

The clitics mechanism in French and Italian

PIETER A. M. SEUREN

Abstract

The article concentrates on the question of the COMPOSITION, the internal ORDERING and the PLACEMENT of clitic-clusters (C-clusters) in French and Italian, though clitic data from other languages are drawn in occasionally. The system proposed is top-down transformational, in the terms of Semantic Syntax (Seuren 1996). Clitics are taken to originate in underlying structure as canonical argument terms or adverbial constituents of clauses. During the process of transformation from semantic to surface form, nonfocus, nonsubject, pronominal argument terms are assigned values for the features of animacy ([±AN]), dative status ([±DAT]) and reflexivity ([±REFL]). On the basis of these, the rule feature ⟨CM⟩, inducing CLITIC MOVEMENT, is assigned or withheld. Plus-values increase, and minus-values reduce, the “semantic weight” of the clitics in question. Pronouns without the feature ⟨CM⟩ are not cliticised and stay in their canonical term position in their full phonological form. Pronouns with the feature ⟨CM⟩ are attached to the nearest verb form giving rise to clitic clusters, which accounts for the COMPOSITION of well-formed C-clusters. The attachment of clitics to a cluster occurs in a fixed order, which accounts for the ORDERING of clitics in well-formed clusters. BRANCHING DIRECTIONALITY, together with a theory of complementation, accounts for the PLACEMENT of C-clusters. Clitics often take on a reduced phonological form. It is argued that, in French and Italian, which are languages with a right-branching syntax and a left-branching flectional morphology, postverbal clitics, or enclitics, are part of left-branching structures and hence fit naturally into the morphology. They are best categorised as affixes. Occasionally, as in Italian glielo, dative clitics (e.g., gli) turn preceding lighter clitics (e.g., lo) into affixes, resulting in the left-branching structure glielo, where -lo is an affix. In a brief Intermezzo, instances are shown of the irregular but revealing lui-le-lui phenomenon in French, and its much less frequent analog in Italian. On these assumptions, supported by

the official orthographies, the clitic systems of French and Italian largely coincide. This new analysis of the facts in question invites further reflection on the interface between syntax and morphology. The final section deals with reflexive clitics. There, the system begins to be unable to account for the observed facts. At this end, therefore, the system is allowed to remain fraying, till further research brings greater clarity.

1. Introduction¹

Clitics are problematic for the theory of language in many different ways. To begin with, it has so far proved impossible to give a precise definition of what clitics actually are. In practice, linguists are guided by a number of criteria which are neither universally applicable nor universally recognised. Arnold Zwicky writes (1994: 571):

In the broadest sense, clitics are elements that have some characteristics of independent words and some of affixes within words. There are several ways in which an element can be word-like or affix-like – in its phonological, morphological, syntactic, or semantic properties, in particular – so that the label “clitic” covers a diverse collection of phenomena, which are unlikely to constitute a unified class for the purposes of theorizing about the nature of grammar. But the fact that such elements are very widespread, occurring apparently in all languages, together with the fact that the description of any type of clitic involves simultaneous reference to two or more components of a grammar, makes them central to linguistic theory.

We know that a clitic cannot occur as a free form, in the sense of Bloomfield (1933: 160, 178). We also know that a clitic is either unstressed *per se* or amalgamates with the support word or constituent (often called “host”) into one phonological accent unit, in which case word accent may fall on a clitic or an extra accent may be assigned.² (The word *clitic* derives from the Greek

1. I am indebted to Brigitte Bauer of the University of Texas at Austin for useful critique and suggestions. I must also thank my Italian and French informants, who patiently submitted to my questioning. The comments made by two anonymous reviewers have substantially improved the paper.

2. This appears to occur only with enclitics, never with proclitics. A much-quoted example (e.g., Anderson 2005: 43) is the Modern Greek imperative *dhóse móu to* ‘give-me-it’, where the Greek accent rules assign an extra accent to *móu* ‘me’. Likewise in, for example, *ta dhiavatíriá sas* ‘your passports’, where the noun *dhiavatíria* ‘passports’ gets an extra final accent owing to the following enclitic *sas* ‘your-pl’, which, in this case, remains unaccented. See also Miller and Monachesi (2003: 99) for a similar example from the dialect of Naples. If our argument, presented in section 8, that enclitics are closer to being suffixes than to being clitics is correct, then the generalisation holds that clitics never receive accent, though suffixes may do so as a result of word phonology.

klitikón, literally ‘leaning’.) No clitic can be under sentential focus accent. A clitic may precede the support, in which case it is a PROCLITIC; when it follows it is an ENCLITIC.

Clitics often occur in clusters (C-CLUSTERS), which occupy a specific position in the sentence and are subject to strict conditions as regards both the combinability and the ordering of the clitics occurring in them. In the traditional literature, the combinability and ordering conditions are mostly treated in a purely taxonomic manner, which lacks explanatory force and fails to link up the grammar of clitics with other parts of the grammar. Perlmutter (1971) was a first attempt at providing a more theory-based account of clitic behaviour, but it remained stuck in taxonomy. After that, many attempts have been made to fit clitics into existing theories, but, on the whole, with only limited success. It is fair to say that, compared with other, more successful, areas of grammatical research, the situation regarding the analysis and description of clitics and C-clusters is disappointing.

No attempt is made in the present article to define the notion of clitic beyond the more or less fluid criteria that have so far served linguists to identify them. Nor will I enter into the question of the phonological form of clitics as opposed to their counterparts occurring as full canonical argument terms. I will, instead, focus on the composition, internal ordering, and placement conditions of clitic clusters, and that only in standard French and standard Italian. It is taken for granted that the elements dealt with in this study are indeed for the most part clitics, and that most of them are “words” in the ill-defined sense still current in the profession. The syntactic category “clitic” (label “CL”) is thus provisionally defined by what it does for its members in the machinery of grammar. I will, however, support the hypothesis, often presented in the literature (e.g., Miller and Monachesi 2003: 97), that clitics represent a transitional stage in a process that leads from full word status to the status of morphological (fleclional) affix.

A major distinction must be made first. When one takes a world-wide view, it would seem that those expressions that one would wish to call clitics occur in two main varieties. First, there are the clitics that gravitate towards what is known as the WACKERNAGEL POSITION. This position, however, is ill-defined. Wackernagel defined the position as second after the first word or constituent in the sentence, but Fraenkel (1932/1933) redefined it as second in any “colon” – a notion described as a semantico-intonational unit (see Adams 1994). And modern phonological or morphological theory has not brought greater clarity either. What does appear is that, in some cases, it is the position after the first word of a clause or sentence, but sometimes it is not the first word but a larger unit that defines the position in question. Latin *-que* (‘and’) and its epic Greek etymological counterpart *te* (also ‘and’) (the former being standardly treated as an affix and the latter as a word, with or without good reason), as well as the Greek particles *mén* ‘on the one hand’ and *de* ‘but, on the other hand’ are

examples of particles that must occur right after the first word, not the first “colon,” of the second conjunct. By contrast, Latin *autem* ‘however’ and *igitur* ‘therefore’, and English *however*, can be reckoned to belong to the second class of Wackernagel clitics: for these it is not the first word but some larger unit, possibly Adams’s “colon,” that defines the position – though *however* may also, though not preferably, occur in first position, followed by a pause, just like Latin *ergo* ‘therefore’. Yet, despite such unclarities, the phenomenon is clearly real. Typical Wackernagel clitics, of either variety, are found in virtually all Australian Aboriginal languages (Dixon 1980: 362) and in, for example, Tagalog (Schachter and Otnes 1972: 183–193). Some Tagalog examples are given in (1) (clitics are printed in small capitals):

- (1) a. *Hindi* SIYA *masaya ngayon*.
not he happy today
‘He is not happy today.’
b. *Bakit* SIYA *hindi masaya ngayon?*
why he not happy today
‘Why isn’t he happy today?’
c. *Tatlo* LAMANG *sila*.
three only they
‘There are only three of them.’

In the light of what is known about Wackernagel clitics, it is not clear whether these are subject to combinatorial or ordering constraints, as is common for the second category of clitics, to which we turn now.

Clitics of the second category, V-clitics, gravitate towards verbs, to which they are either proclitic or enclitic.³ They are normally pronominal and sometimes also adverbial (such as the clitic form *ne* for the negation in French). Examples are the clitics of the Romance languages, Modern Greek, Serbo-Croat, and perhaps also, to some extent, the Germanic languages. In English, for example, cliticisation may be responsible for the fact that, while the internal indirect object normally precedes the direct object, as in (2a), the order is inverted when both objects are pronominal, as in (2b):

- (2) a. *I gave Rose the book*.
b. *I gave IT HER*.

One is inclined to surmise that the light pronominal form *it* is attracted by the verb form owing to a form of cliticisation, and that the heavier pronoun *her* is

3. Exceptions are the expressions for ‘here it is’ or ‘there you are’ in French and Italian. In French one has *le voici/voilà*, where the forms *voici* and *voilà* go back to the old imperative *vois* ‘see’. Italian has *eccolo*, composed of *ecco* ‘there you are’ and the clitic or affix *-lo* ‘it, him’.

either not a clitic or less strongly attracted. The same is found in Dutch and German. German, however, looks as if it also has some form of Wackernagel clitics, often, but not always, in dependent clauses (if complementizers may count as clause-initial support):

- (3) a. ... *daß* ES IHM *niemand hat verzeihen wollen*
... that it him no-one has forgive want
'... that no-one has wanted to forgive it him'
b. *Hat* DICH *der Mann geschlagen?*
has you the man beaten
'Has the man beaten you?'

In the present article, however, we are concerned only with clusters of V-clitics in standard French and standard Italian, though material from other languages is also discussed occasionally.

The problem of how to account for the composition, internal ordering, and positioning, of clustered V-clitics – henceforth the CLITIC CLUSTER PROBLEM – has remained unsolved in all existing theories of grammar. Perlmutter (1971) already observed that grammatical theory of his day – that is, transformational grammar – was unable to provide an answer to the clitic cluster problem for the Romance languages, an observation repeated in Seuren (1976). In the meantime, many other grammatical theories have sprung up and transformational grammar itself has undergone many changes and diversifications. Yet since the mid-1970s, the interest of syntacticians has focused not so much on a solution of the clitic cluster problem as on other semantic and structural properties of clitics, such as their phonological and intonational features. For Legendre (2000: 426) the fact that no syntactic solution to the clitic cluster problem has so far been presented is sufficient reason to conclude that cliticisation phenomena cannot be of a syntactic nature and must, therefore, be treated in the phonology and/or morphology. Referring to the apparently whimsical behaviour of clitics – that is, the clitic cluster problem – she writes (Legendre 2000: 434):

No syntactic explanation has ever been proposed for this idiosyncrasy [sic] because there is none: irregularities of this sort are a hallmark of morphology, not syntax.

This statement is remarkable not only because calling facts for which no rule-based system has so far been found “irregularities” is an amusing *non sequitur*, but also because the author subsequently omits explaining how clitic behaviour can be handled in morphological terms. Worse, she fails to show clitic behaviour.

Syntax, however, cannot be missed in a proper account of clitic behaviour. Clitic behaviour clearly follows syntactic structure, as appears, for example,

from what is still standardly regarded as the CLIMBING of clitics – a phenomenon exhaustively explained by independently motivated syntactic rules (see the final part of Section 4), or from the fact that nonfocus nonsubject pronouns are sometimes not cliticised and are then treated as canonical, syntactically defined, argument terms – a phenomenon discussed in Section 3. Following the trend of the day (see Note 10), Legendre does pay some attention to PRONOUN RESUMPTION, but without any mention of the fact that pronominal resumption, whether optional or obligatory, is a widespread phenomenon in the languages of the world, including languages that do not cliticise their pronouns. But for this author, pronoun resumption simply counts as “clitic doubling.”

Legendre (2000) is only one example of a large number of more or less recent publications that deal with clitics one way or another but fail to look at the facts that are relevant to the clitic cluster problem. Another striking example is the recent study by Anderson (Anderson 2005), a book of 287 pages of text, which is largely devoted to clitics but makes no mention at all of the fact that there is a problem regarding clitic clustering. The book does devote some attention to the question of the placement of clitics (clitic clusters) in sentential structures, but even in this respect no precise account is presented that would have any degree of predictivity. The same holds, one way or another, for all publications on clitics that have appeared since the mid-1970s. Haverkort (1993), for example, is an attempt at making cliticisation phenomena fit into Chomsky’s Principles and Parameters framework without having to build in an extension of the theory, but it fails to look at such phenomena in any detail, which makes the attempt less convincing, to say the least. Belletti (1999) does the same with regard to Chomsky’s Minimalist Program. Many more such studies have appeared between 1970 and now, aiming to shore up a particular variety of syntactic theory in a general sense but failing to account for the, often complex, phenomena themselves. (Moreover, the relevance of such attempts co-varies with the length of time the theories or frameworks in question stay alive.) Other studies only look at one particular subproblem, leaving the wider perspective of the clitics cluster problem out of account. Cardinaletti and Shlonsky (2004), for example, though going under a rather general title, in fact looks in detail at the various conditions on what the authors still consider to be clitic climbing in Italian. Van Riemsdijk (1999), a survey of clitics in the languages of Europe, contains many interesting studies by many competent authors, some of whom mention the clitic cluster problem but all of whom fail to come up with a solution to this specific problem. In Miller and Monachesi (2003), a seminal study of Romance clitics, the authors admit defeat as regards the clitic cluster problem and fall back on a taxonomic template representation, while stressing that such an account is inadequate.

Cardinaletti (1999) is likewise frank about the absence of a solution to the clitic cluster problem. In a survey article on Germanic and Romance clitics,

under the heading “Open questions”, she writes (1999: 68–69):

The second topic which calls for a detailed account is the analysis of sequences of pronouns and, more specifically, the internal structure of clitic clusters. Two issues arise in this respect. First, whether the order of arguments is reflected in the order of pronouns or not ... The second issue concerns possible and impossible clusters, for which Case and person features seem to play an important role ... These still highly mysterious restrictions have been partly discussed in the literature devoted to Romance ... Apart from the few observations in § 5.1, above, no discussion of clitic clusters exists for Germanic languages.

Cardinaletti’s suggestion that “Case and person features seem to play an important role” is interesting in that it is claimed in the present article that the cliticisation machinery is driven by the (appropriately defined) features of animacy ([±AN]), dative status ([±DAT]) and reflexivity ([±REFL]).

Hopefully, the above references and quotations suffice to prove my case that there has hardly been any progress since the mid-1970s as regards the specific problem here called the clitic cluster problem. The conclusion in Perlmutter (1971) and Seuren (1976) still stands: an extension of the grammatical machinery in any of the existing theories is inevitable if grammars and linguistic theory are to account for the facts of clitic clusters. It is the purpose of the present study to contribute to, and help prepare the ground for, a general, universally valid, truly explanatory theory of clitics. It must be admitted, though, that we are still far removed from that ideal. All we can do here is look at a few cliticising languages, in particular French and Italian, in minute detail and try to gain a precise idea of their clitic systems, hoping that, when the same is done for other languages, certain generalisations will emerge. One result of this paper is anyway that strong generalisations have come to light for the two languages concerned, whose clitic system is shown to be much more uniform than surface phenomena suggest. But at this stage, there is no claim as regards clitic systems in other languages than French and Italian, despite the occasional foray into Portuguese or Modern Greek. The justification for this method of proceeding is that premature generalisations tend to lull researchers into believing that the code has been cracked and that no further research is needed. The justification for the selection of French and Italian is that the clitic systems of these languages, especially Italian, are so rich that one may expect much or most of what constitutes cliticisation in universal grammar to turn up at an observable surface level in these languages.

It is not the purpose of this study to answer questions relating to the historical or functional factors that may have led to the phenomena of cliticisation and clitic clustering. Such questions are of great interest, but they are not our direct concern here, though it is not unlikely that the system proposed here may in the end, contribute to a better overall understanding of the functionality, and

perhaps even the history, of cliticisation phenomena. For the moment, we accept the facts as they are, and merely try to capture them in a transformational (but non-Chomskyan) mechanism, the main concern being (a) to get the facts right, and (b) to do so in terms of the simplest and most general possible machinery.⁴ This new system falls back on Seuren (1976), reprinted as Chapter 11 in Seuren (2001) and improved in Seuren (1996) and Seuren (2002). The main advantage of the revised treatment presented here over its older versions is that it has proved possible to treat clitic clustering in modern standard French and Italian as being constrained by the grammaticalized semantic features of animacy [\pm AN], dative case [\pm DAT] and reflexivity [\pm REFL], which gives a semantic motivation to the system. Although this study is essentially restricted to French and Italian, the stabs taken at a few other languages suggest striking similarities with the Franco-Italian system.

