake, Barry J. (1994). *Case*. Cambridge: Cambridge University Press.
- Bok-Bennema, Reineke, and Kampers-Manhe Brigitte (1994). Transparency effects in the Romance Languages. In M. Mazzola (1994), *Issues and theory in Romance linguistics: Selected papers from the linguistic symposium on Romance languages XXII 1993*. Washington, DC: Georgetown University Press.
- Bonet, Eulalia (1991). *Morphology after syntax: Pronominal clitic in Romance*. Doctoral dissertation, MIT.
- Bonet, Eulalia (1995a). The where and how of clitic order. *Revue québécoise de linguistique*, 24, 61-81.
- Bonet, Eulalia (1995b). Feature structure of Romance clitics. *Natural Language and Linguistic Theory*, 13, 607-647.
- Borer, Hagit (1984). *Parametric syntax. Case studies in Semitic and Romance languages*. Dordrecht: Foris Publications.
- Borer, Hagit, ed. (1986). *The syntax of pronominal clitics* (=Syntax and semantics 19). Orlando, FLA: Academic Press.
- Boskovic, Zeljko (1997). Superiority effects with multiple wh-fronting in Serbo-Croatian. *Lingua*, 102, 1-20.
- Bouchard, Denis (1984). *On the content of empty categories*. Dordrecht: Foris.
- Bouchard, Denis (1994). *Heading: the wrong way*. Ms. UQAM.
- Bouchard, Denis (1995). *The semantics of syntax*. Chicago: The University of Chicago Press.
- Bouchard, Denis and Paul Hirschbuhler (1987). French Quoi and its clitic Allomorph Que. In Neidle, Carol and Rafael A. Núñez Cedeño (eds.), *Studies in Romance Languages*, 39-60. Providence, USA: Foris Publications.
- Brandi L. and P. Cordin (1989). Two Italian dialects and the null subject parameter. In Jaeggli, O. and K. Safir (eds.), *The null subject parameter*, 111-142. Dordrecht: Kluwer.

- Brody, Michael (1997). Perfect chains. In Haegeman, Liliane (ed.), *Elements of grammar*. Boston: Kluwer Academic Press.
- Bromberger, Sylvain and Morris Halle (1989). Why phonology is different. *Linguistic Inquiry*, 20, 51-70.
- Burston, Jack (1983). Clitic object sequence and cooccurrence restrictions in French. *Linguistic analysis*, 11, 3, 247-275.
- Buzio, Luigi (1986). *Italian syntax. A Government and Binding approach*. Reidel Publishing Company: Dordrecht.
- Cadely, Jean-Robert (1994). *Aspects de la morphologie du créole haïtien*. Thèse de doctoral. UQAM.
- Chomsky, Noam (1973). Conditions on transformations. In Stephen Anderson and Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232-286. New York: Holt, Rinehart and Winston.
- Chomsky, Noam (1980). On Binding. *Linguistic Inquiry*, 11. 1-46.
- Chomsky, Noam (1981). *Lecture on government and binding*. Dordrecht: Foris.
- Chomsky, Noam (1986a). *Knowledge of language, its nature, origin and use*. New York: Praeger.
- Chomsky, Noam (1986b). *Barriers*. Cambridge MA: MIT Press.
- Chomsky, Noam (1991). Some notes on economy of derivation and representation. In Freidin, R. (ed.), *Principles and parameters in comparative grammar*, 417-454. Cambridge: MIT Press.
- Chomsky, Noam (1992). A minimalist program for linguistic theory. *MIT working papers in linguistic 1*.
- Chomsky, Noam (1994). Bare phrase structure. *MIT working papers in linguistic 5*.
- Chomsky, Noam (1995). *The Minimalism Program*. Cambridge MA: MIT Press.
- Chomsky, Noam (1998). *Minimalist inquiries*. Ms. MIT.
- Chomsky, Noam and Morris Halle (1968). *The sound pattern of English*. New York: Harper & Row.
- Clements, George N, James McCloskey, Joan Maling and Annie Zaenen (1983). String-vacuous rule application, *Linguistic Inquiry*, 14, 1-17.

