

*Clitics in Word Grammar**

RICHARD HUDSON

Abstract

Clitics are a challenge for any view of the architecture of grammar because they straddle the boundaries between words and morphemes and between syntax and morphology. The paper shows that clitics are syntactic words which also serve as word-parts, so their presence is explained in terms of syntactic dependencies, but their position follows morphological rules. The general analytical framework which is proposed builds on the theory of Word Grammar. As expected, clitics do demand a collection of special analytical categories - the word-classes Clitic and Hostword, and the relationships 'host', 'clitic', 'finite verb' and 'extension' - but (unlike other current theories of cliticization) they do not need any extra theoretical apparatus. The paper considers simple clitics in English and special clitics in French and Serbo-Croat.

1 The Resources of Word Grammar

The aim of this paper is to show that clitics (defined in section 2 below) can be accommodated in a grammatical framework such as Word Grammar (**WG**¹) which offers the following austere analytical apparatus:

- The word is the only unit of syntax, with phrases treated as epiphenomena; this is typical of dependency analysis but conflicts with phrase structure, and rules out any analysis which involves:
 - rebracketing (Sproat 1988) for the simple reason that there are no brackets (or anything equivalent);

*This paper arose out of a discussion with Amela Camdzic, who asked what WG has to say about clitics. I should like to thank her for stimulating discussion, and Chet Creider for comments on an earlier version.

¹The most relevant publications on Word Grammar are: Creider and Hudson 1999; Hudson 1984; Hudson 1990; Hudson 1998; Hudson 2000b; Hudson 2000a; Hudson 2000c. More information is available on the WG website (<http://www.phon.ucl.ac.uk/home/dick/wg.htm>), including a freely downloadable 130-page "Encyclopedia of English grammar and Word Grammar" which is updated every other year.

- a contrast between bar-levels or heads and maximal projections (Grohmann 2000);
- rules which locate material in relation to clause boundaries (Anderson 1993).
- There is only one kind of word, and only one ‘level of analysis’, so for clitics we cannot invoke differences such as:
 - that between PF words and SS words (Sproat 1988; Marantz 1988),
 - that between c-structure words and functional words (Bresnan 2001:93),
 - that between syntactic and morphological words (Sadock 1991).
- There are no empty categories, including functional categories, so the analysis of clitics cannot involve these (Grohmann 2000; Terzi 1999).
- There are no processes, which rules out analyses in terms of:
 - movement to spec of CP or (long) head movement (Rudin, Kramer, Billings, and Baerman 1999; Borsley and Rivero 1994; Rudin, Kramer, Billings, and Baerman 1999; Wilder and Cavar 1994),
 - incorporation or head-incorporation (Borsley and Rivero 1994; Rudin, Kramer, Billings, and Baerman 1999),
 - adjunction (Terzi 1999),
 - spell-out rules (Bonet 1995)
- The organising principles of syntax and morphology are quite different, syntax being organised in terms of word-word dependencies while morphology is based on whole-part relationships; this rules out analyses in which the order of clitics is determined by purely syntactic principles and encourages analyses in which they fit into rigid morphological ‘templates’ (Miller and Sag 1997).

Alongside this very strict limitation on the kinds of analytic machinery that are available, WG offers considerable flexibility in two areas where there is ample independent justification for richness:

- It allows indefinitely rich **sub-classification**, so we shall be able to recognise as many sub-classes as are needed for words and for morphological forms. Complex intersecting classes can be handled by means of multiple default inheritance which allows an item to inherit by default from several different super-categories; for example, a clitic pronoun can inherit from Pronoun and from Clitic.
- It allows new **relationships** to be postulated as needed. The basic idea behind WG is that language is a cognitive network whose nodes are words, word-types, morphemes, semantic concepts and so on. These nodes are linked by means of typed relationships like the attributes of an attribute-value matrix, but unlike other such

systems the attribute types are themselves part of a sub-classification (‘ISA’) hierarchy which may be extended downwards and sideways as needed (Hudson 2000a).

The analysis of clitics offered below will use a small number of special sub-types of word and a somewhat larger set of special relationships. Otherwise it will stay within the sparse architectural specification for WG grammars.

2 The challenges of clitics

A clitic is “... an independent syntactic constituent which shows up phonologically as part of a derived word.” (Marantz 1988:253). In other words, a clitic is a unit which is:

- a distinct word for syntax, but
- a mere morpheme for morphology and phonology.

For example, *you’re* contains the clitic verb *’re* in (1).

(1) You’re wrong.

This unit *’re* must be a separate word because it is a verb; in short, the sentence has precisely the same syntactic structure as (2).

(2) You are wrong.

However *you’re* is also a word, with *’re* as one of its parts. The main evidence for this is phonological: the sequence /jɔ:/ is indivisible and not composed regularly out of the pronunciations of the two words - in contrast, say, with *you’re* as in (3).

(3) The pictures of you’re good.

It follows that the pronunciation of the first *you’re* must be stored ready-made as a single ‘word’.

The first challenge is to reconcile these two conflicting claims about word-hood. The solution is obvious: to recognise different kinds of word, with *’re* as one kind of word and *you’re* as a different kind. However we cannot simply invent new kinds of

words and claim to have solved the problem; we must show how the different categories fit together in the theory. In my own earlier discussions of clitics (e.g. Hudson 1990:104) I have claimed that both clitics and the units that contain them (e.g. *you're*) are simply 'words', albeit not very typical words. After some uncertainty I now think this is still the most promising approach, as I shall explain below.

The second challenge is to accommodate Zwicky's 'special clitics' (Zwicky 1977), which are special in terms of their position in the sentence. The kind of example that I have discussed repeatedly (Hudson 1990: 325) is (4), from French.

- (4) Paul en mange deux.
 Paul of-it eats two
 'Paul eats two of them.'

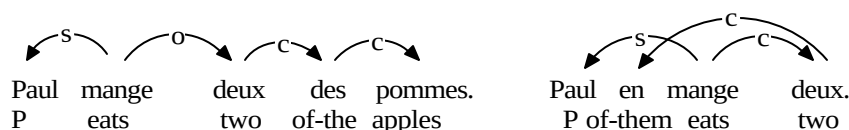
In this example the clitic pronoun *en*, 'of it', is attached to the verb *mange* according to the normal rules for clitic pronouns. However it is in complementary distribution with an ordinary prepositional phrase such as *des pommes*, 'of the apples', which would be placed by the normal rules of syntax after *deux*.

- (5) Paul mange deux des pommes.
 'Paul eats two of the apples.'

The point of this example is that *en* must depend on *deux*, and not on *mange*. In French, a quantity noun such as *deux* requires a complement, so (unlike English), (6) is not permitted.

- (6) *Paul mange deux.

This is why *en* must depend on *deux*, as its complement. The syntactic structure must be as shown in Figure 1, which contrasts it with that of (5)².



² Strictly speaking, *des* should be analysed syntactically as two words, *de* and *les*, which share a single word-form, but this point is irrelevant here.

Figure 1

The dependency link from *deux* to *en* also rules out another attractive analysis, according to which clitic pronouns are simply affixes which absorb the verb's valency requirements (Miller and Sag 1997); this cannot be the right analysis for (4), because the clitic does not affect the verb's valency, but does affect that of the verb's object, *deux*.

Special clitics such as this French clitic pronoun offer three challenges to grammatical theory:

- How to define their host (e.g. *mange*) in cases where the clitic and the host are not directly linked by a dependency.
- How to define their position in relation to their host and to other clitic pronouns attached to the same host.
- How to explain the discontinuity (shown by the tangling dependency arrows in Figure 1) that their special position may produce.

The challenges are the same, but harder to solve, in languages such as Serbo-Croat in which numerous clitics - auxiliary verbs, pronouns and particles - are attached in 'second position' within the clause (Spencer 1991:351-8), as in (7).

- (7) a. Ja *mu ga* dajem svaki dan.
 I *him it* I-give every day
 'I give it to him every day.'
 b. Svaki dan *mu ga* dajem.
 c. Dajem *mu ga* svaki dan.

Although the order of elements within a clause is in general free, the clitics (italicized) must follow the first element (and may even be attached to the first word of the first phrase).

The main point of this paper is to develop a theory of clitics within the general theoretical framework of Word Grammar. This is by no means the first attempt. Apart from my own rather brief discussions (Hudson 1984:48-50, 86-7; Hudson 1990:106, 118-9), there is a full PhD dissertation on the very complex Italian clitic pronouns (Volino 1990). The theory proposed here can be seen in many respects as a development of this early work, in which:

- clitics are treated as separate syntactic words with ordinary dependency relations to other words (in contrast with more recent analyses in which clitic pronouns are treated as mere inflectional affixes - Sag and Godard 1994; Miller and Sag 1997),
- their position is accounted for by postulating extra syntactic relationships which allow 'clitic climbing' (discussed below).

However Volino's analysis is purely syntactic, whereas I shall invoke a morphological relationship to a larger word, as explained below. After a discussion of the units of grammar, I shall consider simple clitics and then special clitics.

3 Compound words and morphemes in English

My present view is that the only linguistic units that are relevant to clitics are the following:

- Word
- Morpheme and Word-form, subsumed under Form
- Phoneme

To put this discussion into historical perspective, in 1984 the only linguistic units that I recognised were Word and Phoneme - I explicitly denied the need for Morpheme (Hudson 1984:55). However I did accept the need for Morpheme in 1990 (Hudson 1990:85), and I still think this is correct. This means that words, as such, do not have phonological properties; for example, the stem of the word CAT is the morpheme {cat}, and it is the latter, rather than the word, that begins with /k/.

Another issue where I still hold the beliefs I expressed in my earlier work is the analysis of compounds (Hudson 1984:50-2), which I think are words that contain smaller words. For example, FIELDMOUSE can be recognised as a single word alongside its component words FIELD and MOUSE. Since the components are ordinary words, this analysis allows us to recognise a syntactic relationship between them, as shown in Figure 2 (where as usual the small triangle signals the supercategory in an 'isa' classificatory link). This syntactic relationship has all its expected consequences for both word order and meaning. The only difference between the analysis shown and the analysis that would have been given for the two-word combination *field mouse* is the presence of the word FIELDMOUSE and its 'part' relationships. (In words, the word FIELDMOUSE has two parts which are instances of the words FIELD and MOUSE. It also has the sense

Fieldmouse, which is a Mouse - the sense of MOUSE. Inside FIELDMOUSE, FIELD depends on MOUSE just as in *field mouse*.)

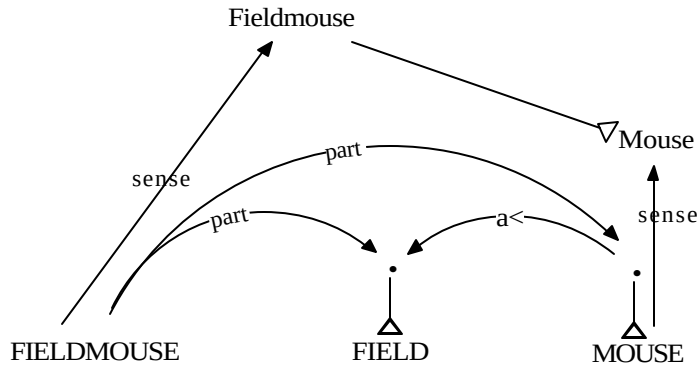


Figure 2

Some arguments for this syntactic analysis within the word are:

- It explains why the relationships between the component words are exactly the same as they would have been if the words had been separate (i.e. *field mouse*): they have the same syntax, semantics and phonology, so the only observable difference is in the spelling.
- In particular, by recognising the ordinary word MOUSE as a part of FIELDMOUSE we explain why they both have the same irregular plural: *fieldmice*.
- The analysis explains why both linguists and writers find compound words so hard to distinguish from word-word combinations; as we all know there is great uncertainty about the use of word spaces and hyphens in such cases.