Skeptics might say that grammatical theories do not have to worry too much about C-clusters, as such clusters are finite in number and can be specified by simple enumeration in the form of “templates,” which was also essentially the method employed in Perlmutter (1971). Such template-type accounts are, in principle, no better than the standard taxonomic treatment of C-clusters in traditional grammars. They are unsatisfactory, not only because they tend to be observationally inadequate, especially as regards Italian, with its reflexive clitics that cut across templates (see Section 9; Seuren 2001: 273–276), but also because cliticisation is clearly structure-dependent and therefore requires an analysis that incorporates it into a system of structure-driven rules and principles.

4. Many modern authors reject the canonical forms of transformational analysis and description developed during the heyday of transformational grammar. Anderson, for example, writes (Anderson 2005: 83):

The notion of a set of special syntactic rules of the sort implied by this typology is clearly a relic of the more general conception of syntax as grounded in a collection of construction-specific rules (“Passive,” “Dative Movement,” “Subject-to-Object Raising,” etc.). Most modern theories of syntax deplore such rules and attempt to derive the properties of individual constructions from more general properties of structure and displacement operations.

This is mere rhetoric and, one will agree, a caricature of those grammatical theories that work with the vintage rules of transformational grammar like PASSIVE, DATIVE EXTRACTION, SUBJECT RAISING, and the like, since if these theories have had any central concern since the 1960s, it has been precisely the urge to “derive the properties of individual constructions from more general properties of structure,” though, of course, with full respect of the facts, unlike the run of more recent theories. Generalising while neglecting the precise facts is a vacuous exercise leading to nothing. It may be true that “most modern theories of syntax deplore such rules”, but it would help if Anderson, or any other author with universalist aspirations, presented a viable alternative that could boast the same degree of factual correctness, precision, predictivity, and generality while needing less machinery. Needless to say, one invariably waits in vain for such help. In linguistics it is obviously false to think that new always means better.

In discussing grammatical analysis and description one can follow two opposed methods, the bottom-up or parsing method and the top-down or generative method. The parsing method attempts to develop an algorithmic procedure that starts from the surface structures of a language and provides a semantic interpretation for them by reducing them to expressions in a semantically analytical language, which, in the best cases, is a variety of the established formal language of predicate calculus. The product of this reduction is called LOGICAL FORM or SEMANTIC ANALYSIS, and is considered to be open to direct semantic interpretation either in terms of some variety of logical model theory, as in model-theoretic formal semantics, or in terms of a more psychologically oriented form of semantic interpretation. By contrast, the top-down method of analysis and description attempts to develop an algorithmic procedure that specifies, for any given “deep structure” input, which strings of symbols are admissible and which are not. This latter method is usually called “generative,” for historical reasons. The bottom-up parsing method is not followed here, if only because the interpretation of given C-clusters in terms of corresponding nonclitic structures is, on the whole, unproblematic (see Seuren 2004a: 83–85 for more general arguments). The real problem rests with the top-down or generative specification of the correct C-clusters in cliticising languages. The problem will, therefore, be dealt with in the generative or top-down fashion.

As is well-known, however, the generative method is far from uniform. According to some (including this author), the structures that are fed into the generative algorithmic rules are semantically analytic forms, SEMANTIC ANALYSES or SAs, which originate in cognition as propositional “thoughts” under speech act operators. All the generative grammar does, in this view, is mechanically transform semantic analyses to corresponding surface structures (Seuren 1996, 2004a). In this MEDIATIONAL view of grammar, a grammar “mediates” between thought and sound. According to the Chomskyans, however, grammars generate sentences arbitrarily, assigning them a semantic interpretation once they have been fully or partly generated. This is the RANDOM-GENERATOR view of grammar, which, in the present author’s view, is unrealistic from a psychological point of view and therefore lacks much of the explanatory power a theory of grammar is reasonably expected to have. Yet this theoretical difference need not detain us here. I agree with the vast majority of analyses in any of the varieties of generative transformational grammar that have seen the light so far that, in principle at least, clitics should be seen as transformationally derived from underlying forms where the clitics are not clitics yet, but standard arguments to verbs (or adverbial adjuncts).⁵ This simple and reasonable assumption suffices

5. This does not apply, of course, to lexically encoded clitics, as in French *s’en aller* ‘go away’. Such cases are considered to be parasitic on the strictly syntactico-morphological constructions.

to bridge the gap between mediational and random generative grammar in the present context.

No position need be taken here on the issue of INNATE restrictions versus performance-driven FUNCTIONALITY of grammatical rules and structures (Newmeyer 1998; Hawkins 2004). In the present author's view, the two conceptions of grammar are not mutually exclusive, as functional features may become inbred. In fact, functional motivation will count in favour of any theory or hypothesis regarding innate universal restrictions on grammars, even if one often comes across rules that are patently counterfunctional yet present in a great many languages and frequently used in them.⁶ The main focus of the present study is the question of how the clitics machinery, whether functionality-driven or restricted by innate principles and parameters or both, actually works for certain languages. An adequate and detailed account of that machinery will further the debate on the nature and functionality of the human language capacity in general, whether viewed from a competence or a performance angle. Meanwhile, one has no choice but to search for the widest possible intralinguistic and cross-linguistic generalisations.

What this paper is not, and cannot be, neutral on is the MODULARITY THESIS for grammars – the thesis that at least top-down grammatical processing takes place via an input-output device, a module, which is near-deterministic, automatic, rapid, insulated from external interference and closed to introspection (Fodor 1983; Seuren 2004b). (Whether bottom-up processing, or parsing, is also modular in the sense indicated, or whether it underdetermines output and needs additional information from cognition, is a moot question, though the latter seems more likely.) There is no way the grammatical mechanisms and processes proposed here can be taken to be subject to the constraints (and operating costs) of what is consciously accessible and controllable in the human mind. Whoever wishes to uphold the thesis that the mechanisms and processes at play in grammar are subject to conscious control seriously underestimates the complexity of the object and the explanatory power of the, inevitably technical, solutions proposed. No modularity claim is made regarding the semantic processing and contextual integration of utterances, whether active or passive, though some of these processes, such as lexical access, are no doubt largely automatic – and largely mysterious.

6. One thinks, for example, of the rule of PREDICATE RAISING, discussed in Section 4, which makes for semantic opaqueness owing to the amalgamation of verb forms on the one hand and of argument terms on the other, yet is present in a vast number of languages and belongs to the most commonly used registers in those languages.

2. A taxonomic survey of the facts in French and Italian (with a little Greek)

Before the data regarding French and Italian nonfocus oblique case pronominal clitics are discussed, some notational conventions for the relevant parameters of cliticisable pronouns must be agreed upon. The following notation is used:

‘**A**’ stands for ACCUSATIVE ‘**1**’ stands for FIRST PERSON
 ‘**D**’ stands for DATIVE ‘**2**’ stands for SECOND PERSON
 ‘**R**’ stands for REFLEXIVE ‘**3**’ stands for THIRD PERSON

As regards the POSITION of C-clusters in the languages concerned, they are proclitic to verb forms, except for positive imperatives and positive first-person plural (1-pl) adhortatives (for example *allons-nous-en*, ‘let us go away’), which have enclitic clusters. In Italian, moreover, C-clusters attached to infinitives, participles and gerunds are also obligatorily enclitic. The same goes for third person reflexives in certain idiomatic expressions such as *vendesi* ‘for sale’, literally ‘sells itself’. (Cases like French *le voici* or Italian *eccolo*, discussed in Note 3, are not taken into consideration.) As for the INTERNAL STRUCTURE of C-clusters, the following overall taxonomic generalities hold:

- (i) No two clitics with identical phonological form, case or semantic function are allowed in one cluster. (In Italian, *si si* dissimilates into *ci si*, probably because this combination is frequently called for, owing to the multiple functions of reflexives; see Section 9).
- (ii) **A**-pronouns are always cliticised (except, sometimes, when reflexives are involved). **A3**-clitics can be combined with all other oblique case pronominal clitics (which must be **D**-clitics). **A1/A2/A3^R**-clitics can be combined with no other oblique case pronominal clitic, except in certain southern Italian dialects, which pose fewer restrictions on the combinability of pronominal clitics.⁷ Pronouns prevented from cliticisation are treated as full lexical arguments.

7. Similar phenomena have been observed for some French speakers, who accept clusters which their understanding of grammar tells them should be permissible. Miller and Monachesi (2003) gives (i) as acceptable to “une partie des locuteurs”. Other speakers, however, will prefer the standard (ii):

(i) *Pierre me te semble fidèle.*
 ‘Pierre seems to me to be faithful to you.’

(ii) *Pierre me semble fidèle à toi.*

Tasmowski (1985) quotes (iii) as fully acceptable, which it may be to some academically trained speakers who press for what they see as consistency in their grammar:

(iii) *Pierre me la lui a fait raconter.*

- (iii) **D** precedes **A**. This holds generally for both French and Italian (as it does for many other cliticising languages, such as Modern Greek and Serbo-Croat). In French, however, the order **A3–D3** is required preverbally and **A** always precedes **D** postverbally. In postverbal final position, the French weak clitic forms *me* and *te* occur in their strong forms *moi* and *toi*, respectively, owing to their carrying phonological accent. In Italian (and Modern Greek), enclitic clusters are identical to proclitic clusters.
- (iv) The French adverbial negation clitic *ne* combines with all clusters and precedes all other clitics. *Ne* tends not to cliticise in clusters attached to infinitives, though cliticisation does occur (see Seuren 1996: 203–206 for a syntactic account).⁸ Italian (and Modern Greek) have no clitic expression for negation.
- (v) The French adverbial clitics *y* ('there') and *en* ('thereof') follow all other clitics, in that order. The Italian counterpart of French *en*, the clitic *ne* ('thereof/-from') follows all other clitics except **A3** clitics, which it precedes (see Note 19). Italian adverbial *ci* is often used colloquially for dative *gli* or *le*, and even the plural dative *loro*. (Modern Greek has no adverbial clitics.)
- (vi) Italian preverbal clusters **D3–A3** or **D3–ne** are written as one single word. All Italian postverbal C-clusters are represented in the standard orthography as forming one single word with their verbal support. The Italian plural third person pronoun *loro* ('they/them') normally does not cliticise, though it allows for uses that suggest an incipient cliticisation.⁹

As it stands, (iii) seems ambiguous between the readings (a) and (b):

- (a) 'Pierre made me tell it to him/her.'
 (b) 'Pierre made him/her tell it to me.'

Less sophisticated speakers will prefer (iv) for reading (a), and (v) for reading (b):

- (iv) *Pierre la lui a fait raconter par moi.*
 'Pierre had it told to him/her by me.'
 (v) *Pierre me l'a fait raconter par lui/elle.*
 'Pierre had it told to me by him/her.'

Or indeed they will seek a different locution, as the accumulation of pronouns feels uncomfortable.

8. French has, besides the canonical *pour ne pas le voir* 'so as not to see it', also the less preferred *pour ne le voir pas*, with cliticised *ne*. One reviewer pointed out that cliticisation of *ne* is fully acceptable with gerunds: *en ne le voyant pas* ('while not seeing it') seems perfectly acceptable. This was confirmed by my native informants, who considered *en ne pas le voyant* clearly ungrammatical.

9. Cardinaletti and Starke (1999: 149, 166) and Egerland (2005) point out that *loro* can be used as in (i) and (ii), that is, without DATIVE EXTRACTION (see Section 3):

- (vii) Reflexive pronouns sometimes appear to behave in a deviant manner. This complication is discussed in Section 9.

3. Some top-down Semantic Syntax for French

It follows that in French only those preverbal C-clusters are allowed that conform to the pattern shown in (i):

- (i) Pattern for preverbal clitics in French:
ne + 1^(R)/2^(R)/3^R + A3 + D3 + y + en
 (ne + me/nous/te/vous/se + le/la/les + lui/leur + y + en)

Postverbally, C-clusters follow pattern (ii) (**3^R** is impossible with imperatives, which are second person by definition, and with 1-pl adhortatives):

- (ii) Pattern for postverbal clitics in French:
A3 + 1^(R)/2^(R)/D3 + y + en
 (le/la/les + moi/nous/toi/vous/lui/leur + y + en)

Before cliticisation, nonfocus pronouns are taken to occupy the canonical positions for argument terms.¹⁰ Not only does this assumption have a generalising and simplifying effect on the grammar, it also provides an immediate

-
- (i) *Non metterò mai loro il cappuccio.*
 not will-put ever them the cap
 'I will never put the cap on them.'
- (ii) *Non dirò mai loro tutto.*
 not will-say ever them all
 'I will never say everything to them.'

Yet (i) and (ii) co-exist with *Non metterò mai il cappuccio a loro* and *Non dirò mai tutto a loro*, where DATIVE EXTRACTION has taken place in the normal way. This use of *loro* instantiates what has been termed "weak cliticisation" (Holmberg 1991). According to Egerland (2005: 1112), it "is considered rather archaic in the modern language, or, at best, is limited to the written register. It is reminiscent of thirteenth and fourteenth century Italian." The form *lo*, a truncated form of *loro*, occurs in some central varieties of Medieval Italian, especially in texts originating from Siena, as a normal preverbal clitic with the same properties as the singular **D3** pronouns *gli/le* (Egerland 2005). This use, however, did not spread and became obsolete. Apparently, the dodging of DATIVE EXTRACTION is an initial sign of cliticisation. One reviewer pointed out that *loro*, as a weak clitic, occurs immediately after at least one verb form, so that one has, next to the canonical *Doveva aver parlato a loro* (with no cliticisation for *loro*), meaning '(s)he must have spoken to them', all of the following: *Doveva loro aver parlato*, *Doveva aver loro parlato* and *Doveva aver parlato loro*.

10. Various authors speak of CLITIC DOUBLING (e.g., Kayne 2001; Cardinaletti 1999: 42–43; Emonds 1999: 352–354), in cases such as:

- (i) *Jean me connaît moi.*
 Jean me knows me

rationale for the fact that pronouns that do not get cliticised stay behind in their canonical argument term position, as shown in (4):

- (4) *Je vous présenterai à lui.*
 I you will-introduce to him
 'I will introduce you to him.'
- (5) a. **Je vous lui présenterai.*
 b. **Je lui vous présenterai.*
 c. **Je lui présenterai vous.*

In the analysis presented below, (5a–c) turn out ungrammatical as a result of the procedure for the assignment of the rule feature CLITIC MOVEMENT (CM) specified in Figure 1. This assignment procedure entails that once an accusative animate pronoun, such as *vous* in (4), has been assigned the rule feature CM, no dative pronoun can be assigned that feature, so that the dative pronoun will stay behind in its canonical position. Therefore, only (4) can be reckoned to be grammatically well-formed and (5a–c) are unwellformed.

Datives are taken to be internal datives in SA (Seuren 1996: 59–60, 157–218). Dative argument terms are either extracted – that is, turned into an external dative under the dative preposition *à* ('to') – or cliticised.¹¹ The structures (6a) and (6b), corresponding to (7a) and (7b), respectively, occur at that level of representation at which all V-induced cyclic rules under S₂ have been implemented and only structure-induced cyclic rules, such as DATIVE EXTRACTION,

-
- (ii) *La conosco quella.*
 her I-know her

This seems to me to be mistaken. There is, most probably, some form of clitic doubling in the *lui-le-lui* cases cited in Section 7, but not in cases like (i) or (ii), which simply manifest a form of resumptive right dislocation. This appears not only from the falling intonation, preferably after a pause, on the final NP, pronominal or not, but also from the fact that full lexical NPs are freely used in that final position. Some authors try to dissociate right dislocation, which makes no difference between a full lexical NP and an accented pronoun in final position, from what they see as clitic doubling. Emonds (1999: 353), for example, rightly assigns ungrammaticality to **Marie l'a suggéré le livre à Paul* 'Marie has suggested the book to Paul' but fails to mention that *Marie l'a suggéré à Paul le livre* is perfectly grammatical, with falling tone on *le livre*. Kayne (2001: 191) declares *Jean la connaît Marie* 'Jean knows Marie' ungrammatical as opposed to *Jean la connaît elle* 'Jean knows her', apparently wanting to secure clitic doubling as a process distinct from right dislocation. But, as Kayne himself admits, the former sentence is all right as an instance of right dislocation. In fact, both sentences are all right, but only with a falling tone on the final NP. The implicit claim that *Jean la connaît elle* is acceptable without the falling tone on *elle* is flatly denied by my informants.