- Cummins, Sarah and Yves Roberge (1994). A morphosyntactic analysis of Romance clitic constructions. In Mazzola, Michael (ed.), *Issues and theory in Romance Linguistics: Selected papers from the linguistic symposium on Romance Languages XXIII*. Washington, D.C.: Georgetown University Press, 239-257.
- Dahl, Östen and Kari Fraurud (1996). Animacy in grammar and discourse. In Fretheim, Thorstein, and Jeanette K. Gundel (eds.), *Reference and referent accessibility*. Philadelphia: John Benjamins Publishing Company.
- Desouvrey, Louis-H. (1994). *Les classes de verbes en langue des signes québécoise*. Master thesis, UQÀM.
- Desouvrey, Louis-H. (1995/1996). *Syntaxe non-linéaire et interaction de contraintes: aspects de la relativization en français*. Predoctoral exam, UQAM.
- Desouvrey, Louis-H. (1996a). Effet du PCO en syntaxe: le mouvement des clitiques en français. To appear in *Actes des onzièmes journées de linguistique*. Université Laval, Québec.
- Desouvrey, Louis-H. (1996b). *Case tier, clause structure and the nature of the complementizer trace effect*. Ms. UQAM.
- Desouvrey, Louis-H. (1997). Relativization in French without complementizer. *Proceedings of the 1996 annual conference of the Canadian Linguistic Association*. University of Calgary Working Papers in Linguistics.
- Desouvrey, Louis-H. (2001a). Effects of Ambiguity in Syntax: A Study of Wh-Operators in English and French. Ms.
- Desouvrey, Louis-H. (2001b). Adverbs, Negation and OCP Effects. Ms.
- Desouvrey, Louis-H. (2002). The Proper Treatment of Coreference Relation. Ms., <http://semanticsarchive.net>.
- Desouvrey, Louis-H. (2005). Romance Clitic Clusters: The Case Connection, In Heggie Lorie, and Francisco Ordonez (eds), *Clitic and Affix Combinations: Theoretical Perspectives*. John Benjamins.
- Di Sciullo, Anna Maria (1982). Opérateurs et liage: cl, AGR. *Revue de l' association québécoise de linguistique*, 2, 2, 47-56.
- Di Sciullo, Anna Maria (1983). Théories du liage et expressions idiomatiques. *Revue de l' association québécoise de linguistique* 3, 2, 51-58.
- Di Sciullo, Anna Maria (1991). Multi-level saturation. *MIT Working Papers in Linguistic*, 19.

- Di Sciullo, Anna Maria (1998). *Local asymmetries: consequences for morphological configurations and paradigms*. Paper distributed at the 8th International Conference on Morphology, Budapest.
- Di Sciullo, Anna Maria (1999). The local asymmetry connection. In *The local asymmetry connection and lexical knowledge*. MIT Working Papers in Linguistic.
- Di Sciullo, Anna Maria and Edwin Williams (1985). Noun incorporation vs. Cliticization. *NELS*, 16.
- Di Sciullo, Anna Maria and Edwin Williams (1987). *On the definition of word*. Cambridge: MIT Press.
- Di Sciullo, Anna-Maria (1990). On the properties of clitic. In DiSciullo, A.M. and Anne Rochette (eds.), *Binding in Romance. Essays in honour of Judith McA' nulty* 209-232. Ottawa: Special publication of the Canadian Linguistic Association.
- Di Sciullo, Anna-Maria (1995). X-bar selection. In Roorick, Johan and Laurie Zaring (eds.), *Phrase structure in the lexicon*. Dordrecht: Kluwer.
- Dufresne, Monique (1993). *L' articulation syntaxique et phonologique de la cliticisation: le cas des pronoms sujets en moyen français*. Doctoral dissertation, UQAM.
- Emonds, Joseph (1978). The verbal complex V' -V in French *Linguistic Inquiry*, 9, 151-175.
- Everett, Daniel (1996). *Why there are no clitics*. Stanford CA: Center for the Study of Language and Information.
- Fagan, Sarah M. (1988). The English middle. *Linguistic Inquiry*, 19, 181-203.
- Fagan, Sarah M. (1992). *The syntax and semantics of middle constructions*. Cambridge: Cambridge University Press.
- Givón, Talmy (1984). *Syntax: a functional typological-introduction*, Vol. 2. Amsterdam: John Benjamins.
- Goldsmith, John (1976). *Autosegmental phonology*. Doctoral dissertation, MIT.
- Goldsmith, John (1990). *Autosegmental and metrical phonology*. Cambridge MA: Blackwell.
- Grimshaw, Jane (1982). On the lexical representation of Romance reflexive clitics. In Bresnan, Joan (ed.), *The mental representation of grammatical relations*, 87-148. Cambridge MA: MIT Press.
- Grimshaw, Jane (1991). *Argument structure*. MIT Press, Cambridge.