One consequence of this analysis of compounds, which I did not comment on in earlier work, is that the stem of a compound word must be a single unit with the characteristics of a morpheme (e.g. the ability to serve as a word's stem). We can show this morpheme-like characteristic by using the traditional morpheme brackets: {fieldmouse}. But what do we call {fieldmouse}? It would be unhelpful to call it a morpheme, because this term is traditionally used only for 'minimal' morphological units. Instead I shall use the rather vague word 'FORM' as a technical term for all units at this level of analysis, with Morpheme as one particular manifestation. Other types of form include the form of the entire word, complete with inflections ('WORD FORM'), and possibly other forms such as the intermediate forms recognised in traditional inflectional morphology (e.g. the various tense forms in Latin to which person and number affixes

are added). This hierarchical subclassification of forms is clearly different from any hierarchy that would apply to words, as can be seen from Figure 3. (In prose, a word may be noun, verb, etc and it may also be a clitic; for example, the possessive 'S is both a noun and a clitic, while the reduced form of *is* is both a verb and a clitic; these two examples will be discussed below. In contrast, a Form may be an affix, a compound-form or an entire word-form; affixes are sub-classified in the familiar way.)

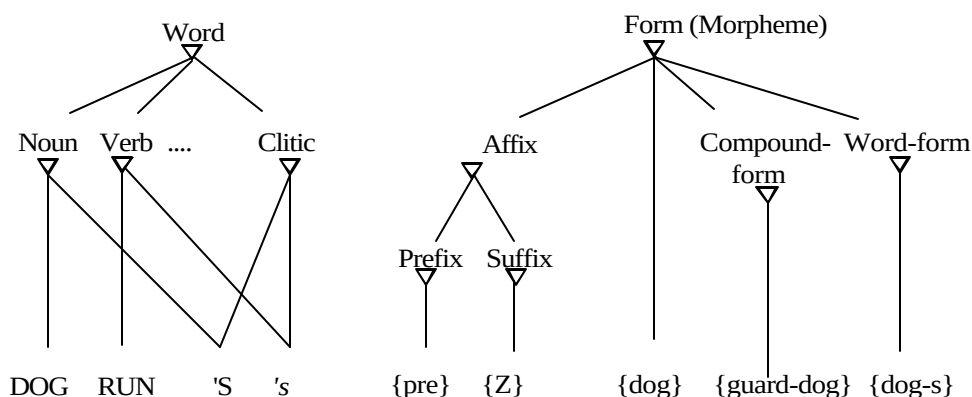


Figure 3

Figure 3 identifies 'Form' with 'Morpheme', on the assumption that the typical form is a simple content morpheme such as {dog}. This kind of analysis is possible thanks to DEFAULT INHERITANCE, where sub-cases override defaults; for example, we can maintain the generalisation that a typical form consists directly of phonemes in spite of the fact that a compound form does not. One perhaps strange consequence of this analysis is that the terms 'form' and 'morpheme' are interchangeable, so we could call {guard-dog} and {dog-s} either forms or morphemes. As explained above, I shall avoid this potentially confusing use of 'morpheme', but this leaves 'form' as the only word needed. A morpheme is now simply a default form.

Another area where I shall continue with my earlier assumptions is in defining relationships between words and their constituent forms. These relationships are called MORPHOLOGICAL FUNCTIONS (Hudson 1990:182; Hudson 2000b), because they each relate a word uniquely to one of its parts. The function 'STEM' relates it to the part which is contributed by its lexeme, and 'WHOLE' to the entire word-form, including inflections (Hudson 1990:181). The analysis of *fieldmouse* is in Figure 4, which shows the internal structure (defined as 'part1' and 'part2') of both the word *FIELDMOUSE* and the compound form {field-mouse}.

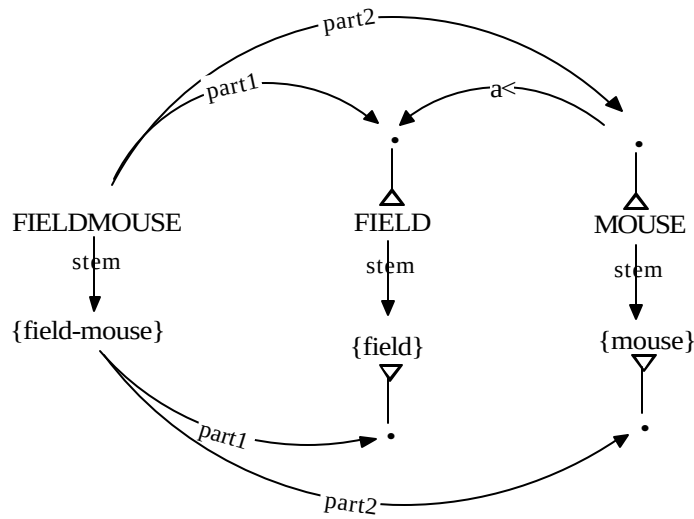


Figure 4

This analysis is firmly embedded in the ‘Word and Paradigm’ tradition of morphology in which the internal structure of words is independent of their syntactic relationships. In relation to this tradition, the only theoretical point which is at all controversial is the hierarchy of different-sized ‘forms’ which accommodates every unit recognised in a morphological analysis as well as single morphemes. There are several reasons for recognising this hierarchy.

- A formal argument rests on the WG assumption that all of grammar is declarative. In a purely declarative grammar, complex morphological forms such as Latin *am-aba-t* ‘He used to love’ cannot be defined by a series of ordered rules which gradually build the final form via a series of intermediate forms which have no special status in the grammar. Instead these intermediate forms (e.g. the imperfect base {am-aba}) have to be recognised as distinct ‘forms’. Thus the hierarchy of forms is needed to preserve the declarative format.
- Another argument involves language learning. It seems likely that when we first learn a word we may not be aware of its morphological structure, and only later do we ‘notice’ its partial similarities to other words; e.g. we record {fieldmouse} at first on a par with {elephant}, as an example of Form. Once the structure is recognised, the learner simply adds the extra links from the form {field-mouse} to {field} and {mouse} and the one to Compound-form. The analysis therefore avoids the psychologically implausible ‘unlearning’ of the initial analysis.

- It solves the problem of ‘cranberry morphs’, the phonological residue which is left over after part of the stem is recognised as a genuine morpheme: {cranberry} is a form that contains {berry} as one part, but there is no theoretical obligation to recognise the rest of it as another form {cran} as well. Similarly, we can link {dormouse} to {mouse} without positing {dor}³.

DORMOUSE is a particularly good example for this analysis, because its plural is the irregular *dormice*, which strongly supports the recognition of not just the morpheme {mouse} but even the word MOUSE. The analysis is shown in Figure 5.

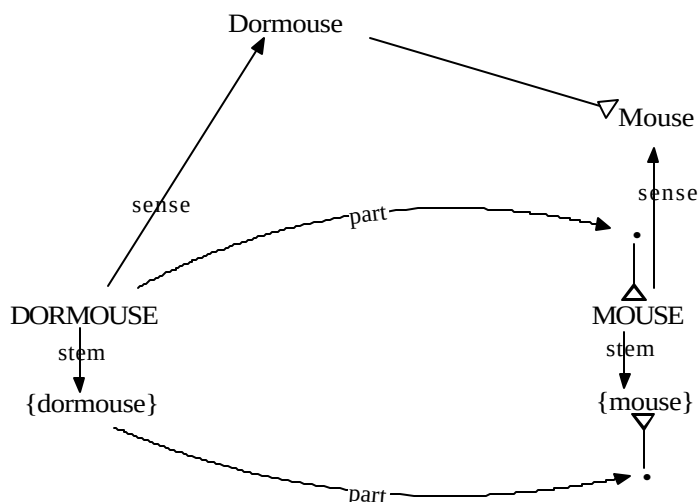


Figure 5

The outcome of this discussion is that, regardless of clitics, we have shown the need for the following kinds of units:

- compound words, which contain smaller words which are related syntactically (by ordinary syntactic dependencies);
- internally complex forms which contain smaller forms.

This is the foundation on which the following analyses of clitics rest.

³ Historically, of course, *dormouse* originated as French *dormeuse*, ‘sleeping (mouse)’, so there is no historical basis for {dor}.

4 Simple clitics and local 'host' relations in English

4.1 The possessive 'S and the definition of 'clitic'

We start with the simplest kind of clitic, in which the clitic word looks and sounds more like an affix than a word, but has ordinary syntax. Many such clitics alternate with full non-clitic forms, which raises the question of how to handle these relationships, so we start with a clitic which has no full version: the English possessive 's, as in (8) and (9).

- (8) The man's name is John.
 (9) The man in the hat's name is John.

I shall refer to this as the pronoun 'S, following the analysis justified elsewhere (Hudson 1990: 276-82); the idea is that 'S has just the same syntax as the possessive pronouns such as HIS, but it also takes an obligatory preceding dependent, which is (the head of) the possessor noun phrase. The dependency structure of (9) is shown in Figure 6, whose main point is that 's is treated as a word which has two dependents, *the* and *name*, and which is the head of the whole noun phrase *the man in the hat's name*.

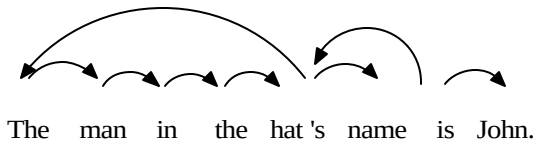


Figure 6

The clitic 'S is technically simple because its position is exactly as expected from its syntactic structure. However we shall see below that the link between the clitic and the preceding word goes beyond mere precedence.

The most obviously clitic-like characteristic of 'S is its morphology: its whole is like an affix rather than a typical word-form, which normally would consist of at least one syllable and a (potentially stressed) vowel. Indeed, better still, we can say that its form is an affix, and more precisely that it is exactly the same affix as we find suffixed to plural nouns (e.g. *songs*) and singular present-tense verbs (e.g. *sings*). In terms of morphological functions, the whole of 'S is the affix {Z}. This suggests a compact definition of a clitic:

Clitic

A clitic is a word whose whole is an affix.

These relationships are shown in the network of Figure 7.

In words, Figure 7 shows that the typical form has a part which is a stressed vowel, whereas an affix does not. (Strictly speaking it allows affixes to have stressed vowels as well, but to rule them out would involve technicalities about phonology that are irrelevant here.) The whole of a word is typically a form, but that of a clitic is just an affix. The possessive pronoun 'Z is an example of a clitic, and has the affix {z} as its whole. This little network expresses the special link between the word-class Clitic and the form-class Affix.

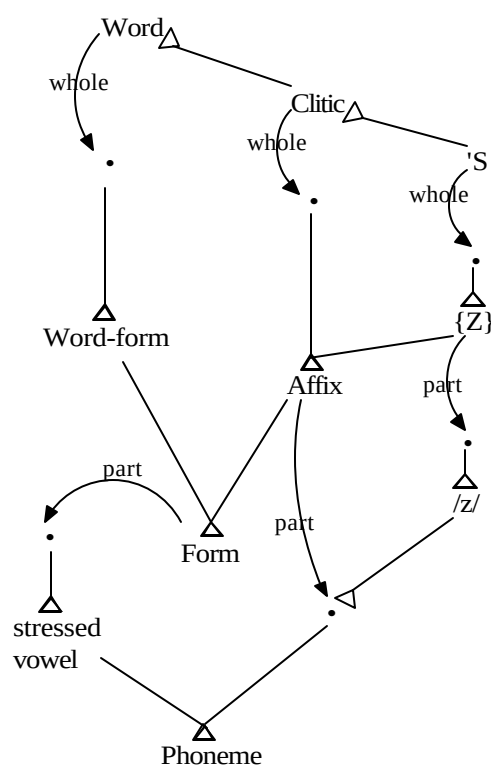


Figure 7

4.2 Hosts and hostwords

As I hinted earlier, we also need to define the relationship between a clitic and the word to which it is affixed, its 'host'. In terms of morphology, it is part of the host - in short, cliticization constitutes a major mismatch between syntactic words and morphological word-forms. There are a number of reasons for believing that the clitic is part of its host, rather than remaining as a separate word form which happens to be next to the host:

- The phonology treats it just like an affix within the host. This is particularly clear with 'Z, which behaves phonologically just like the suffix {Z} in plural nouns and singular verbs. This similarity is as expected if its whole is an ordinary example of {Z} and is therefore morphologically integrated with the host.
- In some cases there are interactions of form between the clitic and its host. This is easily illustrated from the possessive 'S because this interacts with a preceding plural suffix: *my father's mother* but *my parents' mothers*. In short, when the host already contains {Z}, the {Z} of the possessive merges with it. This would follow naturally from a general ban on two examples of the same suffix in a single word-form.
- Some of these interactions are irregular and must be stored. The irregularity is more characteristic of patterns found within words than between words, and indeed is more characteristic of inflections than of clitics (Zwicky and Pullum 1983). For example, the reduced auxiliaries of English have some irregular forms which are used only after specific pronouns (e.g. *you're*). We shall return to these and other examples below.