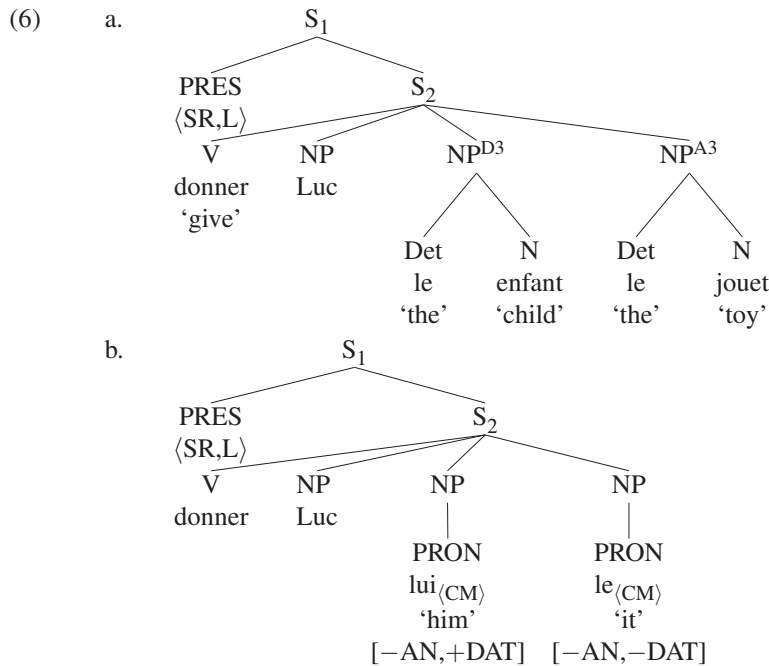
11. In Seuren (1996: 59–60, 99–101) DATIVE EXTRACTION is generalised to TERM EXTRACTION, along with the extraction of the subject term and the formation of a, mostly optional, passive agent phrase in passive sentences.

may still have to be carried out. Let this be called the S-COMPLETENESS (S-COMPL) of the S in question. During the Cycle, S-COMPL moves up, whereby the argument structure of the S in question may change by the addition or deletion of elements. DATIVE EXTRACTION applies during the Cycle at the earliest S-COMPL. It applies to all dative NPs without the feature $\langle CM \rangle$ and is, in a simplified form, described as follows:

DATIVE EXTRACTION:

NP^D is detached from its position and placed under the preposition \grave{a} , after which the new preposition phrase is re-attached to the right under S.

DATIVE EXTRACTION ensures that the internal dative of (6a) is turned into the external dative of (7a). In (6b) the dative pronoun has been assigned the rule feature $\langle CM \rangle$ for CLITIC MOVEMENT, which prevents DATIVE EXTRACTION.



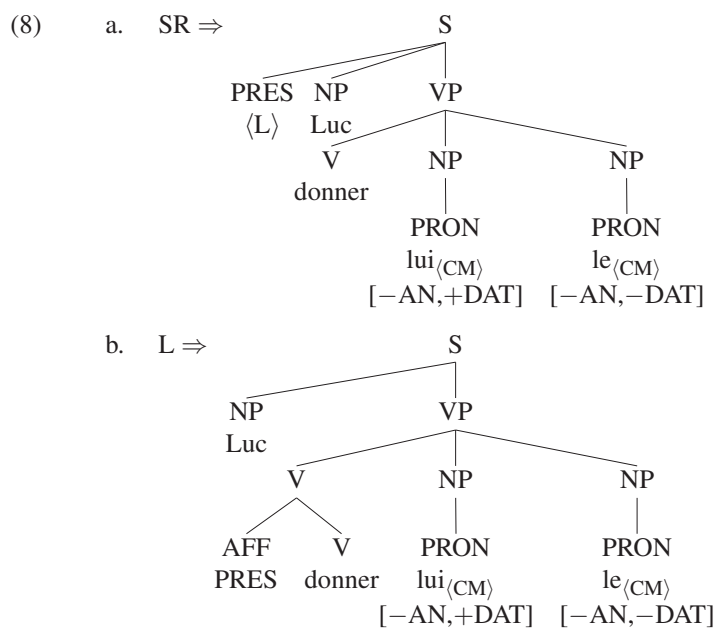
- (7) a. *Luc donne le jouet à l'enfant.*
 'Luc gives the toy to the child/Luc gives the child the toy.'
 b. *Luc le lui donne.*
 'Luc gives it (to) him/her.'

It is assumed throughout (in accordance with McCawley 1970) that the input structures to the grammar (semantic analyses or SAs) are of the form V–S–IO–DO (verb–subject–indirect object–direct object), usually shortened to V–S–O. The TENSE ROUTINE turns the V–S–O-structure of (6a,b) into the NP–VP structure canonical for western European and many other languages. The TENSE ROUTINE is part of the AUXILIARY SYSTEM of the grammar and is induced by the tense predicates PRES or PAST, which carry the rule features SUBJECT RAISING (SR) and LOWERING (L):

SUBJECT RAISING (SR):

The subject-NP of an embedded S is detached and placed in the position of its own immediately dominating S, which is moved one position to the right and is reduced to VP-status in virtue of the universal CONVENTION 1: An S whose subject term is removed is relabelled VP.

LOWERING, in this case, consists in ADOPTION (see below) by the lower verb *donner*. In the process, the predicate PRES (as also PAST) is relabelled AFF(ix) (see Seuren 1996: Ch. 2). (6b) thus becomes (8a) by SR, then (8b) by L. (8b) is the so-called SHALLOW STRUCTURE: it marks the end of the Cycle and the beginning of the Postcycle.



FEATURE ASSIGNMENT, required for cliticisation, takes place at the earliest S-COMPL of any S – in the case at hand, at stage (6b). It applies to nonsubject,

nonfocus oblique pronouns at every S-COMPL, preparing the structure for post-cyclic CLITIC MOVEMENT. The pronouns in question are marked for the features $[\pm\text{AN}]$, $[\pm\text{DAT}]$ and $[\pm\text{REFL}]$:

- $1^{(R)}/2^{(R)}/3^R \rightarrow [+AN]$. (All reflexive pronouns are marked $[+AN]$.)
 Dative pronouns $\rightarrow [+DAT]$
 Reflexive pronouns $\rightarrow [+REFL]$.

Pronouns not marked $[+F]$ for some feature F are $[-F]$. Accusative pronouns are marked $[-DAT]$. One notes that nonreflexive **A3** and **D3** pronouns are marked $[-AN]$, even if they may refer to an animate being.

There is, moreover, the rule feature $\langle\text{CM}\rangle$ inducing CLITIC MOVEMENT. A RULE FEATURE induces the application of the rule in question. $\langle\text{CM}\rangle$ is assigned according to the following $\langle\text{CM}\rangle$ ASSIGNMENT PROCEDURE, which corresponds to the flowchart of Figure 1.¹²

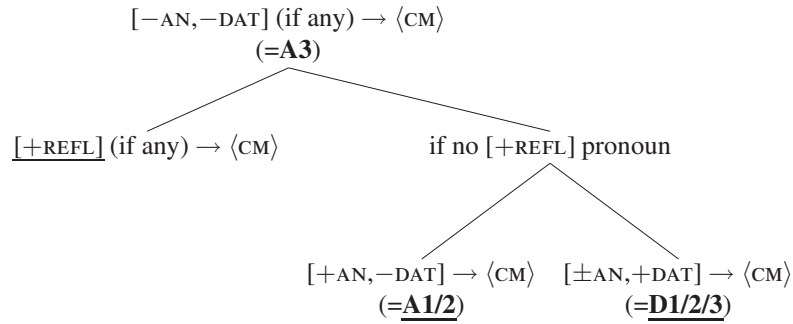
$\langle\text{CM}\rangle$ ASSIGNMENT PROCEDURE:

- a. $[-AN, -DAT]$ pronouns receive the rule feature $\langle\text{CM}\rangle$.
- b. $[+REFL]$ pronouns receive the rule feature $\langle\text{CM}\rangle$. Any further assignment of $\langle\text{CM}\rangle$ to pronouns not already marked $\langle\text{CM}\rangle$ is then blocked.
- c. If there are no $[+REFL]$ pronouns, $[+AN, -DAT]$ pronouns receive the feature $\langle\text{CM}\rangle$. Any further assignment of $\langle\text{CM}\rangle$ to $[+DAT]$ pronouns is then blocked.
- d. In the absence of any $[+AN, -DAT]$ pronouns, any remaining $[\pm AN, +DAT]$ pronoun receives the rule feature $\langle\text{CM}\rangle$.

If one is sensitive to the metaphor of semantic or functional weight, with + values adding weight, one may say that cliticisation is subject to a WEIGHT CONSTRAINT, in the sense that the lightest possible pronouns ($[-AN, -DAT]$) are always cliticised, $[+AN, -DAT]$ pronouns are cliticised only when there is no (relatively heavy) reflexive pronoun around, and $[\pm AN, +DAT]$ pronouns become clitics only when there is no $[+AN, -DAT]$ pronoun around.

Pronouns that do not receive the rule feature $\langle\text{CM}\rangle$ stay in position and are treated as canonical lexical argument terms. A dative pronoun not provided

12. One reviewer pointed out that forms like *pour le bien comprendre* 'to understand it well' occur in (formal) French and forms like *mi sempre chiama* 'he always calls me' in Calabrian. Apparently, short adverbs like *bien* or *sempre* are open to cliticisation in these languages, but only preverbally, never postverbally, which is probably why *pour le bien comprendre* has no equivalent in any variety of Italian. This is interesting, because it suggests that preverbal clitics are still word-like to the extent that full lexical adverbs may, exceptionally, become preverbal clitics, whereas postverbal cliticisation has become absorbed too much into affixal morphology for such wholesale lexical incorporation to be allowed.



“___”: no further <CM>-assignment

Figure 1. Top-down, left-to-right flowchart of the <CM> ASSIGNMENT PROCEDURE

with the feature <CM> will then undergo DATIVE EXTRACTION, as, for example, *à lui* in (4).

This being so, one would surmise that any system, whether or not feature-based, that caters for the phonological form of the pronominal arguments in question (cp. Bonet 1995) should be taken to be activated, in successive stages, after the feature <CM> has been assigned. The question, however, of the phonological form of clitics is only marginally at issue in the present paper.

Since, within the constraints of this system, no cluster can contain a double occurrence of [-DAT] or [+DAT] clitics, the flowchart of Figure 1 ensures that when a [+AN] or a [+DAT] pronoun is marked <CM>, any other pronoun under the same S also marked <CM> must be [-AN, -DAT] (=A3). Put differently, a C-cluster containing a [+AN] or a [+DAT] clitic can only co-occur with a [-AN, -DAT] (=A3) clitic.

CLITIC MOVEMENT being a postcyclic rule, the rule feature <CM> is put into action after (8b).¹³ It is defined as follows:

CLITIC MOVEMENT consists in ADOPTION by the highest lexical V-constituent under the first dominating VP.

13. <CM> is normally postcyclic, but cyclic for reflexives in complement-Ss. See Section 9.

A C-CONSTITUENT is a constituent under a syntactic category node specifying a lexical category, such as V(erb), N(oun), A(djective), and also PRON(oun), but not CL(itic) or AFF(ix). (“C” in “C-constituent” is thus a variable for “V”, “N”, “Adj”, etc.)

A C-constituent is LEXICAL when it is a lexical filler, possibly with affixes, but without clitics.

A constituent *A* is HIGHER than a constituent *B* when *A* dominates or commands *B*; when *A* and *B* are sisters, so that they command each other, *A* must stand to the left of *B* in a right-branching context and vice versa.

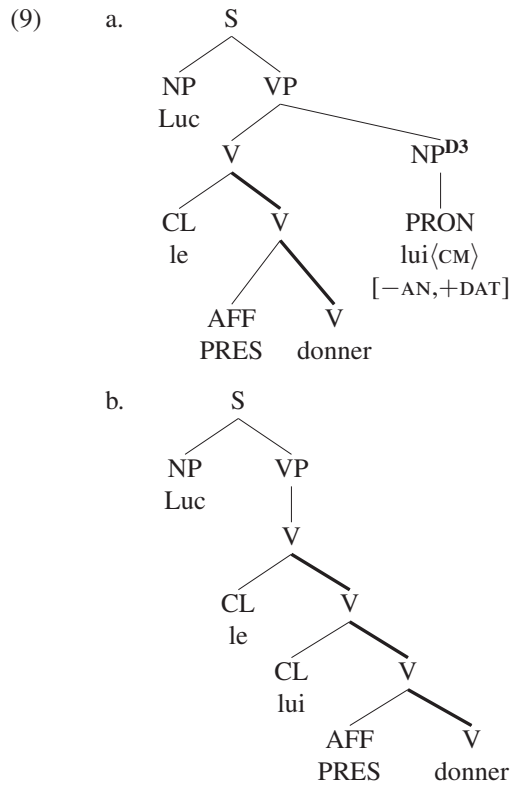
ADOPTION (Seuren 1996: 55) is a universal routine applied whenever a subtree *X* attracts a subtree *Y* as part of a rule specification. It is defined as follows:

ADOPTION takes place whenever a subtree *X* attracts a subtree *Y* as part of a rule specification, which indicates either *X* or *Y* as the INDUCER *I* of the process. The process consists of the following steps:

- a. A copy *X'* of the node label *X* is created and placed in the position immediately dominating *X*. *X'* inherits branching directionality from *X*, unless otherwise specified.
- b. The subtree *Y* is detached from its dominating node.
- c. *Y* is attached to *X'*, either to the left or to the right of the attracting subtree *X*, depending on the directionality of *X'*. If *X'* is set for right-branching, *I* (either *X* or *Y*) is placed left; if *X'* is set for left-branching, *I* is placed right.
- d. The attracted subtree *Y* is relabelled to its surface category.

BRANCHING DIRECTIONALITY is defined for (sub)constituents. In French syntax, as in that of most European languages, syntactic (sub)constituents are in principle right-branching. Imperatives and nonnegative 1-pl adhortatives are exceptional in that they are set for left-branching. In flecional morphology, left-branching is dominant, universally, as far as is known.

On the basis of the feature assignments in (6b), given a right-branching setting for the category V, adoption, at Shallow Structure, of the inducing subtree PRON[*le*] by the right-branching lexical V-constituent $v_{[AFF[PRES]V[donner]]}$ results in (9a), with *le* on the left (directionality is indicated by heavy lines). The attracted subtree PRON[*le*] is relabelled to its surface category CL. Subsequent cliticisation of the inducing PRON[*lui*] is supported by the same V-constituent $v_{[AFF[PRES]V[donner]]}$ and results in (9b), with *lui* below *le*. One sees that, in order to get the correct ordering *le-lui*, as in (7b), *le* must be cliticised before *lui*.



The question now is: what order of application of CM for those pronouns that have been fitted out with the rule feature $\langle CM \rangle$ helps to guarantee that all and only the correct C-clusters are generated in the simplest possible way? Since, in French (and Italian), clusters cannot contain more than one dative or accusative pronoun, the application order of French preverbal cliticisation must be specified as follows:

Ordering constraint for preverbal CLITIC MOVEMENT in French:
 (ne —) [+AN,+DAT] — [\pm AN,-DAT] — [-AN,+DAT] (— y — en)

But note that [+AN,-DAT] — [-AN,+DAT] (e.g., *me lui*) and [+AN,+DAT] — [+AN,-DAT] (e.g., *me vous*) are ruled out by the $\langle CM \rangle$ ASSIGNMENT PROCEDURE of Figure 1. The order of postverbal cliticisation is discussed in Section 8.

4. Some aspects of complementation in French

Clitic behaviour in complementation shows that cliticisation is structure-dependent. Three main cyclic rules are assumed in the complementation system of French (and Italian):

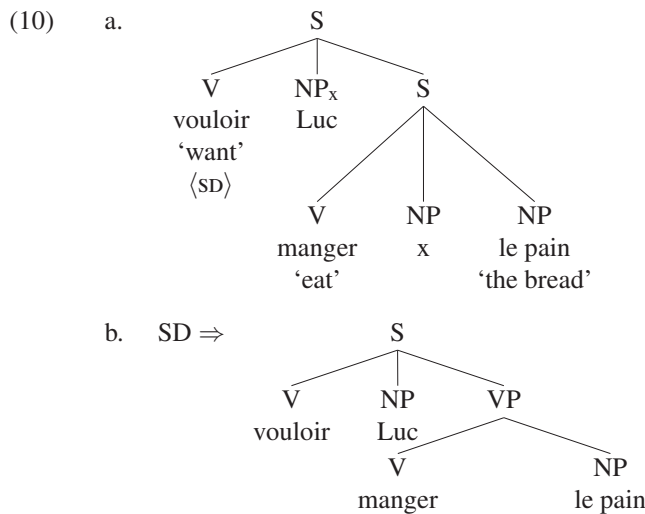
- (a) SUBJECT DELETION or SD (default for lower-subject referential identity)
- (b) SUBJECT RAISING or SR (default for subject complementation)
- (c) PREDICATE RAISING or PR (default for object complementation)

Each rule is induced by the embedding predicate (verb), which is lexically marked for any of the three rules. (As before, the marking is indicated by angled brackets.)