- Grimshaw, Jane (1997a). Projection, head and optimality, *Linguistic Inquiry*, 28, 373-422.
- Grimshaw, Jane (1997b). The best clitic: constraint conflict in morphosyntax. In Haegeman, Liliane (ed.), *Elements of Grammar*. Kluwer Academic Publishers, Boston.
- Gross, Maurice (1968). *Grammaire transformationnelle du français: syntaxe du verbe*. Paris: Larousse.
- Gruber, Jeffrey (1996). *Lexical structures in syntax and semantics*. Amsterdam: North-Holland Publishing Company.
- Guerssel, Mohamed (1977). Constraints on phonological rules. *Linguistic Analysis*, 3, 267-305.
- Guerssel, Mohamed (1978). A condition on assimilation rules. *Linguistic Analysis*, 4, 225-254.
- Guerssel, Mohamed (1995). Berber clitic doubling and syntactic extraction. *Revue québécoise de linguistique*, 24, 1, 111-133.
- Guerssel, Mohamed (no date). *Some notes on the structure of Berber*. Ms.
- Haegeman, Liliane (1995). *The syntax of negation*. Cambridge: Cambridge University Press.
- Halle, Morris (1992). Phonological features. *International encyclopedia of linguistics*, 3, edited by W. Bright, 207-212. Oxford: Oxford University Press.
- Halle, Morris and Alec Marantz (1993). Distributed morphology and the pieces of inflection. In Hale, Kenneth and Samuel Keyser (eds.), *The view from building 20*, Cambridge MA: MIT Press, 111-176.
- Halle, Morris and Jean-Roger Vergnaud (1980). Three-dimensional phonology. *Journal of Linguistic Research*, 1, 83-105.
- Halpern, Aaron (1995). *On the placement and morphology of clitics*. Stanford CA: Center for the Study of Language and Information.
- Hayes, Bruce (1986). Inalterability in CV phonology. *Language*, 62, 321-351.
- Hayes, Bruce (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20, 253-306.
- Heap, David (to appear). *Optimizing Romance clitic sequences*. In Proceedings of LSRL, XXVI, Mexico city, March 1996.
- Hornstein, Norbert (1996). *The logical form: from GB to Minimalism*. Oxford: Blackwell.

- Hornstein, Norbert and Amy Weinberg (1995). The Empty Category Principle. In Webelhuth, Gert (ed.), *Government and Binding Theory and the Minimalism Program*, 241-296. Oxford: Blackwell.
- Huang, C. T. James (1982). 'Move wh in a language without wH movement' *The linguistic Review*, 1: 369-416.
- Huang, James (1995). Logical Form. In Webelhuth, Gert (ed.), *Government and Binding Theory and the Minimalism Program*, 125-175. Oxford: Blackwell.
- Hyman, Larry (1985). *A theory of phonological weight*. Dordrecht: Foris.
- Iatridou, Sabine (1990). About AGR(P). *Linguistic Inquiry*, 21, 4, 551-577.
- Jackendoff, Ray (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jaeggli, Osvaldo (1982). *Topics in Romance syntax*. Dordrecht: Foris.
- Janse, Mark (1996). Clitics and word order since Wackernagel. One hundred years of resarch into clitics and related phenomena. *Orbis*, 36, 389-410.
- Kaye, Jonathan and Jean Lowenstamm (1984). De la syllabicit  . In F. Dell et al. (eds.), *Forme sonore du langage*, 123-159. Paris: Hermann.
- Kayne, R. (1976). French Relative que. In F. Hensey et M. Lujan (eds.), *Current studies in roman linguistics*, Georgetown University Press.
- Kayne, Richard (1972). Subject inversion in French interrogatives. In Casagrande J. and B. Saciuk (eds.), *Generative studies in Romance languages*, Rowley, MA: Newbury house.
- Kayne, Richard (1975). *French syntax. The transformational cycle*. Cambridge MA: MIT Press.
- Kayne, Richard (1984). *Connectedness and binary branching*. Dordrecht: Foris Publications.
- Kayne, Richard (1989). Null subject and clitic climbing. In Jaeggli, Osvaldo and Kenneth Safir, (eds.) *The null subject parameter*, Dordrecht: Kluwer, 239-261.
- Kayne, Richard (1991). Romance clitic, verb movement and PRO. *Linguistic Inquiry*, 22, 647-686.
- Kayne, Richard (1994). *The antisymmetry of syntax*. Cambridge MA: MIT Press.
- Kenstowicz, Michael (1994). *Phonology in generative grammar*. Cambridge MA: Blackwell.
- Keyser, Samuel J. and Thomas Roeper (1984). On the middle and ergative constructions in English. *Linguistic Inquiry*, 15, 381-416.