We must therefore assume that a clitic has an identifiable host, and that its own whole is integrated into a word-form which includes that of the host.

What exactly is the relationship between the clitic and its host? Take *my father's mother*, for example. Traditionally the host is the full word to which the clitic is attached - in this case, *father*; but in that view, the clitic is still separate from its host rather than part of it. There are two ways to integrate the clitic morphologically with the host:

- The clitic is a part of the full word's whole, so the whole of *father* is not just {father}, but {father's}, containing the forms {father} and {s}.
- The clitic and the full word are both part of a larger word, so FATHER and 'S are both parts of a larger word FATHER'S, whose whole combines their respective wholes.

It may be that each solution can be justified in some circumstances, but for the following reasons the second solution seems preferable for the data I am aware of:

- If the clitic was part of the full word's form, combinations like *children's* (as in *the children's parents*) would be problematic, because they would contain a sequence of two suffixes combined at the same level. In general, English words do not allow this - there are no words whose inflection combines two suffixes (with the possible exceptions of *isn't* and *hasn't*).

- Clitics almost always combine with the entire whole of the full word - i.e. they are outside the inflections rather than mixed up with them. For example, in French when clitic pronouns follow the verb they follow all of its inflectional suffixes (e.g. *Donnez-le-moi!*, 'Give me it!', where *donnez* contains the 2nd plural inflectional suffix {ez}). This ordering suggests that the clitics are sisters of the full word, rather than parts of it.
- The second solution parallels the structures that we saw earlier for compound words, in which the word FIELD-MOUSE contains the words FIELD and MOUSE.

In conclusion, then, a clitic and the full word on which it 'leans' are both part of a larger word; so in *my father's mother*, there is a larger word *father's* which contains two words: the full word *father* and the clitic 's.

However, if we do adopt this solution we are left with two questions:

- What kind of word is the larger word?
- Which word is the clitic's host - the full word, or the larger word which contains them both?

As in my earlier work, I believe the larger word belongs to none of the main word classes because it does not follow the rules for such words (Hudson 1990:104). Let us call it simply 'HOSTWORD', and recognise it alongside Clitic as a top-level word type. However, it is a very untypical word type because it has no meaning and unlike Clitic it need not have any lexical members - for example, the word *father's* is an example of Hostword, but is not stored. The same may be true of all hostwords, although we shall see below that some hostwords may in fact be stored.

As for the clitic's host, we have a choice between the full word (e.g. *father*) and the hostword (e.g. *father's*). As implied by its name, the hostword seems the better candidate for the following reasons:

- The clitic is always related directly to the hostword by the 'part' relationship, but it need not have any relationship other than mere adjacency to the full word (e.g. in *the man in the hat's name*, there is no dependency between *hat* and 's).
- In languages that have clitic climbing (discussed below), all clitics typically climb to the same place; this is easily explained if we require them to climb past various full words to a hostword, and only allow one hostword (of the given type) per clause.

In conclusion, then, a clitic's host is a hostword, one of whose other parts is the full word on which the clitic leans; so in the case of possessive 's, its host is the hostword which contains it and the immediately preceding word.

4.3 The analysis of possessive 'S

With these decisions in place Figure 8 presents the relevant bit of the grammar. In words, Clitic and Hostword are recognised as two special and mutually defining kinds of word. A clitic has a host, of which it is one part, alongside some full (i.e. typical, non-clitic) word. In particular, the possessive 'S is a clitic whose whole is an example of the suffix {Z}. When this grammar is applied to an example like *my father's mother*, it has no effect on word order but it has to satisfy the requirements of 'S by finding a hostword; but this in turn needs some ordinary word as its other part. The only ordinary word that can satisfy this requirement is the one immediately before 'S, *father*, so the hostword is *father's*. This is the host of the clitic 'S.

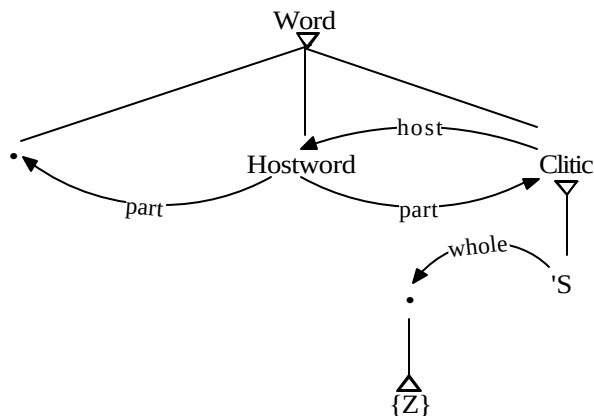


Figure 8

This analysis extends easily to explain the merger of the two {Z} suffixes in plural nouns such as *my parents' mothers*. Figure 9 recognises a special kind of hostword called 'Plural possessive' which combines a plural noun which has the suffix {Z} (unlike, say, *children*) and the possessive 'S. In this very special situation, the two words share the same suffix - i.e. the final *s* doubles up as the plural suffix and the possessive clitic⁴.

⁴ English orthography assumes a different analysis in which the final *s* is just the plural marker, and the possessive clitic is reduced to an inaudible apostrophe: *my parents' mothers*. The only virtue of this analysis

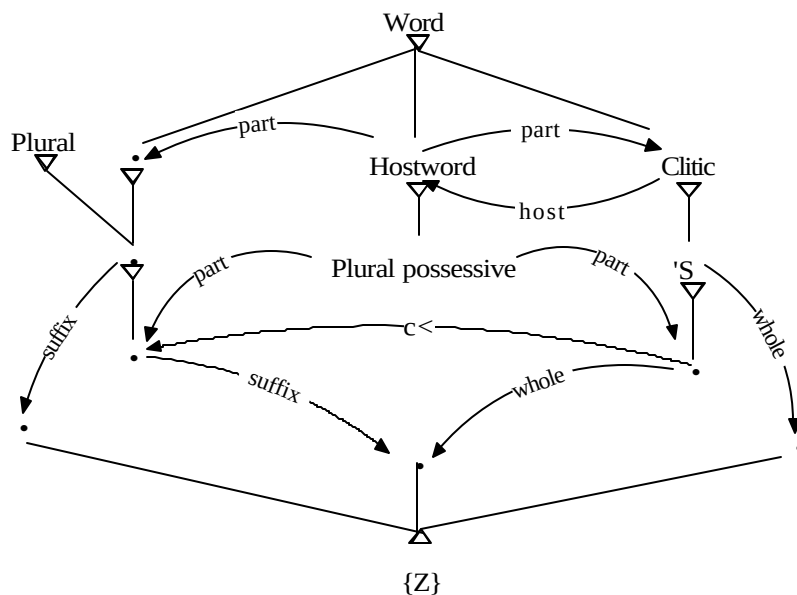


Figure 9

Figure 9 also shows an important extra detail about the possessive 'S, which is that it only merges with a plural suffix if the latter is part of the noun to which it is directly related. If the plural noun is embedded inside the possessor noun phrase, there is no merger. Consider for example what happens if we try to use *the father of my friends* as a possessor. If the {Z} of 'S merged indiscriminately with the {Z} of any plural noun, the result should be **the father of my friends' name*, but this is certainly wrong⁵. This restriction is shown by the syntactic dependency arrow labelled 'c<' (for 'preceding complement'), showing that the two words must be directly related⁶.

4.4 An extension to possessive determiners and pronouns

is that it distinguishes singular and plural. This virtue is outweighed by the perverse treatment of the possessive, which implies that it disappears completely and obscures the fact that two identical suffixes are merging into one.

⁵ It is not clear what the correct form for the possessive is in such examples; a separate {Z} also seems wrong: **the father of my friends's name*. See Zwicky 1987 for discussion of these and similar cases.

⁶ The restriction on merging the {Z}'s is problematic for the WG treatment of determiners, because the head of a phrase containing a determiner is the determiner, not the common noun. Thus in *the father's friend*, the syntactic relationship between *father* and *'s* is actually mediated by *the*. This seems to challenge the assumption that determiners are heads, and further research is called for.

The analysis of 'S can be developed in a more radical direction which exploits the merging analysis applied to the two {Z}s, and which further illustrates the lexical irregularities that may be found with clitics. The proposed analysis extends the analysis for the possessive 'S to all the more familiar possessive determiners and pronouns such as *my* and *mine* by treating them as morphologically irregular combinations of a personal pronoun with 'S; for example, *his* is a single morpheme which corresponds to the two words HIM + 'S. This is possible, even encouraged, under an analysis of clitics which recognises a hostword which combines the clitic with the word on which it leans; in short, the analysis treats *his* as a combination of three words: HIM, 'S and the hostword HIS. This is an example of a lexical item (HIS) which is listed as a hostword - a possibility which I described earlier as unusual but possible. The analysis is displayed in Figure 10. (The alternation between *my* and *mine*, and so on, can be captured in a slightly more complicated version of this diagram.)

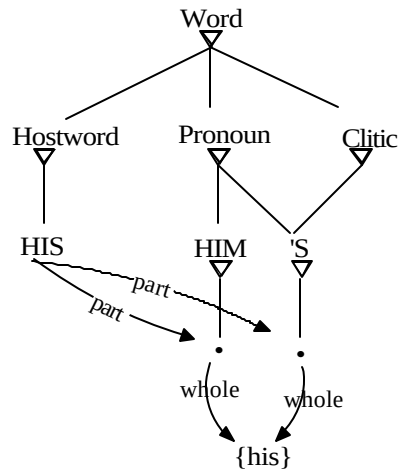


Figure 10

This analysis of possessive pronouns has the great advantage of improving the semantic analysis because it gives the same syntactic and semantic structure to *his* as it does to *John's*. This analysis shows that the possessor depends syntactically on the possessed and thereby explains the familiar facts about binding of such noun phrases:

- (10) John's mother amused herself/*himself.
- (11) John's mother amused him [= John].

These facts are as expected if the subject of *amused* is 's, and if this refers to the mother rather than to John. If we analyse *his* as HIM + 'S, the same explanation applies to examples containing *his*:

- (12) His mother amused herself/*himself.
- (13) His mother amused him [= him].

In the clitic analysis, the merger of HIM and 'S is a fact at the morphological level of much the same character as the merger of the two {Z}s.

4.5 The reduced auxiliaries

This completes the discussion of the possessive 'S, in which we have introduced almost all the apparatus needed for analysing clitics. It generalises easily to other simple clitics, such as the homophonous reduced auxiliary verb 's as in (14).

- (14) The man in the next seat 's asleep.

However, unlike the possessive 'S this form has an unreduced equivalent, the full form *is*:

- (15) The man in the next seat is asleep.

These two sentences have precisely the same syntactic structure, but there is a (well known) syntactic difference between the full and reduced forms which justifies recognising them as distinct word classes: the reduced form is only allowed if the verb 's complement follows it. If the complement is displaced either by ellipsis or by extraction, reduction is not allowed:

- (16) Mary is here and John is/* 's too.
- (17) I wonder where John is/* 's.