SUBJECT DELETION (SD):¹⁴

The subject-NP of an embedded S is deleted under conditions of referential identity with an argument NP of the superordinate S. Given the universal Convention 1 mentioned above, the embedded S is reduced to VP-status.

SD transforms (10a) into (10b), which ultimately gives (11) (tense is omitted here).



- (11) *Luc veut manger le pain.*
 'Luc wants to eat the bread.'

14. As relevant here; but see Seuren (1996: 68–71).

SUBJECT RAISING (SR) has been defined above. When the embedded S occupies the subject position, one speaks of SUBJECT-TO-SUBJECT RAISING or SSR; when it occupies the object position, one speaks of SUBJECT-TO-OBJECT RAISING or SOR. SOR is absent from French (and Italian), other than in exceptional cases (Ruwet 1972: 48–86).¹⁵

Consider sentence (13) with its syntactic analysis as given in (12a–c):

15. Chomskyan grammarians seem to have problems with this statement, probably because there appears to be a deeply ingrained aversion to PREDICATE RAISING (PR) in that school of linguistics. For them, sentences like *Je le crois malade* 'I believe him to be ill' or *Elle m'a vu chanter* 'she has seen me sing' are cases of SOR, not of PR. But *Je le crois malade* is a straightforward case of PR on the adjectival predicate *malade*. Were this not so, one would be hard put to explain sentences like *Je le lui crois fidèle* 'I believe him to be faithful to her'. Under SOR this would have to be either the ungrammatical or anyway undesirable (**)Je le crois fidèle à elle* or the clearly ungrammatical **Je le crois lui fidèle*. The same goes for sentences like *Elle m'a vu chanter*, which are likewise clear cases of PR. If SOR were at play here, one would have **Elle m'a vu le chanter* 'she has seen me sing it', which is ungrammatical. Instead one has *Elle me l'a vu chanter*, with dative *me*, which proves that it is PR, not SOR, which has been at work here.

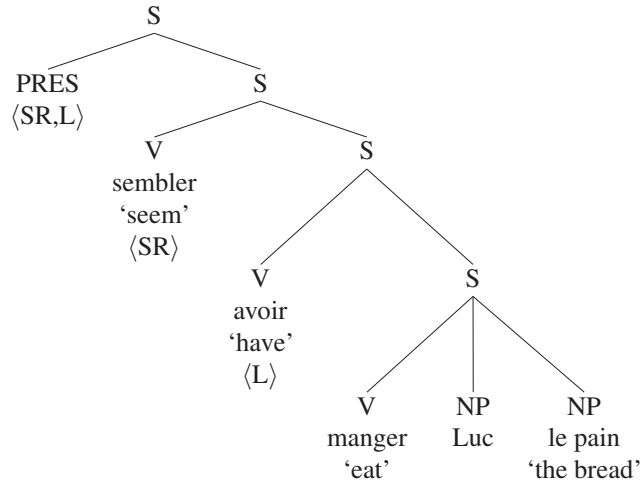
SOR does occasionally occur in French object-complementation, but only as an emergency measure. One such case is the following (Grevisse 1986: § 873.2):

- (i) *Ce mot d'estime LA faisait ME remercier.*
'This word of praise made her thank me.'

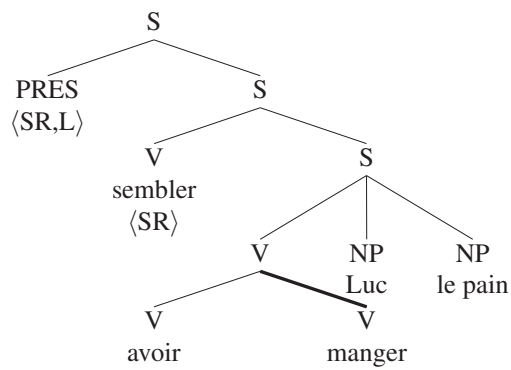
Such cases of exceptional SOR are motivated by the ungrammaticality of the clitic cluster that would come about if the normal syntactic rules were followed: (ii) is ungrammatical, as it combines a [+DAT] with a [+AN] clitic in one cluster, violating the <CM> ASSIGNMENT PROCEDURE:

- (ii) **Ce mot d'estime ME LUI faisait remercier.*

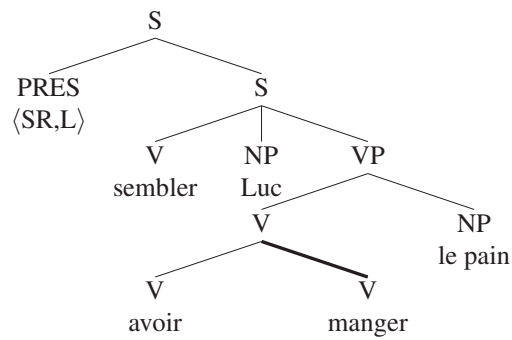
(12) a.



b. L ⇒



c. SR ⇒



- (13) *Luc semble avoir mangé le pain.*
 ‘Luc seems to have eaten the bread.’

LOWERING in (12a) is induced first (cyclically) by v [avoir], which has the surface category V(erb). LOWERING, in this case, consists in adoption of the inducing v [avoir] by the V-constituent of its subject-S, which is right-branching, so that v [avoir] ends up to the left of v [manger], as shown in (12b). (The S above *avoir*, having lost its V, is pruned; past participle formation is glossed over; see Seuren 1996: 48–49.) The next rule to apply is SR, induced by v [sembler], as shown in (12c). TENSE ROUTINE does the rest. One notices the remarkable career of the subject term NP [Luc]: having started out as the subject of the most deeply embedded S, it ends up as the subject of the highest S. Note also that when *le pain* is pronominalised as *le*, the resulting sentence is *Luc semble l’avoir mangé*, and not **Luc le semble avoir mangé* (see the examples (16)–(19) below). This shows that it is indeed SR that has applied here, and not PREDICATE RAISING, which is discussed now.

PREDICATE RAISING (PR):

The V-constituent of an embedded S or VP is adopted by the inducing verb. The lower S- or VP-node is deleted and all remaining material is re-attached higher-up, in virtue of the universal CONVENTION 2:

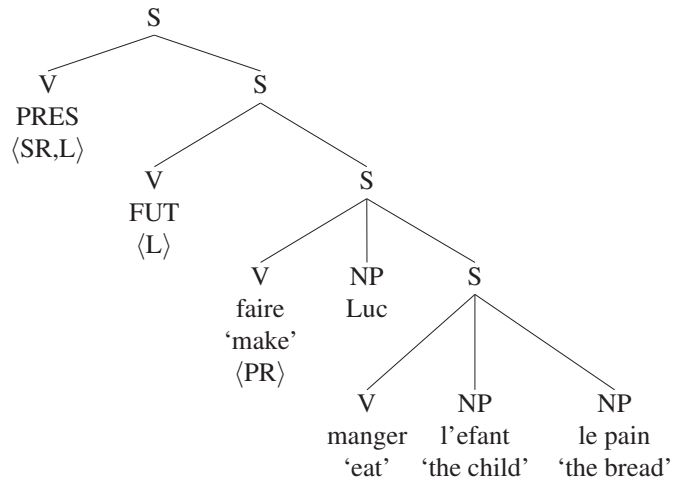
An S or VP whose V-constituent is removed loses its S- or VP-status and is, therefore, no longer a single constituent. All remaining material is re-attached higher-up in the existing order.

PR is the default raising rule for object-Ss in French, as it is in Italian, German, Dutch, Turkish, Japanese, and many other languages. A prominent PR-inducing verb in French is *faire* ‘cause, make’.

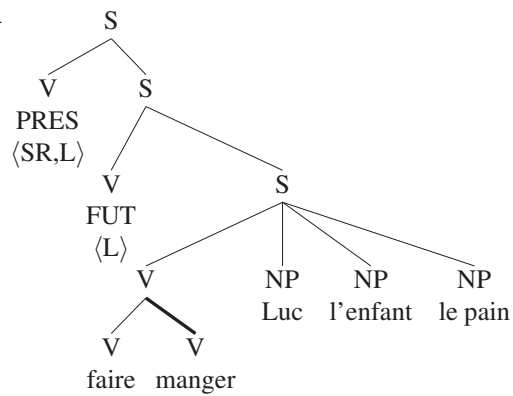
Consider (14), ultimately leading to (15). Note that (14) has two tenses, v [PRES] and v [FUT], both part of the AUXILIARY SYSTEM of the grammar. (14a) is turned into (14b) by PR on the most deeply embedded Cycle. The inducer v [faire], being right-branching, ends up to the left of the attracted constituent v [manger]. Now the lowest S has a new argument structure, with three nominal argument terms. This induces the rule DATIVE EXTRACTION, the result of which is shown in (14c). The next cyclic rule to apply is LOWERING (L) of v [FUT], which is lexically marked in French for the surface category AFF(ix). LOWERING, in this case, consists in adoption of the inducing higher v [FUT] by the highest lexical V-constituent of its subject-S. Since the attracting constituent is right-branching, FUT ends up as AFF[FUT] to the left of v [faire], as in (14d). (The S above FUT, having lost its V, is pruned.) The next cyclic rule to apply is SUBJECT RAISING (SR) on the top Cycle, resulting in (14e), followed by LOWERING (L), also on the top Cycle. PRES being again a surface AFFIX, the result

is (14f). (14f) is the Shallow Structure, as it marks the end of the cyclic part of the syntax, to be followed by the POSTCYCLE.

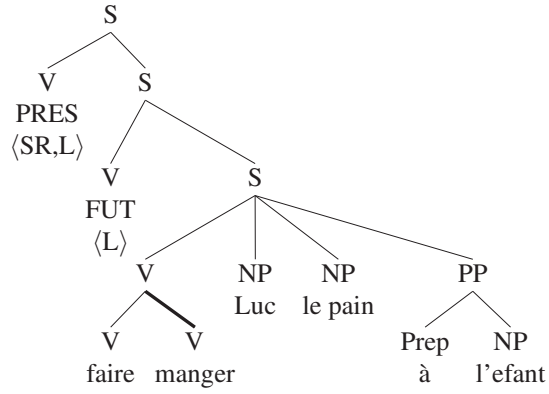
(14) a.



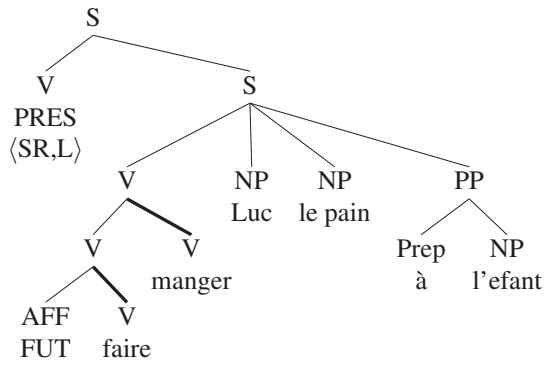
b. PR ⇒



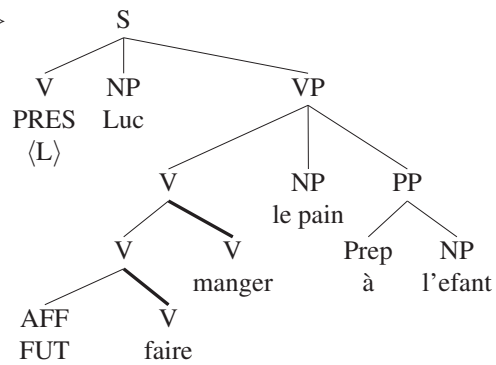
c. DATIVE EXTRACTION \Rightarrow



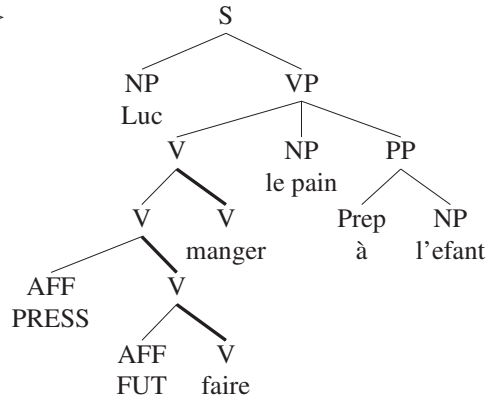
d. L \Rightarrow



e. SR \Rightarrow



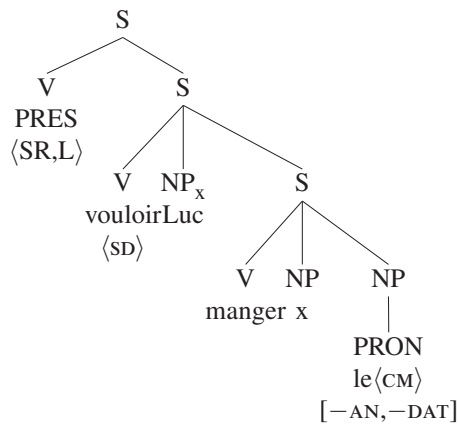
f. L ⇒



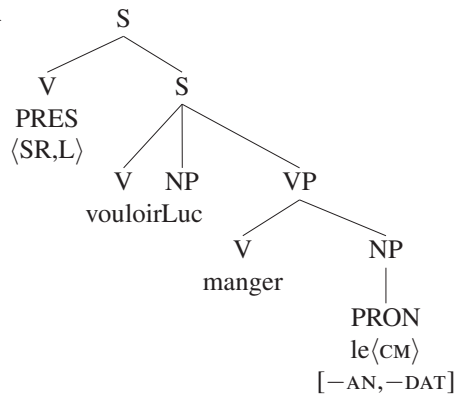
- (15) *Luc fera manger le pain à l'enfant.*
 Luc will-make eat the bread to the child
 'Luc will make the child eat the bread.'

The system as specified so far now predicts the correct positioning of C-clusters with regard to finite verb forms and embedded S-complements. Consider again (10a), now placed under the tense operator v [PRES] and with NP [le pain] replaced with the pronoun $PRON$ [le] which is assigned the features $\langle CM \rangle$ and $[-AN, -DAT]$, as in (16a). SD, SR, and L apply cyclically yielding the shallow NP-VP structure (16c). The constituent $PRON$ [le] is now cliticised in virtue of the feature $\langle CM \rangle$: $PRON$ [le] is adopted by the V-constituent of its own VP v [manger], as shown in (16d), which corresponds to the correct (17).

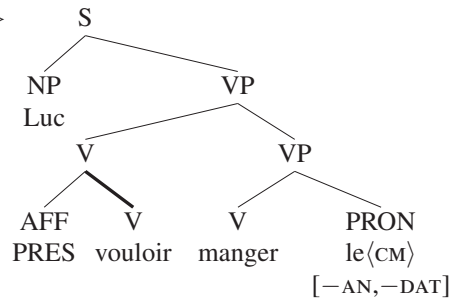
(16) a.



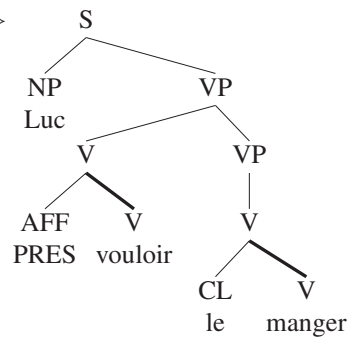
b. SD ⇒



c. SR, L ⇒



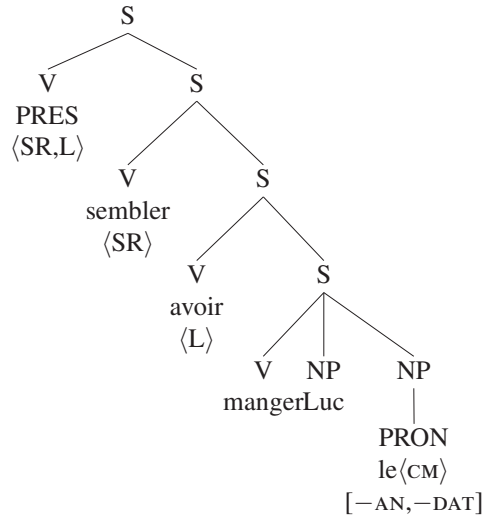
d. <CM> ⇒



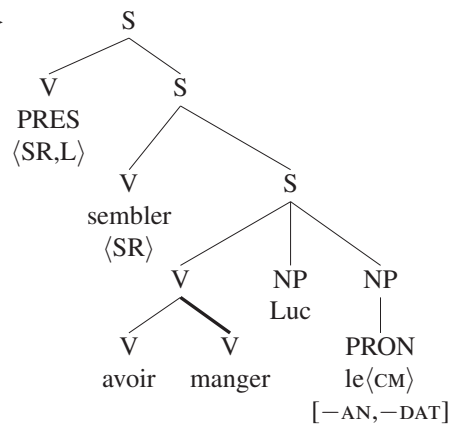
- (17) *Luc veut le manger.*
 Luc wants it eat
 'Luc wants to eat it.'