- Kiparsky, Paul (1982). Lexical phonology and morphology. Yang, I.S. (ed.) *Linguistic in the morning calm*, 3-91. Seoul: Hanshin.
- Kiparsky, Paul (1985). Some consequences of lexical phonology. *Phonology Yearbook*, 2, 83-138.
- Klavans, Judith (1985). The independence of syntax and phonology in cliticization. *Language*, 61, 95-120.
- Koopman, Hilda (1984). *The syntax of verbs*. Dordrecht: Foris.
- Koster, Jan (1984). On binding and control. *Linguistic Inquiry*, 15, 417-459.
- Laenzlinger, Christopher (1993). A syntactic view of Romance pronominal sequences. *Probus*, 5, 241-270.
- Langacker, R.W. (1968). Observations in French Possessives. *Language*, 44, 1.
- Larson, Richard (1988). On the double object construction. *Linguistic Inquiry*, 19, 335-391.
- Leben, William (1973). *Suprasegmental phonology*. Doctoral dissertation, MIT.
- Leclère, Christian (1976). Datifs syntaxiques et datif éthique. In Chevalier, Jean-Claude, and Maurice Gross, (eds.), *Méthodes en grammaire française*. Éditions Klincksieck, Paris.
- Lefebvre, Claire (1982). Qui qui vient? ou qui vient : voilà la question. In C. Lefebvre (ed.), *La syntaxe comparée du français standard et du français populaire : approches formelles et fonctionnelles*, tomes 1-2, Office de la langue française, Montréal.
- Legendre, Geraldine (1996). *Optimal clitic and verb movement in Romanian*. Hand out of talk given at Going Romance, Research Institute for Language and Speech (OTS), Utrecht, DEC. 12-4, 1996.
- Lumsden, John (1987). *Syntactic features. Parametric variation in the history of English*. Doctoral dissertation, MIT.
- Lumsden, John (1992). Underspecification in grammatical and natural gender. *Linguistic Inquiry*, 23, 469-486.
- Lyons, Christopher (1995). Voice, aspect, and arbitrary arguments. In Smith, John Charles, and Martin Maiden (eds.), *Linguistic Theory and the Romance languages*, 78-114. Philadelphia: John Benjamin Publishing Company.
- Manzini, M.Rita (1983). On control and control theory. *Linguistic Inquiry*, 14, 421-446.

- Marantz, Alec (1984). *On the nature of grammatical relations*. Cambridge: MIT press.
- McCarthy, John (1979). *Formal problems in Semitic phonology and morphology*. Doctoral dissertation, MIT.
- McCarthy, John (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, 12, 373-418.
- McCarthy, John (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry*, 17, 207-263.
- Miller, Christopher. (1994). Simultaneous constructions and complex signs in Quebec Sign Language. In Ahlgren, Inger, Brita Bergman and Mary Brennan (eds.), *Perspectives on sign language structure. Papers from the Fifth International Symposium on Sign Language Research*, Salamanca, Spain, 25-30 may 1992, p. 131-147.
- Milner, Jean-Claude. (1982). *Ordres et raisons de langue*. Paris :Seuil.
- Montrul, Silvina (1995). Clitic-doubled dative subject in Spanish. In Zagona, Karen (ed.), *Grammatical theory and Romance languages, LSRL XXV*. Philadelphia: Benjamins Publishing Company.
- Morin, Yves-Charles (1985). On the two French subjectless verbs *voici* and *voilà*. *Language*, 61, 777-820.
- Nespor, Marina and Irene Vogel (1986). *Prosodic phonology*. Dordrecht: Foris Publications.
- Nevis, Joel A., Brian D. Joseph, Dieter Wanner, and Arnold M. Zwicky (1994). *Clitics. A comprehensive bibliography 1892-1991*, Library and information sources in linguistics, 22. Amsterdam: Benjamins.
- Ouhalla, Jamal (1989). Clitic movement and the ECP: evidence from Berber and Romance languages. *Lingua*, 79, 165-215.
- Paradis, Carole, and Darlene Lacharité (1993) (eds.). Constraint-based theory in multilinear phonology. *Canadian Journal of Linguistics*, 38, 2.
- Paradis, Carole, and Jean-François Prunet (1993). A note on velars nasals: the case of Uradhi. *Canadian Journal of Linguistics*, 38, 4, 425-439.
- Paul Hirschbühler and Labelle, Marie (1994). Change in verb position in French negative infinitival clause. *Language Variation and Change*, 6, 149-178.
- Perlmutter, David (1971). *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