The same restriction applies to all reduced verbs, so we can recognise 'reduced auxiliary' as a distinct syntactic word type which combines with the individual finite auxiliaries. Thus whereas *is* is merely 'BE: present, singular' (the present singular inflection of BE), the reduced 's is 'BE: present, singular, reduced'.

Figure 11 shows the WG analysis of the reduced *is* in the context of a more general treatment of reduced auxiliaries. It omits all the details of the morphology of *is* itself, but shows that a reduced auxiliary:

- is a clitic, so it has a host which includes one other word (of any kind);
- is a tensed auxiliary, but unlike typical tensed auxiliaries (and indeed auxiliaries in general) its complement (labelled 'c') is obligatory (numbering 1) rather than optional (numbering either 0 or 1)⁷;
- has a 'source', which is some tensed auxiliary, from which it inherits all its characteristics except for the deviance of its whole.

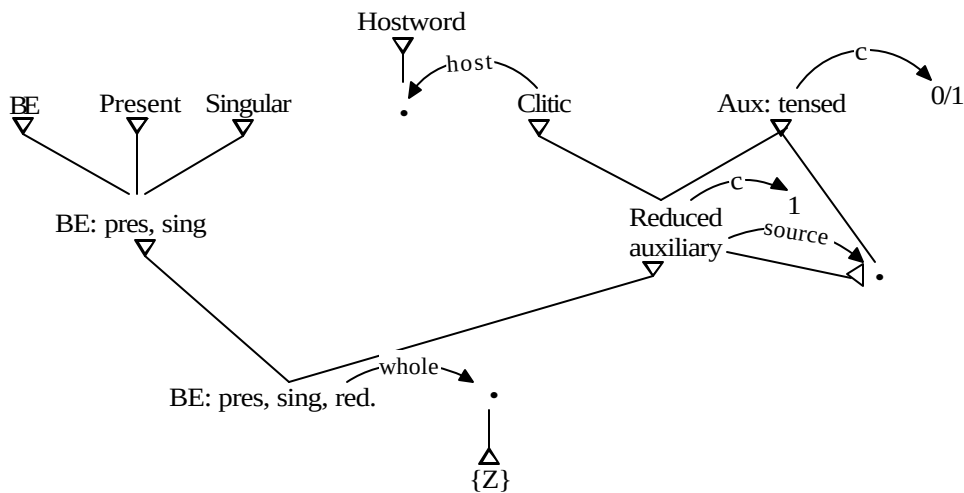


Figure 11

As with the possessive 'S, reduced auxiliaries may show irregular interactions with the word to which they are attached. For example, although the reduced form of *are* is always written 're, this spelling represents an irregular fused pronunciation when it is combined with *you*, giving the single morpheme /jɔ:/ rather than the expected /ju:ə/. Again as with possessive 'S, this reduction is sensitive to syntactic structure, so it does not apply, for example, in (18), although the normal reduced /ə/ is permitted here:

(18) The pictures of you are good.

⁷ This notation is a temporary place-holder for a proper analysis of optionality. The main point to remember is that a node's status as obligatory, optional or impossible is separate from its other characteristics.

The relevant part of the grammar is shown in Figure 12. This shows the default pronunciations for YOU and ‘BE: pres, red’ (i.e. the reduced form of *are*), but it also recognises a special version of this verb-form, called ‘BE: pres, red, you’, which combines with *you* as its subject and shares {jɔ:} as its merged whole.

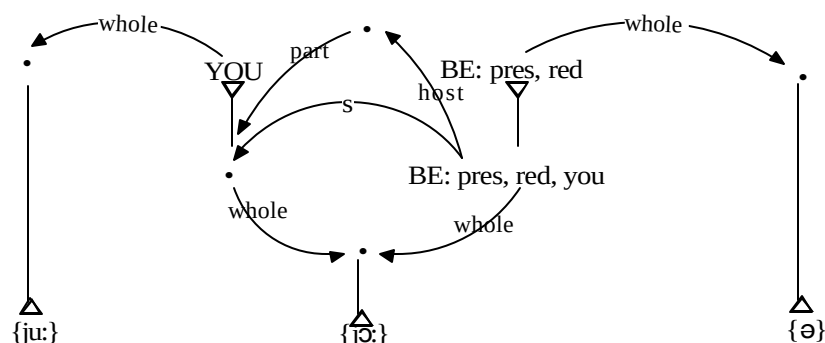


Figure 12

4.6 Summary of the treatment of simple clitics

All the examples discussed so far have been simple clitics, whose position in relation to other words is not affected by their status as clitics. In order to handle these we have introduced the following apparatus, all of which fits comfortably within the general model which we found to be necessary for handling compound words.

- Clitic, a type of word which is distinctive in having an affix as its whole. Any word which is a clitic is always listed as such (i.e. it ‘isa’ Clitic) in the grammar.
- Hostword, a special type of word which is needed by clitics and which (in English) contains a clitic and an ordinary word. Most hostwords are ad hoc combinations, but a few are stored.
- ‘host’, the relationship between a clitic and its hostword.

5 Special clitics and non-local ‘host’ relations in French

5.1 Defining the host

We can now progress to the more complicated examples where cliticization interferes with the normal rules of syntax - the so-called ‘special clitics’. In section 1 I recognised three theoretical challenges:

- How to define a clitic’s host in cases where the clitic and the host are not directly linked by a dependency.
- How to define the clitic’s position in relation to its host and to other clitic pronouns attached to the same host.
- How to explain the discontinuity (shown by the tangling dependency arrows in Figure 1) that a special position may produce.

We shall consider these challenges in turn. The empirical challenge will be the French clitic pronouns, with occasional glances at other languages.

We start with syntax - the dependency relations of clitic pronouns both to other words inside the hostword and to other words outside it. The clitic pronouns in French all attach to a locally definable verb, which need not be the word on which the clitic actually depends. For example, in (19) the pronoun attaches to the auxiliary verb rather than to the verb on which it depends as object.

- (19) Je les ai vus.
I them have seen.
‘I have seen them.’

In contrast, if the object had been a full noun such as *les enfants*, ‘the children’, it would have followed the second verb:

- (20) J’ai vu les enfants.
‘I have seen the children.’

Similarly, the pronoun EN, ‘of it/them’, must be attached to a verb even if it depends on a noun, as in the example quoted earlier:

- (21) Paul en mange deux.
‘Paul eats two of them.’

How can these rules be expressed? The case of EN is rather special so we start with the other pronouns.

In the simplest cases the pronoun attaches to the verb on which it depends. In terms of our analysis of simple clitics, this means that they are both parts of the same hostword. For example, in (22) the pronoun and verb combine to form the hostword *le connaît*, ‘knows him’.

- (22) Paul le connaît.
 Paul him knows.
 ‘Paul knows him.’

I assume that this hostword is the host not only of the pronoun but also of the verb; this assumption will help in analysing more complicated cases. Even this very simple example shows that French pronouns are (technically) special clitics, because the verb’s object is not in its expected position after the verb, as in (23).

- (23) Paul connaît Jean.
 ‘Paul knows Jean.’

In WG, languages such as English and French which have mixed word order (some dependents before the head and others after) are given two basic types of dependent: pre-dependents and post-dependent. For example, an object is an example of a post-dependent, whereas a subject is a pre-dependent. The classification is based on the dependent’s typical position rather than on its actual position; for example, a subject is still a pre-dependent of the verb even when it actually follows the verb thanks to subject-auxiliary inversion. This distinction is important for French clitics because:

- post-dependents cliticize more freely than pre-dependents, as we shall see below;
- cliticization normally overrides the typical position of post-dependents by putting them before the verb; in other words, we can see cliticization (like extraction) as a device for treating post-dependents as though they were pre-dependents.

In WG notation the dependency arrows carry labels which distinguish grammatical functions, so we can use the simple label ‘>’ (symbolising a right-pointing arrow) for the general category ‘post-dependent’, and ‘<’ for pre-dependents. The structures for the last two examples are shown in Figure 13.

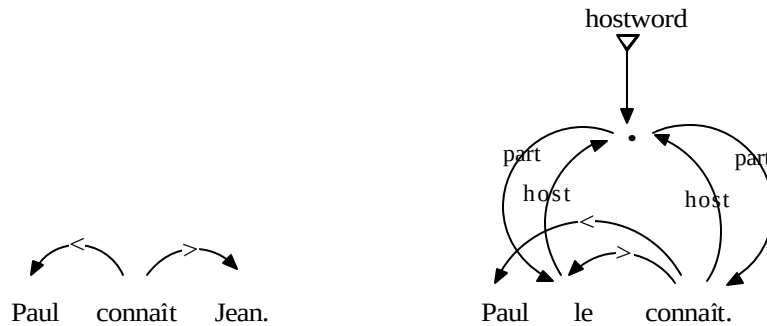


Figure 13

A grammar for such simple examples would be correspondingly simple. It would recognise a class of words called Clitic pronoun, and would attach them to the verb on which they depend as post-dependents by making them both share the same hostword. This is shown in Figure 14. The details of order within this hostword will be discussed in the next section; here we are concerned only with the rules for combining clitics and verbs.

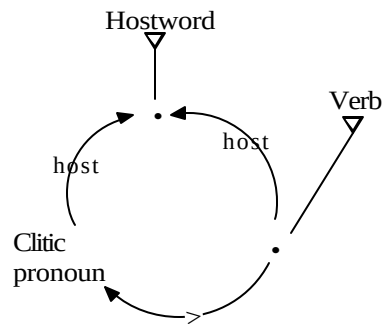


Figure 14

The extra complications which need attention are the following:

- the rule for allowing limited ‘clitic climbing’ from a full verb to an auxiliary (e.g. *l’a connu*, ‘has know him’, not **a le connu*);
- the rule for allowing more ambitious clitic climbing for the pronoun EN, as illustrated in the earlier example *en mange deux* ‘of-them eats two, i.e. eats two of them’;
- the rule which allows very simple cliticization for subject pronouns (e.g. *je le connais*, ‘I know him’).

5.1.1 *Why clitics climb up to an auxiliary.* Clitic climbing allows a clitic pronoun to ‘climb’ from the word on which it depends (as post-dependent) up to a superordinate one. In some languages all clitics can climb quite freely over certain relatively content-free verbs, as they do in the following Italian examples (from Volino 1990):

- (24) Lo ho voluto mangiare.
 it I-have wanted to-eat
 ‘I have wanted to eat it.’
- (25) Lo voglio dover mangiare.
 it I-want to-have to-eat
 ‘I want to have to eat it.’

In modern French, however, clitic climbing is very restricted for most pronouns, and only applies to participles which depend on an auxiliary (‘have’ or ‘be’):

- (26) Paul les a vus.
 Paul them has seen
 ‘Paul has seen them.’
- (27) Paul y est allé.
 Paul there is gone
 ‘Paul has gone there.’
- (28) Le priz lui a été présenté.
 the prize to-him has been presented
 ‘The prize has been presented to him.’

Unlike Italian, climbing is obligatory in these contexts.

The last example is especially interesting as it suggests a recursive process whereby the clitic climbs not only from the full verb to the nearest auxiliary, but also from the latter to the earlier auxiliary - a ‘hopping’ analysis, like the one needed for extraction. Moreover this is the kind of analysis that is certainly indicated for Italian, so I shall assume it is needed in French too. In our example, then, the pronoun *lui* climbs (or hops, according to the preferred metaphor), from *présenté* to *été*, and then from the latter up to *a*; or, to change perspective, it hops down from *a* to *été* and from the latter to the word on which it depends as a post-dependent, *présenté*.

If these two ‘hops’ are examples of the same pattern, they must illustrate the same pattern of relationships, but what is this relationship? In the analysis so far there is no direct relationship at all between a clitic and the word on which it ‘leans’ - e.g. between

lui and *a*. All that holds them together at present is the fact they are both parts of the same hostword. Let us therefore add a new relationship between these words, which we can call simply ‘CLITIC’ (abbreviated as ‘l’, to avoid confusion with the existing ‘c’ for ‘complement’). Thus *lui* is the clitic of *a*, and (assuming a recursive relationship) also of *été*.