A similar treatment applies to (12a) with NP[le pain] replaced with PRON[le], since its shallow structure happens to be analogous to (16c), as shown in (18a–e).

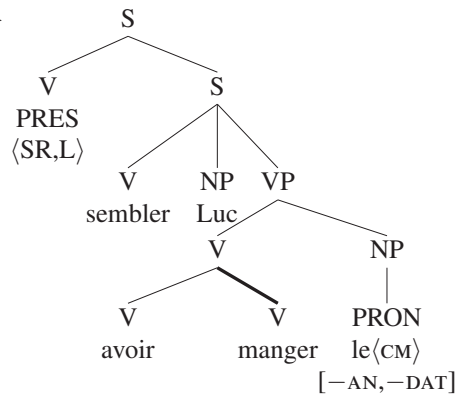
(18) a.



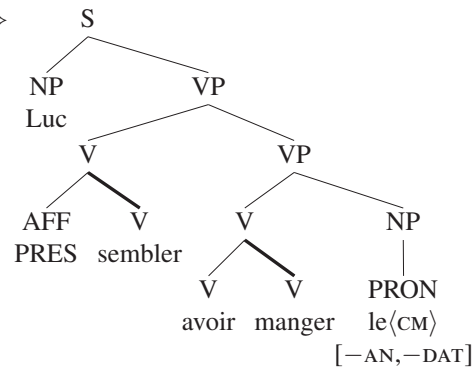
b. L ⇒



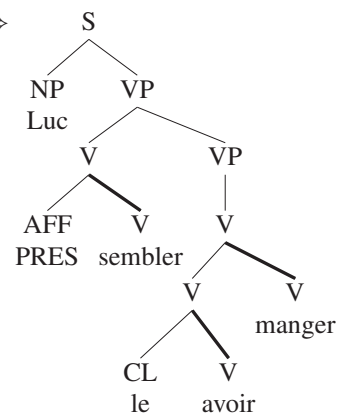
c. SR ⇒



d. SR,L ⇒

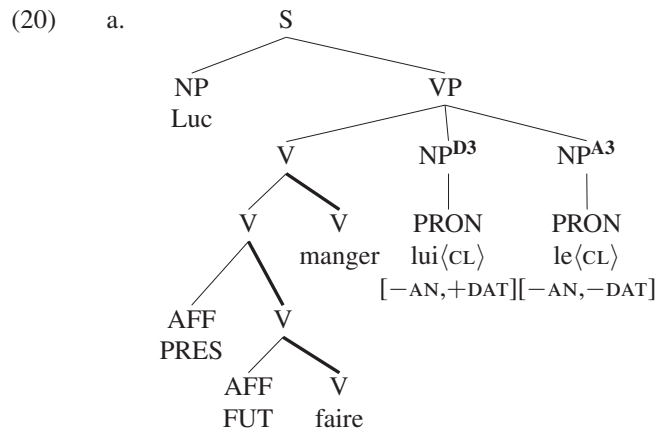


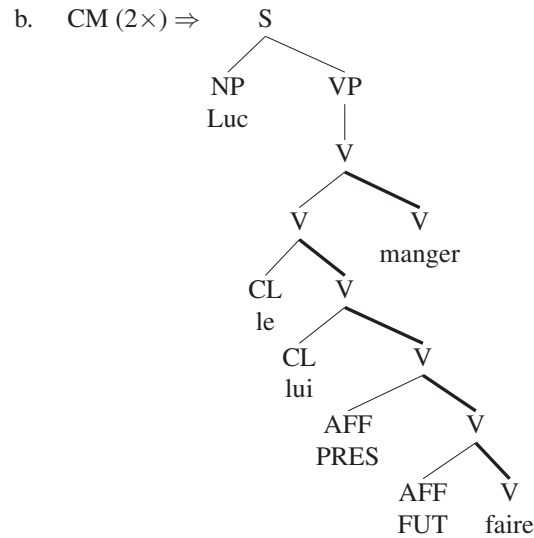
e. CM ⇒



- (19) *Luc semble l'avoir mangé.*
 Luc seems it have eaten
 'Luc seems to have eaten it.'

To see how cliticisation works for (15) one must take (14a) as the point of departure, with $\text{PRON}[\text{lui}]$ and $\text{PRON}[\text{le}]$ for $\text{NP}[\text{l'enfant}]$ and $\text{NP}[\text{le pain}]$, respectively. DATIVE EXTRACTION does not apply since the dative constituent *lui* is marked for the rule feature $\langle \text{CM} \rangle$. Given the end-cyclic or shallow structure (20a), cliticisation of $\text{PRON}[\text{le}]$ followed by $\text{PRON}[\text{lui}]$ gives (20b), which corresponds to the correct (21). It is clear that the rule of PREDICATE RAISING is responsible for the fact that the clitics end up in front of the finite verb *fera*, and not in front of the infinitive *manger*. This is an automatic consequence of the fact that *fera* and *manger* form one V-constituent under one VP: the clitics are adopted by the highest lexical V-constituent, which is $\text{V}[\text{AFF}[\text{PRES}]] \text{V}[\text{AFF}[\text{FUT}]] \text{V}[\text{faire}]$ for both *le* and *lui*.





- (21) *Luc le lui fera manger.*
 Luc it (to) him will make eat
 'Luc will make him eat it.'

Interestingly, the verb *laisser* 'let, allow' occurs in two guises. In one, it induces SD obligatorily (controlled by the higher direct object) and allows for optional PR. In the other it induces PR obligatorily and no SD occurs. The two versions have different argument structures and slightly different meanings, as shown in (22):

- (22) a. $v[\textit{laisser}] + \textit{subj} + \textit{dir-obj} + S$ SD, optional PR allow, permit
 b. $v[\textit{laisser}] + \textit{subj} + S$ PR let, make

The semantic difference is brought out clearly by the pair (23a, b):

- (23) a. *Luc a laissé le visiteur entrer.*
 Luc has let the caller come in
 'Luc has allowed the caller to come in.'
 b. *Luc a laissé entrer le visiteur.*
 Luc has let come in the caller
 'Luc has shown the caller in. / Luc has allowed the caller to come in.'

It follows from the analysis given that when PR has applied, with or without SD, there will be an ambiguity, whereas when only SD has applied, without any PR, as in (23a), there is no ambiguity and only the meaning 'allow, permit' is activated.

There is an interesting difference between French and Italian in the ways clitics are handled with embedded bare infinitives. When no PR is involved, modern French keeps the clitics in front of their own infinitive, as in *Je pourrais le vendre* ‘I could sell it’, *Tu sembles vouloir me le dire* ‘you seem to want to tell me it’. Only when the main verb takes PR, as in *Tu ne m’en feras pas prendre trop* ‘you will not make me take too much of it’ are clitics obligatorily placed before the finite verb form. In early modern French, however, (24a, b) were also allowed besides (17) and (19), respectively:

- (24) a. *Luc le veut manger.*
 Luc it wants eat
 ‘Luc wants to eat it.’
 b. *Luc le semble avoir mangé.*
 Luc it seems have eaten
 ‘Luc seems to have eaten it.’

Modern Standard Italian has both the construction of (17) and (19), and that of (24a, b), in free variation:

- (25) a. *Luca vuole mangiarlo.*
 Lucas wants eat-it
 b. *Luca lo vuole mangiare.*
 Lucas it wants eat
 (26) a. *Luca sembra averlo mangiato.*
 Lucas seems have-it eaten
 b. *Luca lo sembra aver mangiato.*
 Lucas it seems have eaten

But when PR is involved obligatorily, the clitics are placed before the finite verb form, as in French:

- (27) *Luca me lo farà mangiare.*
 Lucas me_{dat} it_{acc} will-make eat
 (**Luca farà mangiarmelo.*)
 ‘Lucas will make me eat it.’

There is a widespread belief, repeated in, for example, Miller and Monachesi (2003: 82–83), that in cases like (25b), (26b), or (27) a rule of CLITIC CLIMBING – sometimes optional, sometimes obligatory – has applied, raising the C-cluster from a lower to a higher VP. The above analysis, however, shows that CLITIC CLIMBING need not be assumed at all, since all we need to account for such facts is the independently motivated assumption that the verbs in question are lexically defined for OPTIONAL PR in the languages (and the periods)

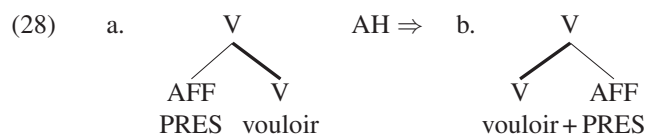
concerned.¹⁶ When PR has applied, the clitics end up in high position, just before the inducing verb. When PR has not applied, they end up in low position, in construction with the infinitival lexical V of the embedded VP. This follows directly from the machinery, in virtue of which *faire*-constructions lead to a single VP, owing to PR, whereas SD (without PR) and SR lead to embedded VPs. The generalisation now holds that cliticisation unites clitics with the highest lexical V-constituent of the first VP above them. CLITIC CLIMBING is now removed from the grammatical machinery and clitics can be taken to stay within their own VP throughout.

5. Affixation

We now specify the postcyclic syntactic treatment of affixes, constituting the transition to the morphology. Affixes induce the obligatory rule of AFFIX HANDLING (AH), which is postcyclic in the syntax but cyclic within lexical C-constituents. AFFIX HANDLING is defined as follows:

AFFIX HANDLING (AH): An AFFIX constituent is adopted by the highest lexical C-constituent commanded by it. The new C-constituent is left-branching if it was not left-branching already, so that the inducing AFFIX constituent is right-adopted. The new left-branching constituent forms one single syntactic word (in tree diagrams, word membership is symbolized as “+”).

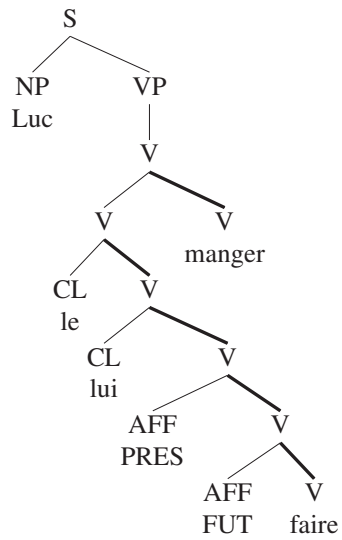
For example, AH turns the V-constituent $v_{[AFF[PRES]v[vouloir]]}$ of (16d), repeated as (28a), into the left-branching (28b), phonologically realized as *veut* (wants).



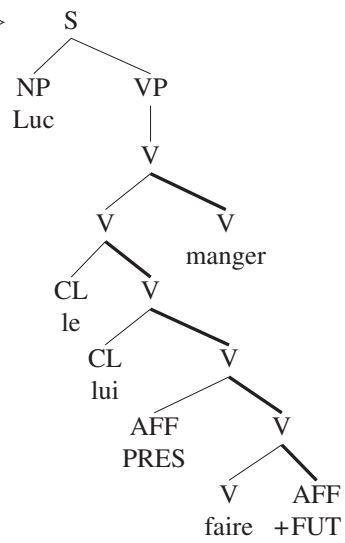
The V-cluster of (20b), repeated as (29a), is a more complex case. (29a) is turned into (29b), then into (29c), realized as *le lui fera manger*.

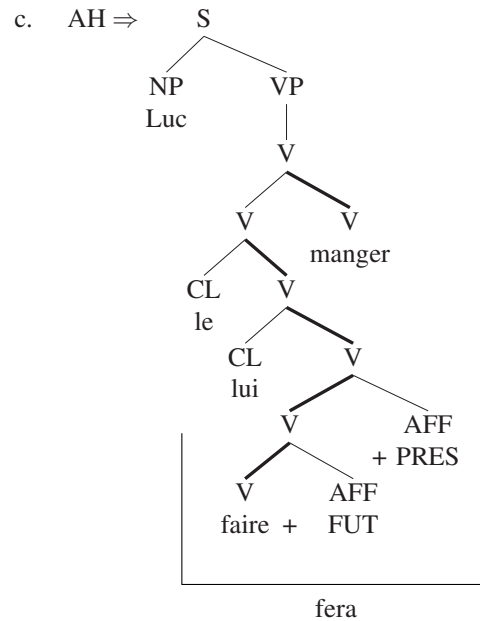
16. The corresponding verbs in German, *wollen* and *scheinen*, and those in Dutch, *willen* and *schijnen*, are lexically defined for obligatory PR (Seuren 2003).

(29) a.



b. AH ⇒





Clitics are normally labelled CL and are thus not affected by AH. Yet it pays to allow for an AFF labelling for clitics in certain cases, as is shown in the following section. Clitics appear to represent an area of transition between syntax and morphology.

6. Cliticisation in Italian (with a little Portuguese)

Preverbal cliticisation in Italian is largely, but not entirely, identical to French preverbal cliticisation. First, the plural third personal pronoun *loro*, unlike its French dative counterpart *leur*, is normally not cliticised but treated as a full lexical NP (but see note 9). So as to avoid not cliticising *loro*, speakers often use the singular dative *gli*, and in colloquial Italian even adverbial *ci*, for a plural reference, which may be resumed by the correct *loro* in final position (see, for example, Cardinaletti and Starke 1999: 169, who give the example *Gliel'ho dato loro* 'I have given it him, them'). As a subject pronoun, *loro* allows for pronominal subject deletion – so-called PRO-DROP – just like other subject pronouns.

Secondly, the **A3–D3** combination has the order reversed when compared with French: the combination of **A3** with **D3** does not result in the order **A3+D3**, as it does in French (*le/la/les–lui*), but in the inverse order **D3+A3**: *glielo* 'to him/her it', *gliela* 'to him/her it', *glieli* 'to him/her them_{masc}', *gliela* 'to him/her

them_{fem}'.¹⁷ All combinations of the clitic dative pronouns *gli* or *le* with following nonreflexive pronominal and adverbial clitics, except *ci/vi* 'there', are written as one single word. The combination of a **D3**-pronoun with *ne* 'thereof/from' therefore also results in a single-word orthography: *gliene* 'to him/her of it/them'. The question is: does Italian orthography mean something substantial in these cases?

One may well marvel at the intuitive analytical powers of those who designed the orthographies of the languages of culture, as their highly cultivated natural feeling for linguistic reality is often confirmed by linguistic analysis. Yet such intuitions can only have a heuristic value, if only because the orthographies based on them are often not entirely consistent. Consequently, orthographical conventions can never serve as a solid ARGUMENT in grammatical analysis (that would be a return to the introspective linguistics of the 1930s), but it is also unwise to ban them altogether from the theoretical arena.

The case at hand is intriguing, in that Italian orthography does point the way towards a welcome generalisation. Let us incorporate the following assumption into the Italian system:¹⁸

Italian preverbal **D3**-clitics are assigned the surface category **CL_{LEX}**, which makes it a lexical C-constituent. Any previously added (higher) pronominal **CL**-node in the same cluster is relabelled as **AFF**.

Given this assumption, we can say that French and Italian treat preverbal nominal clitics identically (disregarding the adverbial clitics *ne* and *ci/vi*):¹⁹

Ordering constraint for preverbal CLITIC MOVEMENT in Italian:
[+AN,+DAT] – [±AN,-DAT] – [-AN,+DAT]

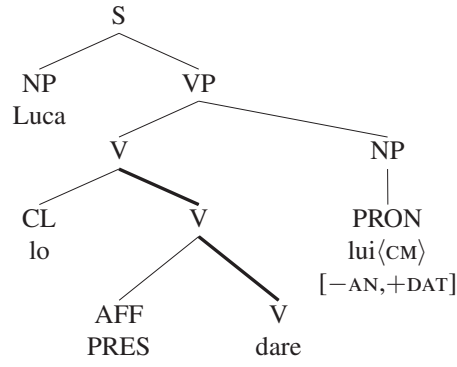
Now **CL_{LEX}[gli]**, being a lexical C-constituent, becomes the support of and will attract previously added clitics. Since **PRON[gli]** (= [-AN,+DAT]) cliticises as **CL_{LEX}[gli]**, **CL[lo]** in (30a) is relabelled **AFF[lo]** supported by **CL_{LEX}[gli]**, as shown in (30b). **AH** then results in (30c), which corresponds to the correct sentence (31).