- Piggot, Glyne (1992). Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory*, 10, 33-78.
- Pollock, Jean-Yves (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry*, 20, 365-424.
- Postal, Paul (1994). The ungrammaticality of subnominal en in French object-raising complements. *Linguistic Inquiry*, 25, 179-186.
- Prince, Alan and Paul Smolensky (1993). *Optimality theory: Constraint interaction in generative grammar*. Ms. To appear, Cambridge MA: MIT Press.
- Pulleyblank, Douglas (1988). Vocalic underspecification in Yoruba. *Linguistic Inquiry*, 19, 233-270.
- Ralli, Angela (1997). *Inflectional features and the morphology component hypothesis*. 20th GLOW.
- Rivero, Maria Luisa (1994). Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory*, 12, 63-120.
- Rizzi, Luigi (1982). *Issues in Italian syntax*. Dordrecht: Foris.
- Rizzi, Luigi (1986a). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17, 501-557.
- Rizzi, Luigi (1986b). On the status of subject clitics in Romance. In Jaeggli, O. and C. Sylva-Corvalan (eds.), *Studies in Romance linguistics*. Dordrecht: Foris, 391-419.
- Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge MA: MIT Press.
- Roberge Yves and Marie-Thérèse Vinet (1989). *La variation dialectale en grammaire universelle*. Montreal: Presses de l' Université de Montréal.
- Roberts, Ian (1991). Excorporation and minimality. *Linguistic Inquiry*, 22, 209-218.
- Rochette, Anne (1988). *Semantic and syntactic aspect of Romance sentential complementation*. Doctoral dissertation, MIT.
- Roldàn, Mercedes (1974). Constraints on clitic insertion in Spanish. In Campbell, Joe, M.G. Goldin, and M. Clayton Wang (eds.), *Linguistic Studies in Romance Languages: Proceedings of the third linguistic symposium on Romance Language*, 124-138.
- Rooryck, Johan (1992). Romance clitic ordering and universal grammar. *The Linguistic Review*, 9, 219-250.
- Rooryck, Johan (1994). Against optional movement in clitic climbing. In Mazzola, Michael L. (ed.), *Issues in Romance Linguistics: Selected papers from the*

- linguistic symposium on Romance languages XXIII, April 1-4, 1993.* Georgetown University Press, Washington, D.C.
- Rooryck, Johan (1994). On 0- and a- underspecification in syntax and phonology. *MIT Working Papers in Linguistics* 22, 197-216.
- Rouveret, Alain (1993). Cliticisation et temps en portugais européen. *Revue des langues romanes*, 93, 2, 337-371.
- Rudin, Catherine (1988). On multiple questions and multiple WH fronting. *Natural Language and Linguistic Theory*, 6., 445-501.
- Ruwet, Nicolas (1972). *Théorie syntaxique et syntaxe du français*. Paris: Éditions du seuils.
- Sadock, Jerrold (1991). *Autolexical syntax. A theory of parallel grammatical representations*. Chicago: University of Chicago Press.
- Sagey, Elizabeth (1986). *The representation of features and relation in nonlinear phonology*. Doctoral dissertation, MIT.
- Schein, Barry and Donca Steriade, (1986). On geminates. *Linguistic Inquiry*, 17, 691-744.
- Selkirk, Elizabeth (1995). Sentence prosody: intonation, stress and phrasing. In Goldsmith, John (ed.), *The handbook of phonological theory*, 550-569. Cambridge, MA: Blackwell Publishers.
- Shieber, Stuart (1986). *An introduction to unification-based approaches to grammar*. Stanford: CSLI Publications.
- Silva-Villar, Luis (1998). Morphology and syntax of Romance imperatives. *Probus*, 10, 115-138.
- Silva-Villar, Luis (1998). Morphology and syntax of Romance imperatives: an incomplete history. *Probus*, 10, 115-138.
- Speas, Margaret (1997). Optimality Theory and Syntax: Null pronouns and control. In Archangeli, Diana, and D. Terence Langendoen (eds.), *Optimality Theory. An Overview*. Blackwell Publishers, Malden (MASS).
- Speas, Margaret (1997). *Phrase structure in natural language*. Dordrecht: Kluwer Academic Publishers.
- Sportiche, Dominique (1989). Le mouvement syntaxique: contraintes et paramètres. *Langages*, 95, 35-80.
- Sportiche, Dominique (1996). Clitic constructions. In Rooryck, Johan and Laurie Zaring (eds.), *Phrase structure and the lexicon*. Dordrecht: Kluwer Academic Publishers.