This relationship will play an important part in the analysis, and is distinct from two other relationships to which it is closely linked:

- It is not quite the same as the relationship between two parts of a hostword. Although the clitics in a hostword (at least in French) are ‘clitics’ of the verb, a clitic may be ‘clitic’ of a verb to which it is not cliticized (e.g. *été*). This will be the proposed analysis for clitics that have ‘climbed’ up to an auxiliary verb.
- Nor does it directly give the clitic’s linear position, because a verb’s ‘clitic’ may either precede or follow it, depending on the verb’s inflectional category. In particular, positive imperative verbs take enclitics rather than proclitics (e.g. *Mangez-le!* ‘Eat it!’).

We shall conclude below that ‘clitic’ is a syntactic dependency which, like other such dependencies, allows generalisations across a range of more concrete patterns.

The diagram in Figure 14 can be revised to show this new relationship (‘l’ for ‘clitic’) instead of the post-dependency (‘>’), giving Figure 14’. (The latter relationship is still relevant, of course, as we shall see in Figure 16 below.)

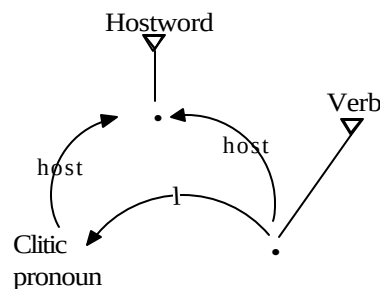


Figure 14'

In our example, *lui* is the ‘clitic’ of *a*; so, given the recursivity of this relationship, it must also be the clitic of *été*. Notice that in the latter case, being ‘clitic’ of a word does not necessarily mean being part of the same hostword - i.e. in traditional terms, being cliticized to it. The analysis of this example is now as shown in Figure 15.

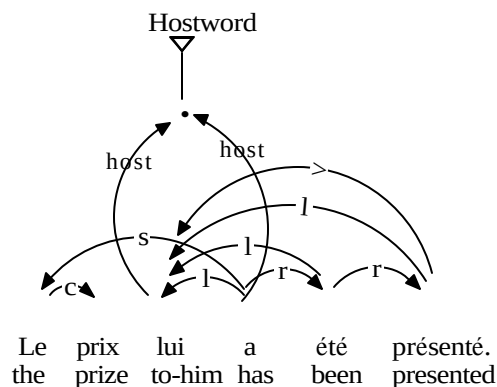


Figure 15

In words, what Figure 15 shows is that *lui* has the same relationship, ‘clitic’, to all three verbs, and it is this shared relationship that allows it to ‘climb’ back along the dependency chain to the first verb. In formal terms this is just the same kind of recursion that we find in other kinds of structure-sharing such as subject-raising and extraction. For instance, in (29) *it* is the subject of all three verbs thanks to a recursive subject-raising pattern which gives two verbs the same subject if the second is the ‘sharer’ of the first (Hudson 1990:117, 235-9).

(29) It has been raining.

The formal rule allows the specified relationship - in our case, ‘clitic’ - to link word A to both B and C, provided that C is a post-dependent of B. The basic rule for clitic climbing is shown in Figure 16.

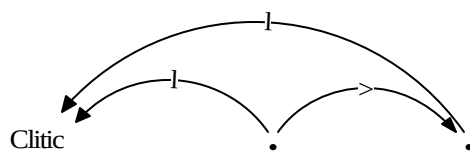


Figure 16

It should be clear how this simple pattern licenses all the ‘clitic’ links in Figure 15, because the clitic is in the same relationship to each of the successive pairs of verbs. However we also need to add the more basic pattern in which the clitic does not climb, but is a post-dependent of the word of which it is ‘clitic’. This extra possibility is shown in Figure 17, in which the recursive pattern is shown as a special case of the simpler one.

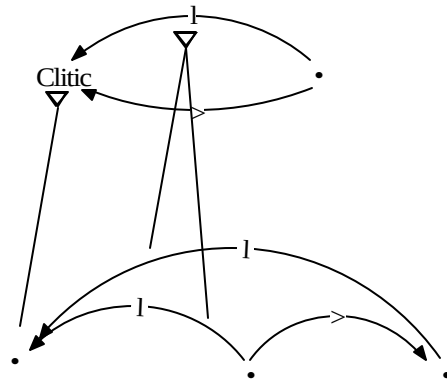


Figure 17

These structure diagrams can be simplified by selecting just one of the dependencies for each word as the only one which is relevant to word order, in contrast with all the other ‘extra’ dependencies that are irrelevant to surface position. This is called the principle of Surface Structure (Hudson 2000c) and allows the simple generalisation that word order restrictions apply to a sub-set of the total structure which contains just one dependency per word. This principle will be relevant when we discuss discontinuities that are produced by cliticization in a later section. On a more practical level, it allows us to improve structure diagrams by showing only the surface structures above the words; this is the structure which is formally equivalent to a phrase structure tree. Any additional dependencies are shown below the words, and are allowed to produce any amount of complexity.

For example, however complex Figure 15 above may look, it is still not complete because it does not show the structure-sharing of the subject; but the two-tier notation makes such structures relatively comprehensible. Figure 18 shows a complete analysis for the same sentence. The complex lines below the words are due to two very similar recursive patterns that share the subject and the clitic across all three verbs.

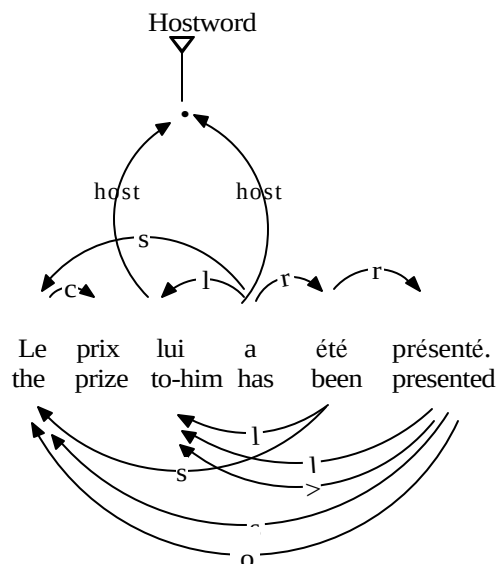


Figure 18

Having established that there is a direct relationship between the clitic and the verb, and having named it ‘clitic’, we must decide what kind of relationship it is. Is it a syntactic dependency, like ‘subject’, or is it some other kind of relationship, like ‘host’? The fact that it relates two words to each other suggests strongly that it is a dependency, though there are word-word relationships which are not dependencies (e.g. that between a pronoun and its antecedent). However the similarity between the two recursive patterns just discussed also points in the same direction; since ‘clitic’ parallels ‘subject’ in these two patterns, it is reasonable to assume that it too is a dependency. Indeed, we shall see in 5.3 that ‘clitic’ is in fact similar to no fewer than three clearly syntactic dependencies. This admittedly weak evidence suggests, therefore, that French has the syntactic dependencies shown in Figure 19 (among others). If true, this conclusion applies only to French, as there is no reason to think that ‘clitic’ is part of English grammar.

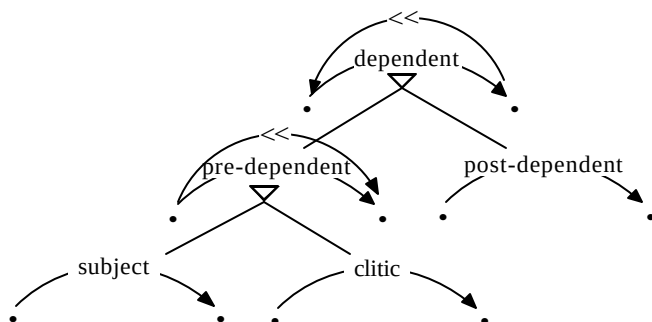


Figure 19

Figure 19 also shows how word order (and more generally all temporal ordering) is shown in WG, given that the grammar takes the form of an order-neutral network. If two nodes A and B are ordered in time, an arc labelled '<<' (meaning 'precedes') connects the two, and points at whichever comes first. Figure 19 shows that by default a word precedes any of its dependents (i.e. French, like English, basically head-initial), but this default is overridden for pre-dependents.

To summarise the analysis so far, the pattern in Figure 17 allows clitic climbing by allowing a clitic to be passed up (or down) a chain of verbs linked by dependency. This is what is needed in order to allow clitics to climb to an auxiliary verb. What it does not yet do is to require them to climb - very important, because (unlike Italian) this kind of clitic climbing is obligatory. To achieve this we can recognise a special kind of verb which combines the following characteristics:

- it is a participle;
- it depends on an auxiliary verb;
- it has a clitic as dependent (i.e. in the 'clitic' relationship).

There is no obvious traditional name for this kind of verb, so we can call them simply PACs (for Participle with an Auxiliary and a Clitic). In the grammar we simply record all these characteristics, plus the requirement that the clitic is structure-shared with the auxiliary. The relevant entry is in Figure 20 (where 'r' stands for the 'sharer' relationship between an auxiliary and the next verb which forces subject- 'sharing').

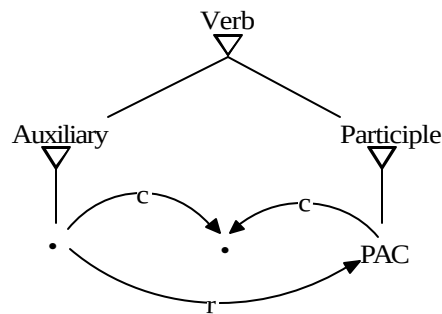


Figure 20

5.1.2 Why EN can climb further. One of the clearest pieces of evidence that clitic pronouns are separate elements of syntax rather than mere inflectional affixes is that the

pronoun EN, ‘of it/them’, can climb up to the verb out of a deeply embedded phrase, as in (30)⁸.

- (30) Paul en sait le nom de l’auteur.
 Paul of-it knows the name of the author
 ‘Paul knows the name of the author of it.’

As expected, EN also climbs to an auxiliary according to the usual rules:

- (31) Paul en a découvert le nom de l’auteur.
 Paul of-it has discovered the name of the author
 ‘Paul has discovered the name of the author of it.’

Why is EN such a good climber compared with the other clitics?

The obvious answer is that the other clitics are genuine pronouns, so they cannot be post-dependents of anything but a verb; in contrast, EN is really a prepositional phrase with much more flexible distribution⁹. The discussion of simple clitics in English (Section 3) suggests an analysis of EN in which it is a single hostword containing two separate words, DE (‘of, from’) and a third person pronoun. (Compare the analysis of HIS in 3.4 as HIM + ’S, and even more relevantly, the analysis of *du* as DE + LE in *du livre*, ‘of the book’.) As a preposition, DE can depend on all sorts of words including nouns and adjectives, hence its versatility as a dependent. In short, the freedom of EN does not need any explanation. It has the same distribution as phrases headed by DE, but, being a clitic, it benefits from the recursive climbing allowed by Figure 17. The structure for sentence (31) is as shown in Figure 21. The main point of the example is the series of repeated ‘l’ (for ‘clitic’) links which bridge the gap between *en* and *auteur* and allow an eventual ‘c’ (for ‘complement’) link between them.

⁸ Thanks to Paul Rowlett for French data used in this section.

⁹ EN is also a noun phrase meaning ‘some of it/them’, as in *Paul en mange*, ‘Paul eats some of it.’ However this is not a peculiarity of EN; the same is true generally of any phrase introduced by DE, such as *de la viande*, ‘some meat’. Any such phrase can be used in places where a noun phrase is expected, as in *Paul mange de la viande*, ‘Paul eats some meat.’ Since DE is clearly the head of the phrase, and DE is (normally) a preposition, this well-known fact is very strange. It is as much of a challenge for WG as for any other theory of grammar and I have no explanation for it.