17. All Italian clitics ending in *-i* and followed by another clitic beginning with *l-* or *n-* have the final *-i* changed into *-e*. For example, *mi lo* 'to me it' becomes *me lo*; *ci lo* 'to us it' becomes *ce lo*. Moreover, *gli lo* 'to him it' and *le lo* 'to her it' both become *glielo* (and likewise for *gliela/glieli/glielo*).

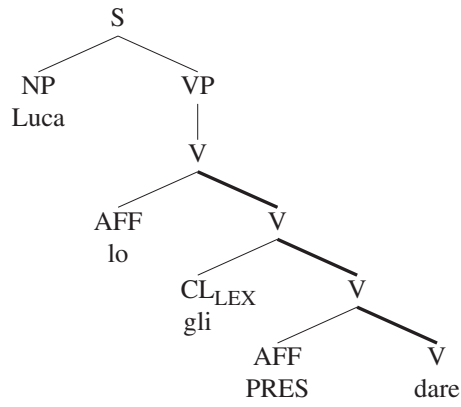
18. The proposal to treat enclitics as affixes was recently made for Portuguese by Duarte and Matos (2000) in terms of the Minimalist Program.

19. For adverbial *ci/vi* 'there' and partitive *ne* 'thereof, therefrom', the rule seems to be that they come last, as in French, and thus form the tail end of any cluster, in the order *ci/vi – ne*, except when combined with **A3** pronouns, which follow *ci/vi – ne*, as in *Ne l'ho tolto* 'I have taken it from it'. I will not pursue the peculiarities of *ci/vi* and *ne* in the present paper. They deserve a separate monograph.

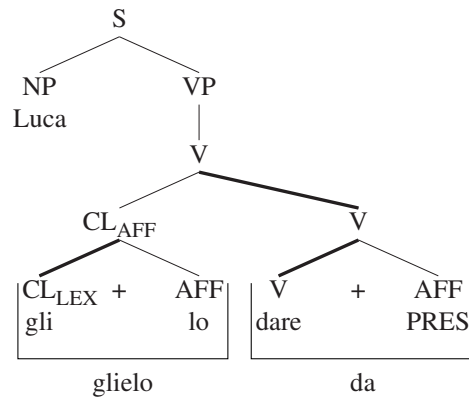
(30) a.



b. CM \Rightarrow



c. AH (2 \times) \Rightarrow



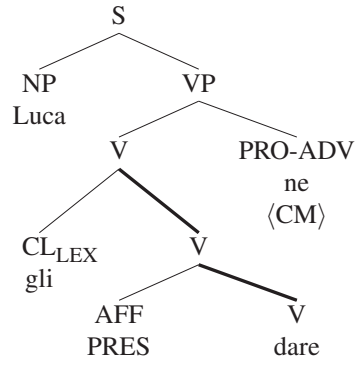
(31) *Luca glielo da.*
 Luca him-it give
 'Lucas gives it to him.'

This analysis has a double advantage. First, the French and Italian systems can now be unified to a greater extent than before, Italian merely having the extra feature of turning accusative clitics preceding **D3** clitics into affixes taking the **D3** clitic as support. Secondly, we have a rationale for the fact that official Italian orthography writes *glielo* etc. as one word.

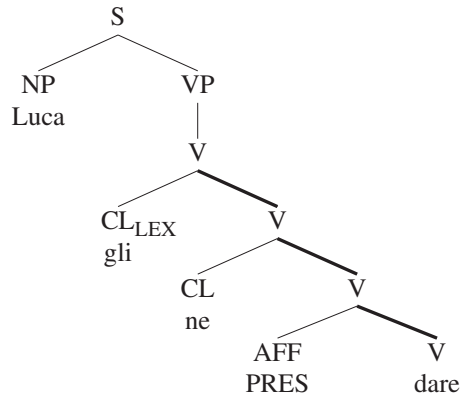
It must be borne in mind that the historical transition from a “weak” word preceding a support word to becoming a suffix of the support word is a widespread and well-known phenomenon. An example is Latin *mecum* ‘with me’, derived from *cum me*, which survives in Italian as *con me*, but in Spanish as *con migo* ‘with me’, where the preposition *con* ‘with’ is repeated as the suffix *-go*, still derived from Latin *cum*. A further example is the future tense paradigm in the Romance languages, where the future tense suffix derives from the Latin verb *habere* ‘have’. Whereas Latin *habere* could occur both before and after the support infinitive, the morphological process in Romance allows only for future tense suffixes, not prefixes. Consider also Homeric Greek, where prepositions can be turned into postpositions but then lose their accent, so that they are united with the support word into one phonological accent unit, not unlike the cases quoted in Note 2. Turkish “prepositions” are all postpositions, often obligatorily or optionally turned into affixes (for example, the postposition *ile* and the affix *-lu*⁴ (both ‘with’) are in free variation). The definite article in Danish and Rumanian is, under certain conditions, turned into an affix. Thus one has Danish *hus-et* ‘the house’ and Rumanian *drum-ul* ‘the road’. English has *homeward* besides *towards home*. And, of course, many more such examples could be produced. This phenomenon is no doubt connected with the fact that, universally, flecional morphology is left-branching and thus takes suffixes rather than prefixes.

In one respect, however, the orthography may be misleading, since **D3** clitics also form one word with *ne* ‘thereof/therefrom’, as in (33), even though *ne* is cliticised after *gli*, as one infers from its position in Italian C-clusters. We can, of course, adapt the definition of affixation to accommodate this fact, but such an extension of the definition lacks further general support, while affixation as described above is found all over the grammars of natural languages all over the world. It looks, therefore, as if we must consider Italian orthography to be at variance with linguistic reality in this detail (perhaps the phonological change from *gli* to *glie* has played a role). We thus take it that (33) is generated as in (32).

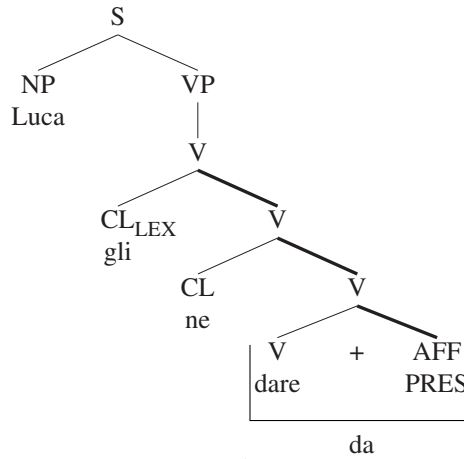
(32) a.



b. CM ⇒

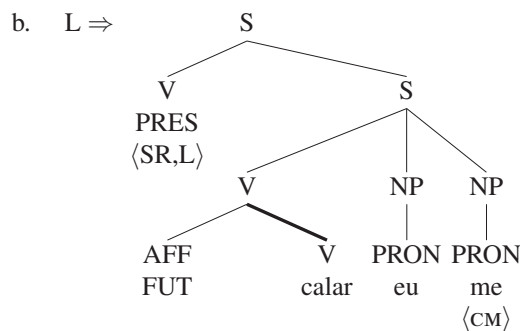
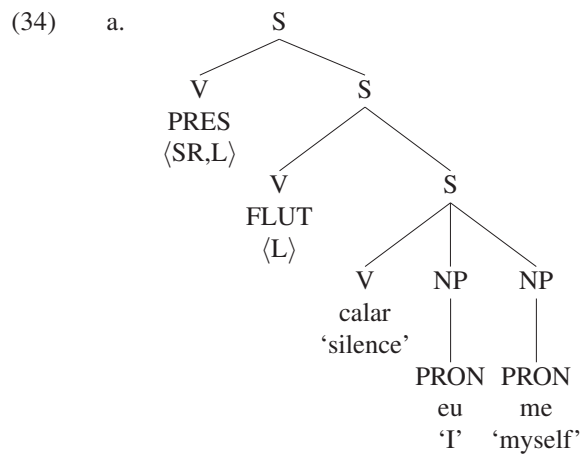


c. AH ⇒

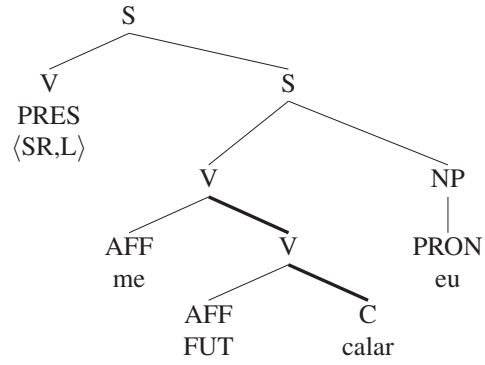


(33) *Luca gliene da.*
 Luca him/her-thereof/therefrom give
 'Lucas gives him/her of it.'

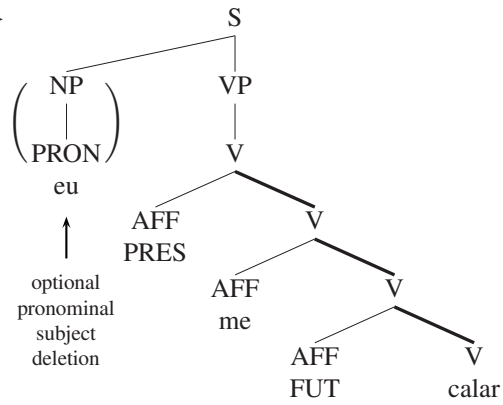
That the assumption of clitics occurring as affixes is not too outlandish appears from a remarkable phenomenon in the standard formal register of European Portuguese, which has so-called *mesoclitics*. Mesoclitics are affixes placed in the middle of a finite verb form belonging to the futurity paradigm, between the affix for futurity and the present or past tense/person affix. The use is ancient. Its history is based on the fact that the present or past future tense/person affix derives from the Latin verb *habere* 'have': 'I would read it' was originally construed as 'read it I had to'. An example is (da Cunha and Cintra 1985: 300) *calar-me-ei* 'I will remain silent', from *calar(se)* 'to remain silent'. *Calar-me-ei* thus derives from an original *calarme habeo* 'silence myself I have to', regrammaticalised as *calar* + FUT + *me* + PRES, where *me* is cliticised as an affix just above AFF[FUT], as in (34c). *Calar-me-ei* may perhaps be taken to be generated as in (34), with cyclic cliticisation, but further research is obviously required.



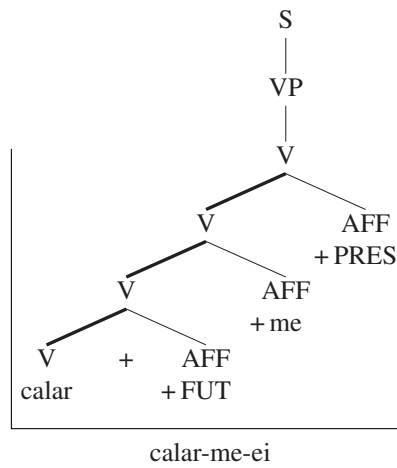
c. CM \Rightarrow



d. SR,L \Rightarrow



e. AH (3x) \Rightarrow



7. Intermezzo: the “lui-le-lui” phenomenon

Remarkably, certain (sociolinguistically deviant) forms of French allow for the combination **D3-A3-D3**, where one occurrence of **D3** is semantically vacuous (clitic doubling indeed). The following examples (complete with solecisms and misspellings) were taken from the internet:

- (35) a. *Devons-nous la laisser sucer son pouce ou doit-on lui le lui enlever une fois endormie?*
 ‘Must we let her suck her thumb or should one take *it* (*her*) out once she has fallen asleep?’
- b. *Revenu dans son bureau il se retrouve face à Aniya qui n’a pas apprécié ses manigances vicieuses et tient à lui le lui faire comprendre.*
 ‘Back in his office he finds himself again facing Aniya who has not liked his vicious tricks and fully intends to let *him* know *it*.’
- c. *Alors moi je ne savais pas si je dois lui le lui dire sans manifester aucune haine bien sûre.*
 ‘But I didn’t know if I should tell *him/her* (*it*), without showing any hatred, for sure.’
- d. *La différence avec les années 80, c’est qu’aujourd’hui vous avez le cran de lui le lui dire: “Oui, tu me déranges, casse-toi!”*
 ‘The difference with the ’80s is that nowadays you have the guts to say *it to him/her*: “Yes, you are disturbing me. Piss off!”’
- e. *Si vous souhaitez lui les lui faire essayer, ...*
 ‘If you wish to let *him/her* try *them*, ...’
- f. *Si José Bové pense que nos cités sont des camps, c’est peut-être sa visite à Ramallah ou Gaza qui lui le lui a inspiré.*
 ‘If José Bové thinks that our cities are camps, it was perhaps his visit to Ramallah or Gaza that gave him this idea (lit. inspired *this to him*.’
- g. *Ce journaliste a oublié qu’on n’est plus dans l’ère hassaniene ou au moment des tensions avec nos voisins nos journaux étaient “sollicités de faire plus preuve de patriotisme”! et on lui le lui a rappelé a ses depends alors ...*
 ‘This journalist has forgotten that we no longer live in the Hassan era when, as soon as there were tensions with our neighbours, our newspapers were “advised to give more proof of their patriotism,” and *he* has been reminded of *it* at his cost like.’
- h. *Et Harry lui le lui rendit.*
 ‘And Harry gave *it* back to *him/her*.’

- i. *Pierre Boussel n'a cessé de leur le leur répéter.*
'Pierre Boussel never stopped repeating *it to them*.'
- j. *Je suis pas vraiment bi mais j'aimerai faire l'amour avec une fille, c'est un gros fantasme quoi, mon copain est au courant, je lui le lui ai dit au bout de ...*
'I am not really bisexual, but I would like to make love to a girl, it's a great dream, like, my boy friend knows, I told *him (it)* after ...'
- k. *Dans la société camerounaise, on ne lui la lui offre pas toujours.*
'In Cameroonian society one does not always offer *her to him*.'
- l. *Quant la polémique concernant Dubé, je pense qu'il faut lui laisser sa chance, comme on lui la lui a laissé.*
'As regards the polemic around Dubé, I think it's proper to let him have his chance, as one has indeed let *him* have *it*.'

Similar examples in Italian appear to be less frequent, but they do occur:

- (36)
- a. *Leo, però, non capisce che Linka prova un debole per lui; cosa invece "avvertita" dalla moglie Lia, quando Leo gliela le presenta.*
'Leo, however, does not notice that Linka has a weak spot for him, something which his wife Lia does realize when Leo introduces *her to her*.'
 - b. *Le prendo le braccia e gliele le piego sul petto.*
'I take her arms and fold *them (her)* across her chest.'
 - c. *Non contento, afferra il paio di mutandine e gliele le strizza sugli occhi sbarrati.*
'Not satisfied, he grabs her panties and wrings *them* out (*to her*) over her wide-open eyes.'

This evidence is interesting in that it suggests that the status of French **A3-D3** clusters is not very robust in the sense that **A3** clitics followed by a **D3** clitic are vacillating between clitic and affix status. The copying of the **D3** clitic before the **A3** clitic may conceivably represent a transitional stage, reminiscent of Spanish *con migo* discussed in section 6. It seems significant that the Italian instances of this phenomenon are much less numerous than the French ones. Anyway, the facts are interesting enough to give one pause, but further research is obviously required.

8. Postverbal cliticisation in French and Italian

Apart from some phonological detail (mainly the diphthongisation of final *-e* into *-oi* in postverbal *me*, *te* and *se*, but not *le*, due to the fact that these cli-

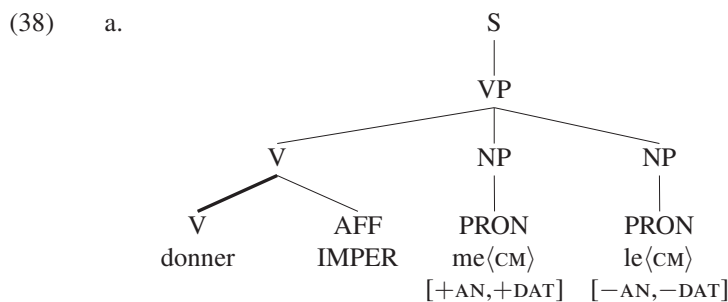
tics cannot receive word accent), French postverbal C-clusters differ from their preverbal counterparts only in the internal ordering of the constituent elements. There is, moreover, good reason to treat all postverbal clitics as suffixes, rather than clitics. In French, postverbal, or enclitic, cliticisation is restricted to positive singular and plural imperatives and positive 1-pl adhortatives, as in (37a–c):

- (37) a. *Donnons-le-lui.*
 let-us-give it him/her
 ‘Let us give it to him/her.’
 b. *Donne-le-moi.*
 give it me
 ‘Give it to me.’
 c. *Parlez-m’en.*²⁰
 speak me of it
 ‘Speak to me about it.’