- Steele, S. (1995). Toward a theory of morphological information. *Language*, 71.
- Stroik, Thomas (1992). Middles and movement. *Linguistic Inquiry*, 23, 127-137.
- Stroik, Thomas (1999). Middles and Reflexivity. *Linguistic Inquiry*, 30, 119-131.
- Suñer, Margarita (1988). The role of agreement in clitic doubled constructions. *Natural Language and Linguistic Theory*, 6, 391-434.
- Suñer, Margarita (1991). Two properties of clitics in clitic doubled construction. In Huang, James and R. May (eds.), *Logical structure and linguistic structure*. Boston: Kluwer Academic Press.
- Terzi, Arhonto (1996a). Clitic climbing from finite clauses and tense raising. *Probus*, 8, 273-295.
- Terzi, Arhonto (1996b). The Linear Correspondence Axiom and the adjunction site of clitics. In Di Sciullo, Anna-Maria (ed.), *Configurations: essays on structure and interpretation*. Somerville MA: Cascadilla Press.
- Terzi, Arhonto (1999). Clitic combinations, their hosts and their ordering. *Natural Language and Linguistic Theory*, 17, 85-121.
- Thomas, Adolphe V. (1956). *Dictionnaire des difficultés de la langue française*. Larousse, Paris.
- Travis, Lisa (1984). *Parameters and effects of word order variation*. Doctoral dissertation, MIT.
- Tremblay, Mireille (1991). *Possession and Datives: Binary branching from the lexicon to syntax*. Doctoral dissertation, McGill University, Montreal.
- Trubetzkoy, Nicolas (1967/1939). *Principes de phonologie*. Paris: Klincksieck.
- Uriagereka, Juan (1995). Aspects of the syntax of clitic placement in western Romance. *Linguistic Inquiry*, 26, 79-123.
- Vergnaud, Jean-Roger (1974). *French relative clauses*. Doctoral dissertation, MIT.
- Watson, Keith (1997). French complement clitic sequences: a template approach. *French Language Studies*, 7, 69-89.
- Webelhuth, Gert (1995) (ed.). *Government and Binding Theory and the Minimalism Program*. Oxford: Blackwell.
- Wehrli, Eric (1986). On some properties of French clitics *se*. In Borer, Hagit (ed.), *The syntax of pronominal clitics*, 263-283. Orlando: Academic Press.
- Williams, Edwin (1980). Predication. *Linguistic Inquiry*, 11, 203-238.

- Williams, Edwin (1994). A reinterpretation of evidence for verb movement in French. In Lightfoot, David, and Norbert Hornstein (eds.), *Verb movement*, 189-205. Cambridge: Cambridge University Press.
- Williams, Edwin (1995). Theta Theory. In Webelhuth, Gert (ed.), *Government and Binding Theory and the Minimalism Program*, 97-124. Cambridge: Blackwell.
- Yip, Moira, Joan Maling and Ray Jackendoff (1987). Case in tiers. *Language*, 63, 217-250.
- Zanuttini, Raffaella (1997). *Negation and clausal structure: a comparative study of Romance languages*. New York: Oxford University Press.
- Zribi-Hertz, Anne (1982). La construction "se moyen" du français et son statut dans le triangle moyen-passif-réfléchi. *Linguisticae Investigationes*, 6, 345-401.
- Zribi-Hertz, Anne (1987). La réflexivité ergative en français moderne. *Le Français Moderne*, 55, 23-54.
- Zwicky, Arnold (1977). *On clitics*. Bloomington: Indiana University Linguistics Club.
- Zwicky, Arnold and Geoffrey Pullum (1983). Clitization vs. inflection: English N'T. *Language*, 59, 502-513.
- Zwicky, Arnold M. (1994). What is a clitic? In Nevis et al. *Clitics. A comprehensive bibliography 1892-1991*, Library and information sources in linguistics, 22. Amsterdam: Benjamins.