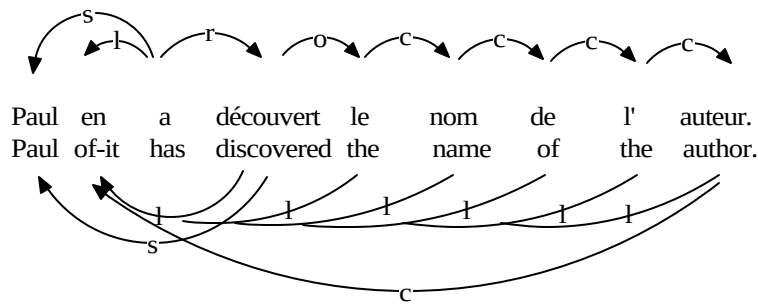


Figure 21

Nor do we need an explanation for why *en* does not climb easily out of subject noun phrases. For example, (32) is not possible.

- (32) *Le patron en a téléphoné hier.
 The boss of-it has phoned yesterday.
 'The boss of it (e.g. the firm) phoned yesterday.'

This follows automatically from the restriction on the non-clitic relationship in Figure 17 which requires it to be a post-dependent. An ordinary subject such as *le (patron)* in this example is not a post-dependent but a pre-dependent, so there is no reason why a clitic should be able to climb across it. Better still, the analysis in Figure 17 explains why EN **can** climb out of the subject of a passive verb¹⁰ as in the following example (Grévisse 1986: 1032).

- (33) La ville ayant décidé de donner à une rue le nom des deux frères, la nouvelle plaque en fut inaugurée ... par ...
 'The town having decided to give a street the name of the two brothers, the new plaque of it (i.e. of the street) was inaugurated ... by ...

This is possible because, according to the WG analysis (Hudson 1990:336-53), the subject of a passive verb is also its object (as it would be without passivization). Thus in example (33) *la (nouvelle plaque)* is not only the subject of *fut* and *inaugurée*, but also

¹⁰ In a language such as Italian in which unaccusative and unergative verbs are clearly different in syntax, unaccusative verbs can be given the same syntax as passive verbs, with the subject doubling up as the object. This explains why NE (the Italian equivalent of French EN) can climb out of the subject of an unaccusative verb. It is unclear (to me) whether this distinction applies in French, and I am fairly sure that it does not apply to English.

object of *inaugurée*, so in spite of the word order there is a chain of post-dependents between *plaque* and *en*:

- *fut* to *inaugurée* (sharer, a post-dependent)
- *inaugurée* to *la* (object, a post-dependent)
- *la* to *plaque* (complement, a post-dependent)
- *plaque* to *en* (post-adjunct (‘>a’), a post-dependent).

The analysis of the last clause is shown in Figure 22.

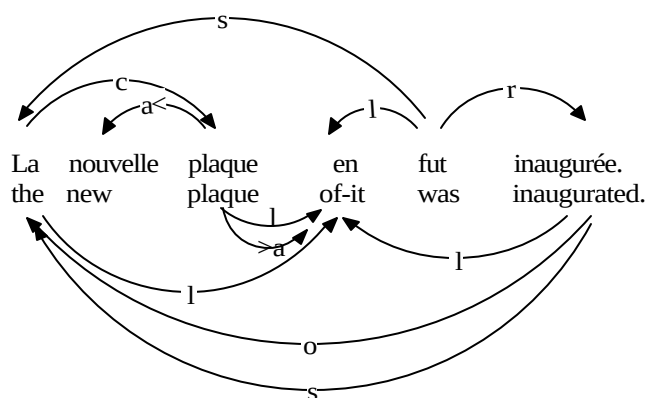


Figure 22

As far as EN is concerned, therefore, its relative freedom is already explained by the grammar given already. In contrast, the restrictions on other clitic pronouns almost mirror those on pronouns in general. Unlike prepositions, they can only carry dependencies which are available for nouns, which means in principle that they can only depend on verbs and prepositions. This immediately explains why they are less free than EN, but it does not explain why they cannot climb up from a prepositional phrase. The grammar given so far wrongly allows examples like (34).

- (34) *Paul la travaille avec.
 Paul her works with.
 ‘Paul works with her.’

The only way to express this idea in French is to use a non-clitic pronoun:

- (35) Paul travaille avec elle.

French might have allowed clitic pronouns to attach directly to prepositions, as they can in Arabic (for example); or it might have allowed them to climb up from a preposition to a verb. In fact, neither of these possibilities is allowed so they must be excluded by stipulation.

The facts are simple: object pronouns (which in this case include *Y*, ‘to/at it/them’) can only depend on verbs, whether as post-dependents or as clitics. In other words, they can only climb around verbs, whether from them, via them or to them. In contrast, *EN* is restricted in this way only as far as the target of climbing is concerned: it can climb from any word which will accept *DE* as post-dependent, and via any kind of word (including determiners/pronouns such as *la* in Figure 22). These facts are easily added to the grammar in Figure 17, giving the one in Figure 23.

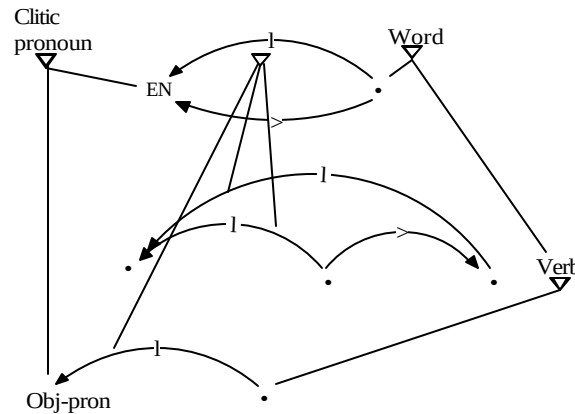


Figure 23

Figure 23 is just the same as Figure 17 with the following differences:

- Clitic pronouns are now divided into two: *EN* and the object pronouns.
- The category ‘clitic pronoun’ has now been moved out of the diagrams so that it merely brings together the two kinds of clitic. (Its main work is in Figure 14’, which shows how clitics cliticize.)
- *EN* is given the default pattern which was shown in Figure 17, which relates the clitic either directly or indirectly to the word of which it is a post-dependent.
- The default pattern allows *EN* to be clitic of any kind of word.
- Object pronouns inherit all the patterns allowed for *EN* except that they can only be the clitic of a verb.

5.1.3 *Subject clitic pronouns.* French (unlike Italian) has clitic pronouns which act as subject of a tensed verb. These are obligatory in the absence of a full pronoun or noun phrase as subject.

- (36) Il dormait.
 he was-sleeping
 ‘He was sleeping.’

Their syntax is much simpler than that of object pronouns and EN because they do not climb at all, at least not as clitics. As subjects, they are of course eligible for the subject-raising illustrated in (29). Thus in (37) *il*, ‘he’, is the subject of all three verbs, but this is not because it is a clitic but because it is a subject.

- (37) Il a été élu.
 he has been elected.
 ‘He was elected.’

The only syntactic difference between clitic subjects and non-clitics is in subject-auxiliary inversion, which is possible only for clitics (in contrast with English and German). This is illustrated by the difference between the following pairs.

- (38) Il finit bientôt.
 he finishes soon
 (39) Finit-il bientôt?
 ‘Does he finish soon?’
 (40) Paul finit bientôt.
 ‘Paul finishes soon.’
 (41) *Finit Paul bientôt?

However this peculiarity of clitics has nothing to do with climbing. It can easily be handled in the rules for ordering clitics and verbs, and fits comfortably alongside the other mood-based peculiarity of imperative verbs whose clitic pronouns follow them (e.g. *Finis le!* ‘Finish it!’).

The grammar given above already prevents subject clitics from climbing simply by allowing climbing only for EN and object pronouns. All we need to do is to expand the sub-classes of clitic pronoun by adding subject pronouns, leaving the rest of the

grammar untouched. For convenience we can show these sub-classes as an extension of the basic cliticization pattern given in Figure 14'. The result is Figure 24.

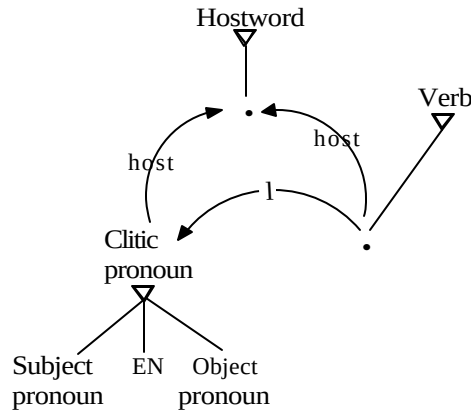


Figure 24

In conclusion, we now have a grammar for French clitic pronouns which handles the syntax of individual clitic pronouns. It does this by defining two words: the word they are cliticized to and the one they depend on as post-dependents (or as subjects). In simple cases these words are the same, but thanks to climbing they may be different. The main elements in this grammar are:

- three distinct sub-classes of clitic pronoun: subject pronoun, object pronoun and EN;
- the dependency 'clitic' which links a clitic pronoun to any relevant word: the one it is cliticized to, the one it is post-dependent of, and any word in the dependency chain between these two;
- the schematic pattern in Figure 24 which handles the cliticization;
- the schematic pattern in Figure 23 which handles climbing.

What I have not tried to do so far is to explain ordering within the hostword - i.e. the position of each clitic in relation to the verb or in relation to other clitics. This is the target for the next section.

5.2 Defining position within the host

The next challenge is to deal with the morphology of the hostword on the basis of the syntactic structures discussed in 4.1. Given a hostword that consists of a verb and one or more clitics, what order can these clitics occur in? The ordering of clitics is extremely

rigid, and is much more characteristic of the rules for ordering inflectional affixes than of those for ordering words in a sentence. Whereas free word order is very common, and perhaps even normal, free affix order is rarely, if ever, found. Similarly, the order of clitics is always rigidly fixed, with very minor and local possibilities of variation. This is why it is best to deal with clitic order as an example of morphological structure.

For French the facts are complex but clear (though there is some marginal variation according to style - Hawkins and Towell 2001:69).

- Most verbs follow all their clitics, but positive imperative verbs precede them.

(42) Vous les mangez.
you them eat.
'You eat them.'

(43) Mangez-les!
'Eat them!'

- When a verb has more than one preceding clitic, they combine in the following order:
 - subject pronoun
 - negative marker NE
 - object pronouns
 - EN

(44) Je ne les en tire pas.
I not them from-it pull not.
'I don't pull them from it.'

- Object pronouns occur in the following order when they precede the verb:
 - ME, TE, SE, NOUS, VOUS (first and second person, third person reflexive)
 - LE, LA, LES (third person direct: 'him, her, them')
 - LUI, LEUR (third person indirect: 'to him, to her, to them')
 - Y (a preposition phrase like EN, but containing the preposition À, 'to, at': 'to/at it/them' or simply 'there')

(45) Paul nous le donne.
Paul to-us it gives
'Paul gives us it.' (or 'Paul gives it us.')

- (46) Paul le lui donne.
 Paul it to-her gives
 'Paul gives her it.' (or 'Paul gives it her.')
- (47) Paul m'y rencontre.
 Paul me there meets
 'Paul meets me there.'

- If object pronouns follow the verb, they occur in the same order as above except that the order of the first two pronoun classes is reversed.

- (48) Donnez-le-nous
 give it to-us
 'Give us it!' (or 'Give it us!')

- The combination of clitic pronouns is further restricted in a more or less arbitrary way:
 - Only one pronoun from each class may be used, even where a combination of two would be convenient semantically. For example, members of the ME list can be used either as direct objects or as indirect objects, but we cannot use one as direct object and also another as indirect object:

- (49) Je me présente à Paul.
 'I present myself to Paul.'
- (50) Je te présente Paul.
 'I present Paul to you.'
- (51) *Je me te présente.
 'I present myself to you.'

- The first and third pronoun classes also cannot combine, suggesting that they too form a single class in spite of their positional difference.

- (52) *Je me lui présente.
 'I present myself to him.'