The order of **A3** and **D1** or **D2** clitics/suffixes is different in preverbal and postverbal position: preverbally the order is **D1/2–A3**, as in *Il me l’a dit* ‘he has said it to me’, but postverbally the order is inverted to **A3–D1/2**, as in *Dis-le-moi* (‘say it to me’). Postverbally, French accusative clitics/suffixes always precede dative clitics/suffixes. The facts are straightforwardly accounted for by (a) assuming left-branching directionality for French positive imperatives and positive 1-pl adhortatives, (b) assigning to French postverbal pronominal elements the surface category AFF, and (c) imposing the following application order of the postverbal cliticisation/affixation rule for French:

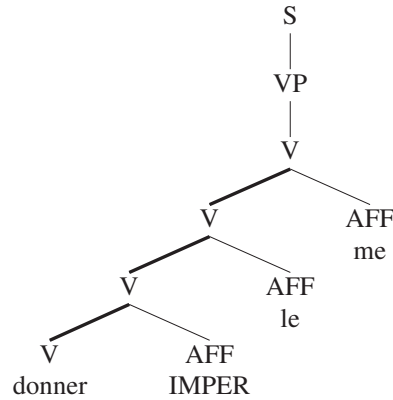
$$[-\text{DAT}] - [+ \text{DAT}] - y - en$$

This gives the derivation (38a–c) for (37b).

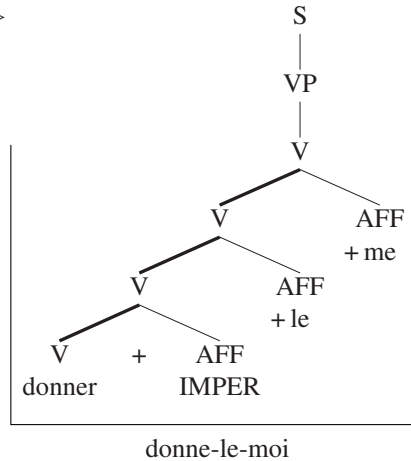


20. In spoken French, forms like *parlez-m'en* are sometimes replaced by *Parlez-moi-s-en*, with the strong form *moi* and an inserted spurious *s*.

b. CM (2×) ⇒



c. AH (3×) ⇒



One notes the crucial difference with the process of preverbal cliticisation. Preverbally, the surface category is CL but postverbally the surface category is AFF. According to the definition given, when a constituent labelled AFF is adopted by a lexical V-constituent, it becomes part of the constituent in question, which means that a subsequently added AFF will COMMAND the AFF constituent adopted earlier. In the case of preverbal cliticisation, however, a clitic adopted by a lexical V-constituent does not become part of the constituent in question, so that a subsequently added CL will BE COMMANDED BY the CL constituent adopted earlier. Therefore, when the [-DAT] constituent _{pron}[le] is adopted by the lexical V-constituent

$v[v[\text{donner}]_{\text{aff}}[\text{IMPER}]]$,

it forms a new lexical V-constituent

$v[v[donner]_{aff}[IMPER]]_{aff}[le]$,

to which the [+DAT] constituent $_{pron}[me]$ can now be added, forming the lexical V-constituent

$v[v[v[donner]_{aff}[IMPER]]_{aff}[le]]_{aff}[me]$.

Each newly added AFF now commands the entire lexical V-constituent, including affixes added earlier.

As regards French orthography, it is noteworthy that postverbal clitics are usually hyphenated with each other and with the supporting verb – a convention which used to be *de rigueur* in better days but is now gradually disappearing. In the analysis presented here, this hyphenation convention may be taken to be grounded in an intuitive awareness of the affixal nature of the postverbal clitics in the language.

The French hyphens are matched by Italian orthography, which writes postverbal clitics all as one word, connected with the verbal host. In Italian, postverbal cliticisation occurs not only with second person nonnegative singular and plural imperatives and nonnegative 1-pl adhortatives, but also with nonfinite verb forms (participles, infinitives, gerunds), as well as in certain fixed expressions involving third person reflexives, such as *vendesi* ('sells itself') or the plural *vendonsi* ('sell themselves', both: 'for sale', seen on billboards and in advertisements). Forms like *dandoglielo* ('giving it to him/her'), *dartene* ('to give you of it'), or *dammelo* ('give it to me') are perfectly current in Italian.²¹

Since, in Italian, all postverbal C-clusters are, as far as order and composition are concerned, identical to their preverbal counterparts, one might be tempted to assume that postverbal clitics are first generated preverbally, then clustered, and finally transposed as a cluster to postverbal position. More intellectually challenging, however, is the hypothesis that (a) Italian nonfinite verbs, nonnegative imperatives and nonnegative 1-pl adhortatives are obligatorily set for left-branching (but see note 21), (b) all postverbal clitics take the surface

21. Negative second person singular imperatives take the form of infinitives preceded by *non* ('not'). Clitics either follow or precede such infinitives, without any semantic difference (though the preverbal variety seems to be more current in the southern dialects):

- (i) *Non farlo!*
not do+it
'Don't do it!'
- (ii) *Non lo fare!*
not it do
'Don't do it!'

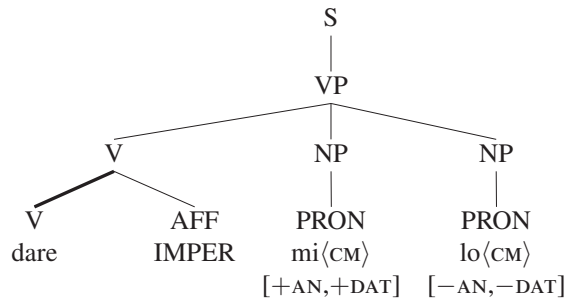
This is accounted for by allowing either left or right branching for such imperatives, depending on the regional variety of the language.

category AFF, and (c) other than in French, the application order of postverbal cliticisation in Italian is:

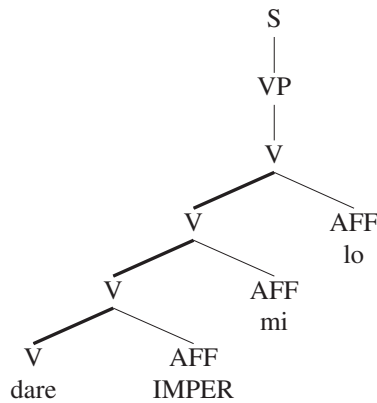
[+DAT] – [–DAT] – *ci/vi* – *ne*

The imperative *dammelo* ('give it to me') is then generated as in (39), with the dative *mi* affixed first, followed by the accusative *lo*.

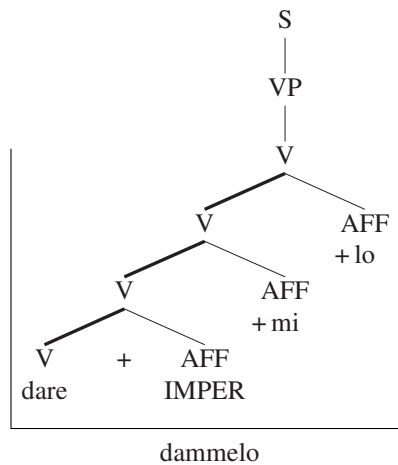
(39) a.



b. CM (2×) ⇒



c. AH (3×) ⇒



One wonders if there is a natural tendency for postverbal clitics to turn into affixes (suffixes), given the fact that morphological affixation is predominantly left-branching in the languages of the world. Since postverbal cliticisation results from the left-branching directionality of the verb forms concerned, AH makes no surface difference in these cases, which makes it natural for enclitics to merge with the category of affixes.

In the verbal C-clusters of Modern Greek, which only take pronominal clitics, datives always precede accusatives, both preverbally and postverbally, but they are always spelled as separate words, as shown in (40a,b):

- (40) a. *Mou to édhose.*
 me it (he-)gave
 ‘He gave it to me.’
 b. *Dhóse móu to.*
 give me it
 ‘Give it to me.’

The simplest solution seems to be to assume that clitics are labelled CL preverbally but AFF postverbally. This is in accordance with the fact, mentioned in note 2 and illustrated in (40b), that postverbal C-clusters and postnominal possessive pronouns partake in the phonologically defined accent distribution rules for words. The preverbal and postverbal orders of application of CM in Modern Greek are then identical, namely [+DAT] – [–DAT], just like the Italian postverbal order. Both preverbally and postverbally, dative *mou* ‘to me’ of (40a, b) is then cliticised first, followed by the accusative *to* ‘it’, which is placed to the right of *mou*, postverbally because *mou* is part of the lexical V-constituent, preverbally because it is not. The facts of official Greek orthography are then simply taken into the bargain.

9. Reflexive cliticisation

Reflexive pronouns pose special problems for any theory of cliticisation.²² Complications already appear to arise in cases without any form of complementation. The following sentences are both grammatical and natural in French and Italian, respectively, where the accusative pronoun may refer to, say, the person’s hands, and where the dative pronoun is an idiomatic so-called “ethical” dative, indicating some form of involvement on the part of the referent of the subject term (note that [+REFL] pronouns are always [+AN]):

22. On the special nature of reflexivisation in natural language, see Seuren (2001: 130–135). The main feature of reflexive pronouns is that they do not refer but reflexivise the predicate: ‘self-admire’, ‘self-hate’, etc.

- (41) a. *Il se les est lavées.*
 [+AN,+DAT,+REFL] [-AN,-DAT,-REFL]
 he to himself them is washed
 'He has washed them (for himself).'
- b. *Se le è lavate.*
 [+AN,+DAT,+REFL] [-AN,-DAT,-REFL]
 to himself them (s)he-is washed
 '(S)he has washed them (for himself/herself).'

In these cases, the reflexive pronoun is a normal part of the C-cluster. The $\langle \text{CM} \rangle$ ASSIGNMENT PROCEDURE of Figure 1 applies: the [+DAT] reflexive clitic *se/si* cannot co-occur with a [+AN] clitic, as shown in (42a, b):

- (42) a. **Il se t'est toujours imaginé comme une*
 he himself you is always imagined as a
personne honnête.
 person honest
 'He has always thought of you as an honest person.'
- b. **Si ti è sempre immaginato come una*
 himself you he-is always imagined as a
persona onesta.
 person honest
 'He has always thought of you as an honest person.'

Yet, contrary to the normal procedure which leaves datives uncliticised in case of a $\langle \text{CM} \rangle$ assignment blocking, the grammatical versions of (42a, b) are (43a, b), respectively, where the dative pronouns have been cliticised and the accusative pronouns have been left as full forms. The corresponding sentences (44a, b), where the inverse procedure has been followed, are grossly ungrammatical:

- (43) a. *Il s'est toujours imaginé toi comme une*
 he himself is always imagined you as a
personne honnête.
 person honest
 'He has always thought of you as an honest person.'
- b. *Si è sempre immaginato te come una*
 himself he-is always imagined you as a
persona onesta.
 person honest
 'He has always thought of you as an honest person.'
- (44) a. **Il t'est toujours imaginé à soi comme une personne honnête.*
 b. **Ti è sempre immaginato a se come una persona onesta.*

For that reason, the $\langle CM \rangle$ ASSIGNMENT PROCEDURE as specified in Figure 1 has been set up in such a way that, once a [+REFL] (and hence [+AN]) pronoun has been assigned the feature $\langle CM \rangle$, no other than a [-AN, -DAT] pronoun can have that feature under the same predicate. [+REFL] clitics can, therefore, be combined only with [-AN, -DAT] clitics. Any other pronoun remains uncliticised.

So far the system seems to work well. There is, however, the fact that in the southern Italian dialects that are not subject to the $\langle CM \rangle$ ASSIGNMENT PROCEDURE a sentence like (45) (in standard Italian orthography) is both grammatical and idiomatic. The sentence means literally 'I recommend myself to you', but is widely used idiomatically for 'Remember, don't forget'. In standard Italian one simply says *Mi raccomando*, without the dative pronoun *ti*:

- (45) *Mi ti raccomando.*
 [+AN, -DAT, +REFL] [+AN, +DAT, -REFL]
 myself to you I-recommend
 'I recommend myself to you.'

The order *mi ti* conflicts with the rule system given so far, which predicts the ungrammatical (46) since [+AN, +DAT] pronouns are cliticised before [+AN, -DAT] pronouns:

- (46) **Ti mi raccomando.*

The conflict can be repaired, without any harm to the system, by extending the preverbal application order of CM in both French and Italian as follows:

$$[+REFL] - [+AN, +DAT, -REFL] - [\pm AN, -DAT, -REFL] - [-AN, +DAT, -REFL]$$

That is, all reflexive pronouns are cliticised first, after which the nonreflexive pronouns are cliticised in the order as specified in Section 3. As a result, [+REFL] pronouns always come first in the preverbal C-cluster.²³ Yet, if this is correct, it does not apply to a special category of impersonal reflexive clitics in Italian, which is discussed below.

Since reflexive pronouns are not referential but reflexivise the verb (see Note 22), reflexive pronouns have a special relation with their verbs. Some verbs can

23. One reviewer called my attention to the fact that Calabrian has sentences like (i) (in standard Italian orthography):

- (i) *Mi ti sei arreso.*
 to me yourself are surrendered
 'You surrendered yourself to me.'

where, contrary to the proposed CM-order, the reflexive *ti* follows the dative *mi*. Obviously, further research is indicated for the dialects in question.

only occur reflexively, such as French *se repentir* and its Italian equivalent *pentirsi* (both ‘repent’; cp. English *perjure oneself*). Others are lexically reflexive: their reflexivity has been forged within the lexicon and has become idiomatic. Examples are French *s’en aller* and Italian *andarsene* (both ‘go away’), where any sense of reflexivity is gone. For such lexically reflexive cases, the clitics have no choice but to stay with the embedded verb in PR-constructions, as in (47a, b). Yet (47a) is not liked by French speakers, while Italian goes further, in this respect: (47b) is considered ungrammatical by native speakers, who do not accept any reflexive pronoun going with the infinitive in PREDICATE-RAISING constructions.

- (47) a. (?) *Je le ferai s’en aller.* (**Je se l’en ferai aller*)
 I him will-make go away
 ‘I will make him go away.’
 b. **Lo farò andarsene.* (**Se ne lo farò andare*)
 him I-will-make go away
 ‘I will make him go away.’

In cases where reflexivisation is due to a productive process and is controlled by the subject of the embedded S, the reflexive pronoun is preferably dropped in French. When it is not dropped, it is, apparently, cliticised during the syntactic Cycle, and not postcyclically, as is demonstrated in (48a). In Italian, the corresponding reflexive pronoun has to be dropped, as in (48b).

- (48) a. *Je lui ferai laver les mains.*
 I him will-make himself wash the hands
 (less preferred: *Je lui ferai se laver les mains*)
 ‘I will make him wash his hands.’
 b. *Gli farò lavare le mani.*
 him I-will-make wash the hands
 (**Gli [farò lavarsi/si farò lavare] le mani*)
 ‘I will make him wash his hands.’

The datives *lui* and *gli* in (48a) and (48b), respectively, show that the rule of PREDICATE RAISING has applied, and that this is not a case of emergency SOR, as in sentence (i) in Note 15. Accordingly, the accusative pronoun *le* is found in (49), since the reflexive *se* does not function as a direct object but is part of the embedded verb which has become intransitive by reflexivisation:

- (49) *Je le ferai (s’)asseoir.* (**Je lui ferai (s’)asseoir*)
 I him will-make sit down
 ‘I will make him sit down.’

Likewise, Italian has (50), with the same meaning, which suggests that the reflexive has been deleted during the Cycle:

(50) *Lo farò sedere. (*Gli farò sedere)*

With the always reflexive *se repentir* and *pentirsi* we get the same, as is shown in (51a, b) (both: 'I will make him repent it'):

(51) a. *Je l'en ferai repentir. (??Je l'en ferai se repentir)*
 b. *Ne lo farò pentire. (*Ne lo farò pentirsi)*

Here, again, French dislikes the reflexive *se* on the infinitive and Italian drops it.

However, when reflexivisation is controlled by the subject term of the superordinate clause, as in (52a, b), the reflexive pronoun is treated the way other pronouns are. Both (52a) and (52b) are assumed to have an underlying semantic structure corresponding to "he_x has caused [x be introduced by the boss]", transformed by PR into "he_x has _v[caused-to-be-introduced] x by the boss", which reflexivises the object term *x* into *se/si* (himself), subsequently cliticised in the normal way – that is, in the Postcycle:

(52) a. *Il s'est fait introduire par le chef.*
 he himself is made introduce by the boss
 'He had himself introduced by the boss.'
 b. *Si è fatto introdurre dal capo.*
 himself is made introduce by the boss
 'He had himself introduced by the boss.'

All this is consistent with the assumption that French pronouns that are reflexive within the embedded S are either deleted or cliticised during the Cycle, while in Italian such pronouns are deleted during the Cycle.