The facts just listed are frustratingly arbitrary when considered from the point of view of syntax and semantics because clitics are classified in ways that conflict with

familiar syntactic and semantic characteristics. For example, the first class of object pronouns includes:

- one third-person pronoun which must be reflexive: SE;
- four first- and second-person pronouns which may be either reflexive or not: ME, TE, NOUS, VOUS. Thus ME translates into English as either ‘me’ or ‘myself’.

Moreover, all these pronouns may be used either as direct objects or as indirect objects; but third-person non-reflexive pronouns are divided between the next two pronoun classes. In short, there is no single syntactic principle, such as person, reflexivity or grammatical function, which explains the order classes. Nor is there a syntactic explanation for the gaps illustrated by (51) and (52). The syntactic and semantic functions of the pronouns are irrelevant; all that counts is the class to which each pronoun has been assigned. All this is strongly reminiscent of morphology, where a single morpheme may have many different syntactic and semantic functions; to take an example discussed earlier, the morpheme {Z} is found in plural nouns, singular verbs, and possessive constructions.

For these reasons we shall handle all these facts as part of the morphology of the hostword which consists of the verb and all the clitics which lean on it; for instance, in (45), *Paul nous le donne*, ‘Paul gives us it.’, there is a hostword *nous le donne* whose internal structure will be controlled by morphological rules. The challenge is to formulate these rules. Let us start by considering the overall morphological structure of a French hostword using the categories developed during the previous sections.

- It has a **whole**, which is an entire word-form containing all the parts.
- It has a **stem**, which is the verb - *donne*.
- It has a number of other **parts** which are the various clitics, and which can be distinguished as distinct morphological functions each of which identifies a ‘positional slot’ which may be filled by a clitic. For example, we can easily distinguish the following functions:
 - ‘s’ - the subject clitic
 - ‘n’ - the negative clitic
 - ‘o’ - object pronouns other than Y
 - ‘y’ - Y
 - ‘e’ - EN

The main complications involve the object pronouns, so we leave these details till later. The other elements can be combined quite easily into the grammar for French hostwords shown in Figure 25.

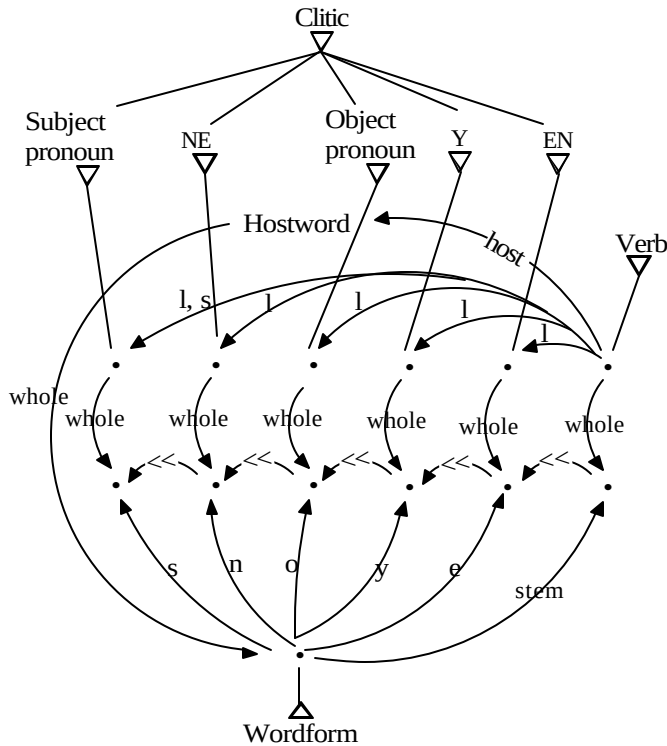


Figure 25

For all its apparent complexity, Figure 25 encodes the facts very directly. It says that a hostword consists of (its whole is) a word-form with up to six parts: a stem and five other parts. The stem is the whole of the verb, while the other parts are the wholes of its various clitics. They occur in the order shown, as indicated by the arrows labelled '<<', and they belong to the classes shown. As there is only one morphological function for each class, only one member of each class is possible. This is correct in general, but is clearly wrong for the object pronouns, to which we now turn. We shall also need to modify some of the details for imperative verbs. As the rules are all morphological rules it will be helpful to recognise the whole of a hostword as a special kind of word-form, a 'CLITIC-CLUSTER'.

For clitic-clusters containing object pronouns we have to handle the following facts:

- There are three sub-classes of pronoun (excluding Y) which we can call Me, Le and Lui after their first member in the earlier lists.
- These occur in the order Me << Le << Lui with most verbs, but Le << Me << Lui with positive imperatives.
- Although Me and Lui occupy distinct positions, they belong to a single co-occurrence class from which only one example may be chosen. There are thus just two co-occurrence classes:
 - Me/Lui (ME, TE, SE, NOUS, VOUS, LUI, LEUR)
 - Le (LE, LA, LES)

The challenge is therefore to recognise three distinct positions but only two co-occurrence classes.

The solution is to distinguish just two morphological functions and two main pronoun classes, and then to split each of them into two by recognising a special position-based sub-function and sub-class. We can call the two main functions '123' (for the class Me/lui, which covers all three persons) and 'd' ('direct', for Le, which contains just third-person direct objects). The special sub-case of '123' is 'i' (for Lui, which is always an indirect object). This analysis is shown in Figure 26, which is intended to replace the area of Figure 25 which deals with object pronouns.

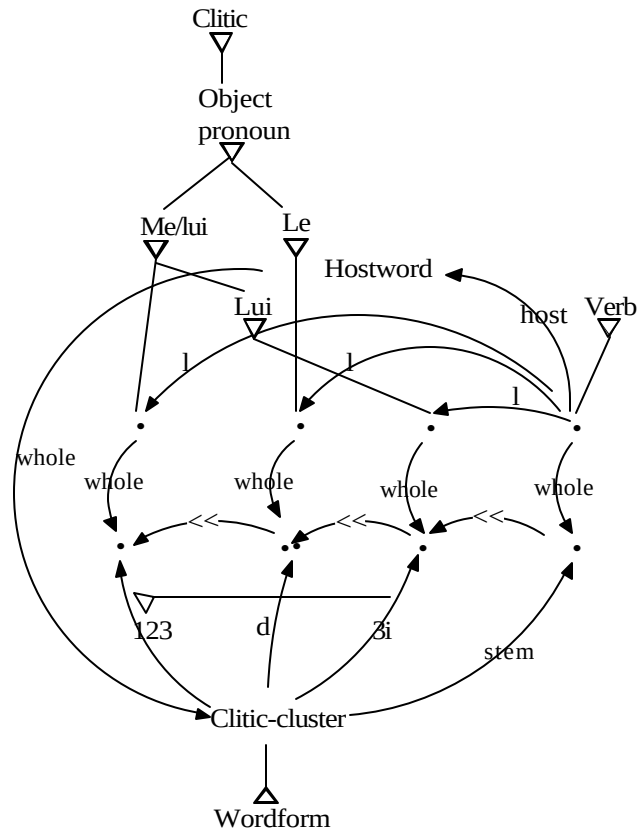


Figure 26

I have to admit that this analysis involves a small amount of theoretical hand-waving because the relevant area of WG theory is still under construction; this is the area hinted at by the numbers '0' and '1' in Figure 11 and footnote 8. The idea to be expressed is that pronoun combinations such as *me lui* are not permitted because they would both stand in the same relation to the word-form: '123', even though one of them involves a special sub-case of this relation. The trouble is that nothing in this diagram actually limits the number of examples of each relationship; all this grammar does is to limit the classes of words and the order in which they occur. In order to add numerical limits we have to add numerical information to the network, such that '123' cannot occur more than once for any given word-form. The system which I described in 1990 did incorporate this kind of information (Hudson 1990:20) but other changes have made this part of that system irrelevant and no proper alternative has been developed as yet. We must therefore treat Figure 26 as a place-holder for a complete analysis.

Finally we have the special case of imperative verbs, whose clitics follow them.

(53) *Donnez-le-nous!*
 ‘Give it us!’

(54) *Donnez-le-lui!*
 ‘Give it him!’

(55) *Donnez-m ‘-en*¹¹!
 ‘Give me some!’

The differences in comparison with the default order are:

- The verb precedes the first clitic rather than following the last one.
- The order of *Me* and *Le* is reversed, as in (53), to give *Le* << *Me* instead of *Me* << *Le*. Since *Lui* also follows *Le*, *Lui* and *Me* now occupy a single position so we no longer need to subdivide the super-class *Me/lui*.

There is no need to pay attention to subject clitics and *NE* because neither of them can occur with a positive imperative verb, nor need we mention *Y* and *EN* because their positions, relative to the other clitics, stay the same. All we need, therefore, is a pair of special ordering rules which override the default orders for ‘123’ and ‘d’, and for ‘d’ and the ‘stem’ verb. These are shown in the middle of Figure 27, which recognises ‘Imperative clitic-cluster’ as a sub-type of clitic-cluster which automatically inherits all the latter’s characteristics except for those which are shown here.

¹¹ As noted in footnote 3, *EN* may be used as a noun phrase to mean ‘some of it’; compare *Donnez-moi de la viande!* ‘Give me some meat!’.

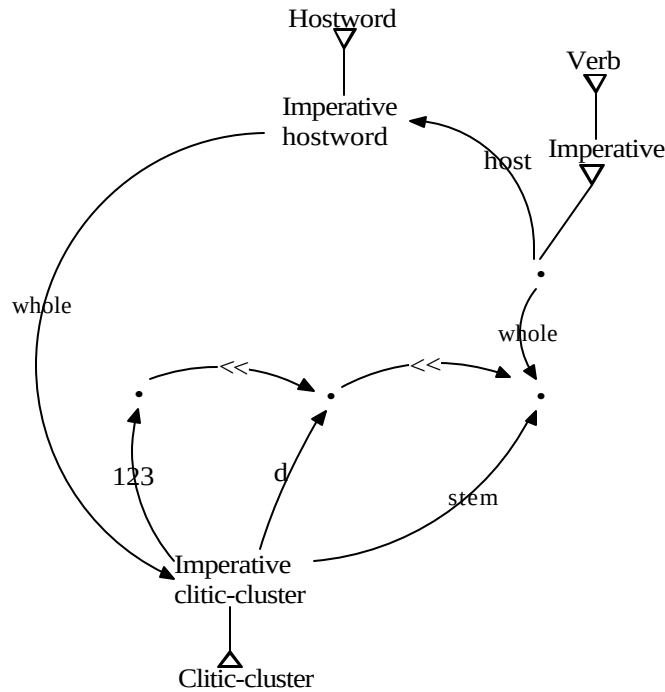


Figure 27

In conclusion, we now have a grammar for French clitic pronouns which handles all the complex facts reviewed above. There are other complexities which we have ignored, of course; in particular, we have ignored the morpho-phonemic variation in the forms of the pronouns whereby ME, for instance, may appear as *m'* or as *moi*, as below:

- (56) Paul m'aide.
 'Paul helps me.'
 (57) Aidez-moi!
 'Help me!'

The grammar involved the following elements which are therefore available, where needed, for the treatment of clitics in other languages.

- Hostword, a special kind of word whose parts are also words, and which (as we have just seen) may itself be sub-classified for special patterns.
- Clitic-cluster, a special kind of word-form which serves as the whole of a hostword and whose parts are other word-forms; like Hostword, it may be sub-classified.
- Clitic, a special kind of word which occurs as part of a hostword.

- ‘Clitic’, the syntactic dependency between a clitic and a full word.

These are the special categories that are needed only for the sake of clitics. They combine with the categories that are needed for ordinary words - ordinary word classes, syntactic dependencies, forms and functions.

5.3 Explaining discontinuities and deviant word order

The last general peculiarity of clitics is that their position may conflict with the most general syntactic principles of word order as well as with more specific rules. The example that I gave earlier is (58).

- (58) Paul en mange deux.
 Paul of-them eats two.
 ‘Paul eats two of them.’