So much for "normal" productive reflexives. Both French and Italian, however, also have a form of reflexivisation which is akin to passivisation and may be called MIDDLE REFLEXIVISATION. This only occurs with transitive verbs and it consists in reflexivising the object term, which then becomes the subject term. The original subject remains unexpressed. Examples are:

(53) a. *I libri tascabili si vendono nelle*
 the books pocketable themselves sell in-the
stazioni ferroviarie.
 stations railway
 'Pocket books are sold at railway stations.'
 b. *Les livres de poche se vendent dans les*
 the books of pocket themselves sell in the
gares du chemin de fer.
 stations of the railway
 'Pocket books are sold at railway stations.'

- (54) a. *Delle volte si rompe un vetro.*
 some times itself breaks a window
 ‘Sometimes a window gets broken.’
 b. *Qualchevolta una finestra se rompe.*
 sometimes a window itself breaks
 ‘Sometimes a window gets broken.’

An example that has no direct counterpart in French (for reasons unknown) is the following:

- (55) *A lui, gli si spezzano le gambe ogni volta*
 to him him themselves break the legs each time
che fa lo sci.
 that he-does the ski
 ‘He, his legs get broken every time he goes skiing.’

The order *gli-si* follows from the rule ordering as specified. *Si* is cliticised first and assigned the category CL; then *gli* is cliticised as CL_{LEX}, changing the labelling of *si* into AFF. AFFIX HANDLING subsequently places *si* after *gli*.

What does not follow from the system as developed so far is the fact that (55) is grammatical in standard Italian. As soon as the feature ⟨CM⟩ has been assigned to a [+REFL] pronoun, the ⟨CM⟩ ASSIGNMENT PROCEDURE forbids CM for any other pronoun. The remedy seems to be to assign the feature [+REFL] only in cases of true reflexivisation, not in middle cases.²⁴ On that assumption, all is well, but for the spelling: one would expect *glisi*, spelled as one single word. This may indicate either that our system is not quite right or that the official spelling is not entirely consistent. The answer is left open for now.

Like canonical, nonmiddle reflexives, Italian cuts out middle *si* from infinitives resulting from PREDICATE RAISING. Thus, Italian has (56), without any middle *si*:

- (56) *A lui, gli farò spezzare le gambe.*
 to him him I-will-make break the legs
 ‘He, I will make sure that he breaks his legs.’

24. The southern Italian dialects, where the ⟨CM⟩ ASSIGNMENT PROCEDURE applies less strictly, allow for “true” reflexive *si* together with *gli*:

- (i) *Gli si è offerta senza condizioni.*
 to him herself she-is offered without conditions
 ‘She offered herself to him unconditionally.’

Again, the order *gli si* follows from the rule ordering as given and from AFFIX HANDLING.

Lest one doubts that (56) contains an underlying middle *si*, one should realize that it derives from an underlying causative construction *Io farò s[spezzare le gambe gli si]* ‘I will cause s[break the legs to him themselves]’, where the middle *si* cannot be missed in the embedded S.

Greater problems are posed by a form of third-person reflexivisation in Italian which may be called IMPERSONAL REFLEXIVISATION. This corresponds semantically, but not grammatically, with French *on* ‘one’, where a human subject, which can be first, second or third person, is semantically understood. Examples are (57a, b), which display a specific form of PRONOUN RESUMPTION: the first person plural subject term is resumed as an impersonal:²⁵

- (57) a. *Noi, il cervello si avrà piccino, ma lo*
 we the brain one will-have small but it
sappiamo adoperare.
 we-know use
 ‘We, we may have little brain, but we know how to use it.’
 b. *Parce que nous, on est là.*
 because we one is there
 ‘Because we, we are there.’

The French sentence (58) differs from (54b) in that (58) implies a wilful act of breaking, whereas (54b) does not:

- (58) *Quelquefois on casse une fenêtre.*
 sometimes one breaks a windows
 ‘Sometimes one breaks a window.’

In Italian, however, middle *si* often coincides at surface level with impersonal *si*, with the result that (54a) is ambiguous: besides the middle meaning given there, it also means ‘sometimes one breaks a window’, just like (58).²⁶

As regards the grammar of cliticisation, Italian impersonal *si* appears to behave differently from the normal reflexive and the middle *si*. The most obvious difference is that impersonal *si* stands close to the verb, from which it can be separated only by the clitic *ne*, as is shown in (59a–e).

- (59) a. *Lo si dice.*
 it one says
 ‘One says so./It is being said.’

25. (57a), which has a Tuscan flavour, is taken from Lo Cascio (1974); (57b) was observed in conversation.

26. One reviewer pointed out that (54a) loses its ambiguity when *un vetro* is preverbal, as in *Delle volte un vetro si rompe*, which only has the meaning ‘Sometimes a window gets broken’, that is, with middle *si*. It would seem that this has to do with the fact that impersonal *si* leaves *un vetro* as a direct object term, while, with middle *si*, it acquires subject status.

- b. *Ci si sveglia presto in montagna.*²⁷
 oneself one wakes-up early in mountain
 ‘One wakes up early in the mountains’
- c. *Se ne mescola il contenuto.*
 one of-it stirs the contents
 ‘One stirs its contents.’
- d. *Non mi se ne parla.*
 not to-me one of-it speaks
 ‘One does not speak to me about that.’
- e. *Da Napoli, se ne partirà domani.*
 from Naples one from-it will-leave tomorrow
 ‘One is due to leave from Naples tomorrow.’

One notes the striking difference between (60a), which contains impersonal *si*, and (60b), where dative *se* is a normal reflexive:

- (60) a. *Lo si toglie.*
 it one removes
 ‘One removes it.’
- b. *Se lo toglie.*
 to-himself it he-removes
 ‘He takes it off.’ (lit.: he removes it from himself)

The sentences (59a) and (60a) illustrate an important difference between impersonal and middle *si*. Impersonal *si*, as in (59a) and (60a), leaves the argument structure of the verb intact: *si* still acts as some sort of subject term, allowing for a direct object with transitive verbs. Middle *si*, on the contrary, turns the object term into a subject term, which is reflexivised on the verb. Therefore, the use of *si* with an intransitive verb automatically restricts the interpretation to impersonal *si*, as in:

- (61) *Si parte subito.*
 one leaves immediately
 ‘One (we/you/they) is (are) leaving immediately.’

Plurality of the middle subject transfers to the finite verb, as in (55) or:

27. *Ci si* in (59b) is a dissimilation from *si si* (see (i) in Section 2 above), where the first *si* is the normal reflexive of the intrinsically reflexive verb *svegliarsi* (‘wake up’), and the second *si* is the impersonal *si*.

- (62) *I soldi prestatì si devono rendere /*
 the moneys borrowed themselves must-pl pay-back /
devono rendersi.
 must-pl pay-back themselves
 'Borrowed money must be paid back.'

But impersonal *si* takes a singular verb form:

- (63) *I soldi prestatì, li si deve rendere.*
 the moneys borrowed, them one must-sg pay-back
 'Borrowed money, that one must pay back.'

Yet participial-adjectival agreement takes the plural form when a plural subject is understood. For example, the perfect tense counterpart of (61) shows a (masculine) plural participle, but with the singular finite form of the perfective auxiliary verb *essere* 'be', (always obligatory with reflexive verbs):

- (64) *Si è partiti da diversi giorni.*
 one is left-pl-masc since several days
 'One (we/you/they) left several days ago.'

On the whole, the grammar (and semantics) of impersonal *si* still hides many secrets. All we can do here is assign it its place in the clitics machinery. It is not to be assigned the feature [+REFL] but it always receives the rule feature <CM> and is cliticised before *ne* ('thereof/-from') but after [\pm AN, -DAT]. It cliticises, in some ill-understood and probably idiosyncratic way, the subject term, but it can only do so in a finite clause where *si* has subject status. Thus a sentence like (65), where *si* has become a direct object, is impossible:²⁸

28. French *on* is likewise impossible in a PR-construction:

- (i) **Jean fera partir on.*
 Jean will-make leave one
 'Jean will make one leave.'

French *on* and Italian impersonal *si* can only occur as subject. No other argument function, whether semantically original or derived, as in (i), is allowed. In some cases, one may think of deletion of impersonal *on/si*, as in (ii) or (iii) (both 'Everything suggests that ...'):

- (ii) *Tout fait penser que ...*
 (iii) *Tutto fa pensare che ...*

But the conditions under which such deletion can take place are unclear.

- (65) **Luca farà partirsi domani / si farà partire*
 Luca will-make leave-one tomorrow / one will-make leave
domani.
 tomorrow
 ‘Lucas will make one leave tomorrow.’

It follows that when there is an ambiguity between impersonal and middle *si* in a sentence involving PR applied in an S-complement, as in (66a), the alternative version without PR, with *si* attached to the lower verb, can only have the middle meaning:

- (66) a. *Non si deve mai rompere un vetro.*
 not itself/one must ever break a window
 ‘One must never break a window. / A window must never get broken.’
 b. *Non deve mai rompersi un vetro.*
 not must ever break-itself a window
 ‘A window must never get broken.’

This is as far as we can go. The grammar and semantics of impersonal, and to some extent also middle, reflexives is still opaque to a considerable extent, despite much valuable work recently done by grammarians. The same goes for the adverbial clitics *y* and *en* in French, and *ci (vi)* and *ne* in Italian. Given this overall lack of clarity, it seems advisable to leave the question of a full account of the clitics in question to later research.

Max Planck Institute for Psycholinguistics, Nijmegen
 pieter.seuren@mpi.nl

References

- Adams, James N. (1994). Wackernagel’s law and the position of unstressed personal pronouns in classical Latin. *Transactions of the Philological Society* 92.2: 103–178.
- Anderson, Stephen (2005). *Aspects of the Theory of Clitics*. Oxford Studies in Theoretical Linguistics 11. Oxford: Oxford University Press.
- Belletti, Adriana (1999). Italian/Romance clitics: Structure and derivation. In *Clitics in the Languages of Europe*, Henk Van Riemsdijk (ed.), 543–579. Berlin-New York: Mouton de Gruyter.
- Bloomfield, Leonard (1933). *Language*. New York: Holt.
- Bonet, Eulalia (1995). Feature structure of Romance clitics. *Natural Language and Linguistic Theory* 13: 607–647.
- Cardinaletti, Anna (1999). Pronouns in Germanic and Romance languages: An overview. In *Clitics in the Languages of Europe*, Henk Van Riemsdijk (ed.), 33–82. Berlin-New York: Mouton de Gruyter.
- Cardinaletti, Anna and Ur Shlonsky (2004). Clitic positions and restructuring in Italian. *Linguistic Inquiry* 35 (4): 519–557.

- Cardinaletti, Anna and Michal Starke (1999). The typology of structural deficiency: a case study of the three classes of pronouns. In *Clitics in the Languages of Europe*, Henk Van Riemsdijk (ed.), 145–233. Berlin-New York: Mouton de Gruyter.
- Cinque, Guglielmo and Giampaolo Salvi (eds.) (2001). *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*. Amsterdam: Elsevier.
- Cutler, Anne, John A. Hawkins, and Gary Gilligan (1985). The suffixing preference: A processing explanation. *Linguistics* 23 (5): 723–758.
- da Cunha, Celso F. and Luis F. L. Cintra (1985). *Nova gramática do Português contemporâneo*. Rio de Janeiro: Nova Fronteira.
- Dixon, Robert M. W. (1980). *The Languages of Australia*. Cambridge: Cambridge University Press.
- Duarte, Ines and Gabriela Matos (2000). Romance clitics and the Minimalist Program. In *Portuguese Syntax. New Comparative Studies*, J. Costa (ed.), 116–142. Oxford: Oxford University Press.
- Egerland, Verner (2005). Diachronic change and pronoun status: Italian dative *loro*. *Linguistics* 43 (6): 1105–1130.
- Emonds, Joseph (1999). How clitics license null phrases: A theory of the lexical interface. In *Clitics in the Languages of Europe*, Henk Van Riemsdijk (ed.), 291–367. Berlin, New York: Mouton de Gruyter.
- Fodor, Jerry A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fraenkel, Ernst (1932). Kolon und Satz: Beobachtungen zur Gliederung des antiken Satzes I. *Abhandlungen der Akademie der Wissenschaften in Göttingen, Philosophisch-Historische Klasse (1932)*, 197–213. Berlin: Weidmannsche Verlagsbuchhandlung.
- Fraenkel, Ernst (1933). Kolon und Satz: Beobachtungen zur Gliederung des antiken Satzes I. *Abhandlungen der Akademie der Wissenschaften in Göttingen, Philosophisch-Historische Klasse (1932)*, 319–354. Berlin: Weidmannsche Verlagsbuchhandlung.
- Grevisse, Maurice (1986). *Le bon usage. Grammaire française*. Douzième édition refondue par André Goosse. Paris-Gembloux: Duculot.
- Haverkort, Marco (1993). Clitics and parametrization. Case studies in the interaction of head-movement phenomena. PhD-thesis, Catholic University of Brabant, Tilburg, The Netherlands.
- Hawkins, John A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Holmberg, Anders (1991). The distribution of Scandinavian weak pronouns. In: *Clitics and Their Hosts*, H. van Riemsdijk and L. Rizzi (eds), 155–174. Geneva/Tilburg: EUROTYPE Working Papers, Theme Group 8: Clitics, Vol. 1.
- Horn, Laurence R. (1985). Metalinguistic negation and pragmatic ambiguity. *Language* 61 (1): 121–174.
- Kayne, Richard S. (2001). A note on clitic doubling in French. In *Current Studies in Italian Syntax. Essays Offered to Lorenzo Renzi*, Guglielmo Cinque and Giampaolo Salvi (eds.), 189–212. Amsterdam: Elsevier.
- Legendre, Géraldine (2000). Morphological and prosodic alignment of Bulgarian clitics. In: *Optimality Theory. Phonology, Syntax, and Acquisition*, Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (eds), 423–462. New York: Oxford University Press.
- Lo Cascio, Enzo (1974). Alcune strutture della frase impersonale italiana. In *Fenomeni morfologici e sintattici nell'italiano contemporaneo*, M. Medici and A. Sangregorio (eds.). Atti del sesto congresso internazionale di studi, Roma 4-6 settembre 1972. Vol. I, Part 1, 167–195. Società Linguistica Italiana. Rome: Bulzoni.
- McCawley, James D. (1970). English as a VSO-language. *Language* 46 (2): 286–299.
- Miller, Philip and Paola Monachesi (2003). Les pronoms clitiques dans les langues romanes. In *Langues romanes. Problèmes de la phrase simple*, D. Godard (ed.), 67–123. Paris: CNRS Éditions.

- Newmeyer, Frederick J. (1998). *Language Form and Language Function*. Cambridge, MA: MIT Press.
- Perlmutter, David (1971). *Deep and Surface Constraints in Syntax*. New York: Holt, Rinehart & Winston.
- Ruwet, Nicolas (1972). *Théorie syntaxique et syntaxe du français*. Paris: Du Seuil.
- Ruwet, Nicolas (1991). *Syntax and Human Experience*. Chicago-London: The University of Chicago Press.
- Schachter, Paul and Fe T. Otanes (1972). *Tagalog Reference Grammar*. Berkeley: University of California Press.
- Seuren, Pieter A. M. (1976). Clitic pronoun clusters. *Italian Linguistics* 2: 7–35. (reprinted in Seuren 2001: 267–290)
- Seuren, Pieter A. M. (1996). *Semantic Syntax*. Oxford: Blackwell.
- Seuren, Pieter A. M. (2001). *A View of Language*. Oxford: Oxford University Press.
- Seuren, Pieter A. M. (2002). Clitic clusters in French and Italian. In *Liber Amicorum Bernard Bichakjian, hem aangeboden ter gelegenheid van zijn 65ste verjaardag*, Haïke Jacobs and Leo Wetzels (eds), 217–233. Maastricht: Shaker.
- Seuren, Pieter A. M. (2003). Verb clusters and branching directionality in German and Dutch. In *Verb Constructions in German and Dutch* (= CILT 242), Pieter A. M. Seuren and Gerard Kempen (eds.), 247–296. Amsterdam/Philadelphia: Benjamins.
- Seuren, Pieter A. M. (2004a). *Chomsky's Minimalism*. New York, Oxford: Oxford University Press.
- Seuren, Pieter A. M. (2004b). The importance of being modular. *Journal of Linguistics* 40 (3): 593–635.
- Van Riemsdijk, Henk (ed.) (1999). *Clitics in the Languages of Europe*. Berlin-New York: Mouton de Gruyter.
- Zwicky, Arnold (1994). Clitics. In *The Encyclopedia of Language and Linguistics*, R. A. Asher and J. M. Y Simpson (eds.), 571–576. Oxford: Pergamon Press.