The general problem is that *en* has climbed out of the phrase headed by *deux*, and is separated from the rest of that phrase by *mange*. This kind of discontinuity is forbidden by one of the most general principles of syntax which can be expressed in various ways; for example, in an analysis which uses arcs to show dependencies we can ban arcs that tangle (Hudson 1998: 20). If we ignore the fact that *en* is a clitic, as in Figure 28, this example clearly defies the ban.

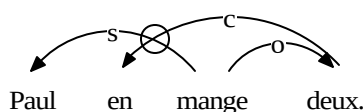


Figure 28

The discontinuity problem is solved in just the same way as other kinds of discontinuity: by postulating an additional syntactic dependency which satisfies the demands of surface structure explained above in connection with Figure 18 (Hudson 2000c). Subject raising, extraction and extraposition all produce a discontinuous phrase which is ‘rescued’ by a dependency which attaches the detached part to a higher verb. For example:

- Subject-raising: in (59), *John* is attached (as subject) to *kept*:

(59) *John kept talking.*

- Extraction: in (60), *who* is attached (as extractee, labelled 'x<') to *do*, *think* and *invited*:

(60) *Who do you think she invited?*

- Extraposition: in (61), *who* is attached (as extraposee, labelled '>x') to *died*:

(61) *Two people died who were riding in the front coach.*

Similarly in (58), *en* depends on *mange* by virtue of the dependency which we called 'clitic'. These dependencies are in addition to those which link the separated word to the rest of the discontinuous phrase. The similarities between the four structures can be seen in Figure 29. The solution is a purely syntactic one in the first three cases, so the same must be true for clitics as well.

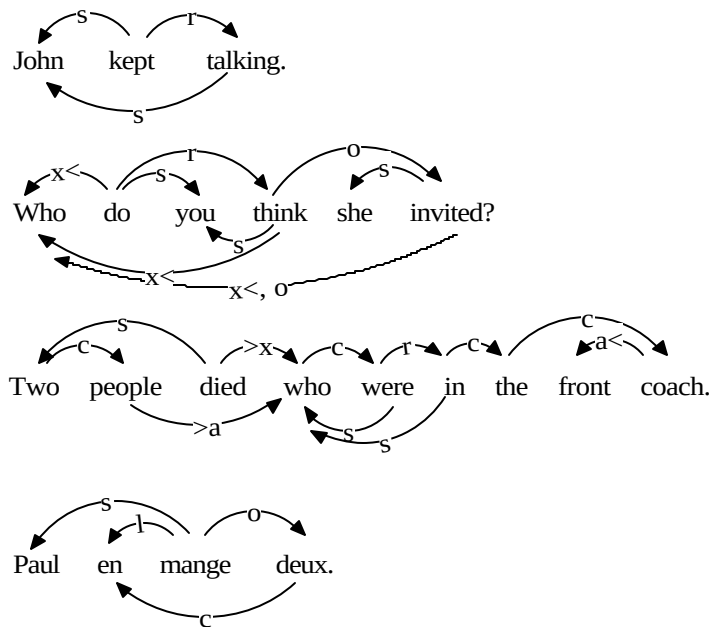


Figure 29

On the other hand, clitics also organise themselves quite differently from ordinary words, so their position both relative to one another and relative to the verb can conflict with the rules of syntax. Here are some examples of precedence rules without purely syntactic precedents that I have already mentioned:

- Most clitics are post-dependents, and yet they precede the verb on which they depend.
- One inflectional category (imperative verbs) are an exception to this rule.
- The relative order of direct and indirect objects depends on their person (e.g. *me le*, 'to-me it' but *le lui* 'it to-him').
- The relative order of all clitics is rigidly fixed, in contrast with the generally flexible order of ordinary words.

The question is not how to formulate these rules - I have already answered this question. Rather, the question is why the rules for clitics should take precedence over the more general ones for syntax. However natural this priority might seem, we still need a formal explanation for it.

There are two possible explanations.

- Default inheritance automatically gives priority to statements about more specific categories. 'Clitic' is more specific than any of the word classes mentioned in syntax because every clitic is defined as a combination of some word class with Clitic. Therefore any rule for clitics will override those for non-clitics.
- Rules for ordering parts always take priority over those for non-parts. Recall that in WG syntax relates words directly to one another without postulating phrases, so words are not parts of anything larger. In contrast, affixes are parts of words, and so are clitics. This principle probably applies outside language - for example, if we know that a screw is part of a machine we do not look for it in the screw drawer - but it is almost certainly a universal of language (Sadock 1991).

It seems pointless at this point to choose between these explanations since they both involve extremely general principles with ample independent justification.

6 A glance at Serbo-Croat

In section 1 I mentioned the following data from Serbo-Croat:

- (61) a. Ja mu ga dajem svaki dan.
 I him it I-give every day
 'I give it to him every day.'
- b. Svaki dan mu ga dajem.

c. Dajem mu ga svaki dan.

The generalisation to be captured is that the clitics cluster in ‘second position’ - a typical Wackernagel pattern.

More precisely, this restriction is defined as “after the first accented constituent (e.g. NP) or after the first accented word” (Spencer 1991:355). For argument’s sake I shall assume that this description means the following:

- Clitics are always associated, in terms of position, with a finite verb; so ‘second position’ means ‘second position within a finite clause’. Let us call this verb simply ‘the finite verb’. In (61) the finite verb is *dajem*, ‘I give’, but it may be an auxiliary as in (62) (ibid:352).

(62) Pisao sam pismo.
written am letter
‘I wrote a letter.’

- The first position implied by the second position is defined in relation to the finite verb: it is:
 - the finite verb itself as in (61c);
 - (if the finite verb is an auxiliary) the full verb that depends on it as complement as in (62);
 - a phrase headed by any dependent of either of these verbs as in (61a,b);
 - a part of such a phrase, as in (63) (ibid:355).

(63) Taj mi je pesnik napisao knjigu.
that to-me is poet written book
‘That poet wrote me a book.’

In short, the first position contains either the finite verb itself or any one word which is subordinate to this verb in a chain of dependencies. Presumably there are further restrictions on this subordination. We shall call the word in the first position simply W1. W1 may be the head of a phrase, so it need not be the first word - e.g. if *svaki*, ‘every’, depends on *dan*, ‘day’, in (61b), then W1 is *dan* but the very first word is *svaki*.

- Either W1 itself or one of its subordinates must be accented.

- Apart from the rigid rules for clitics, the order of words is very free. So if there are no clitics, the finite verb may occur in any position within its clause.

How might this complex system be handled in terms of the WG analysis of clitics developed for English and French? The following suggestions are not intended as a properly worked out analysis, but rather as tentative first guesses.

As in French, any clitic requires a hostword which also contains any other clitics plus one word to act as the hostword's 'stem'. For example, in (61b) the hostword is *dan mu ga*, with *dan* as its stem. However unlike French, the hostword itself is part of the syntactic structure, as though it were an ordinary word, and acts like the verb in a verb-second language¹². In a verb-second language such as German, main clause structure is built round the finite verb which has precisely one extractee and no other pre-dependents. As in English, the extractee dependency combines with some other dependency, so in (64) the extractee is also the object.

- (64) Paul kennen wir gut.
 Paul know we well
 'We know Paul well.'

Similarly in Serbo-Croat: a hostword has precisely one extractee (which is also defined by accent), and all other words must follow it. In simple cases the extractee is also the hostword's stem - e.g. *dan* in (61b), or (most simply) the finite verb *dajem* in (61c), whose structure is shown in Figure 30. For simplicity I have presented the clitic-based properties and the ordinary syntactic dependencies in separate diagrams.

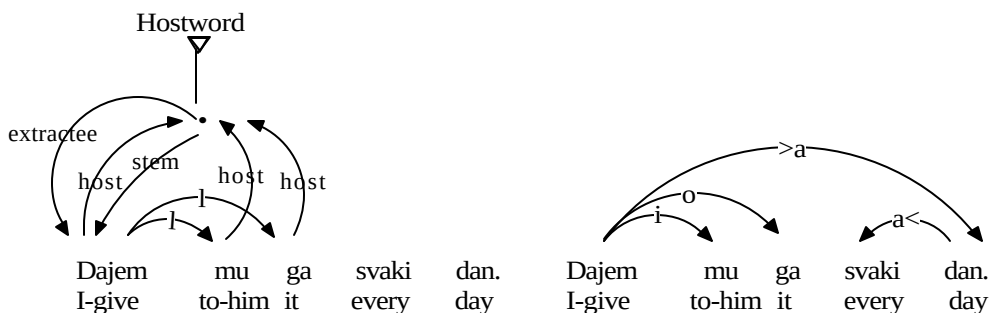


Figure 30

¹² Anderson suggests that the verb-second pattern is a generalisation of the pattern for second-position clitics (Anderson 1993).

This is the basic idea behind the following programmatic analysis, which revolves round the properties of the hostword. The main complexity lies in the very free relations between the hostword's stem (the 'stem word') and the finite verb, with some minor complexity threatening in the relation between the extractee and the stem word. In the example just illustrated the finite verb is also both the stem word and the extractee, but the other examples show that the stem word may be any word which is subordinate to the finite verb in the dependency chain. For example, in (63) the finite verb is itself one of the clitics (*je*, 'is'), so it cannot be the stem word; instead the latter is the adjunct of the subject of the finite verb's complement. The structure is in Figure 31, with the clitic-based and basic dependency-based structures again shown separately.

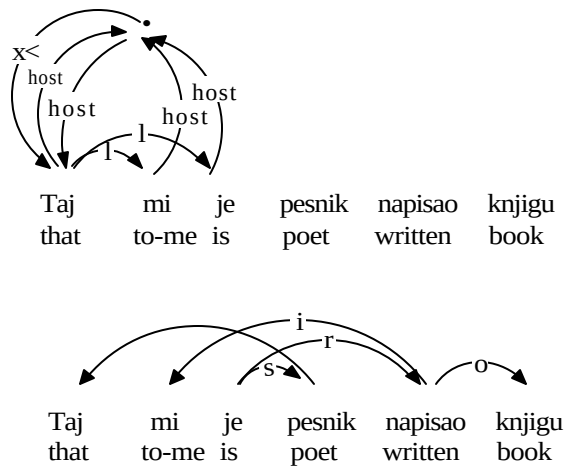


Figure 31

The challenge in this example is to relate the word *taj* and the finite verb *je* so that the former is allowed to be the hostword's stem, given the dependency links that relate them: *pesnik* depends on *je*, and *taj* on *pesnik*. The first step is to introduce 'FINITE-VERB' as a function from the hostword to a finite verb, in order to allow the generalisation that every hostword is associated with a finite verb, and that there can only be one hostword per finite verb - in other words, one clitic cluster per finite clause.

The second step is to define the relationships that are possible between this finite verb and the hostword's stem, the word to which the clitics attach.

- They may be the same word, as in Figure 30, where the finite verb *dajem* doubles up as stem word.

- The finite verb may itself be a clitic auxiliary with one of the auxiliary's dependents as the stem word. This may be the full verb (65), or it may be the subject (66) (ibid:352).

(65) Pisao sam pismo

written am letter.

'I wrote a letter.'

Or it may be the subject as in ...

(66) Devojke su čitale ovu knjigu.

girls are read this book

'The girls were reading this book.'

- If the finite verb is an auxiliary, the stem word may depend not on it but on the full verb, which in turn depends on it; so the stem verb may be two dependency links from the finite verb. No doubt it is possible to increase this number further.

(67) Juče ste čitali knjigu.

yesterday are read book

'Yesterday you read a book.'

(68) Ovu knjigu smo već čitali.

this book are already read

'We have already read this book.'

In short, although in the simplest cases the stem word is the finite verb itself, in other cases it may be what we can call an 'extension' of the finite verb - a dependent, a dependent of a dependent, and so on.

This pattern of relationships requires a new relationship, 'extension', which we can label simply 'x', and a recursive definition which allows a finite verb's extension to be separated from it by any number of dependency links, including none. This pattern, shown in Figure 32, is the same as the formal pattern in Figure 17, but with the addition of a reflexive loop which allows a word to be its own extension.

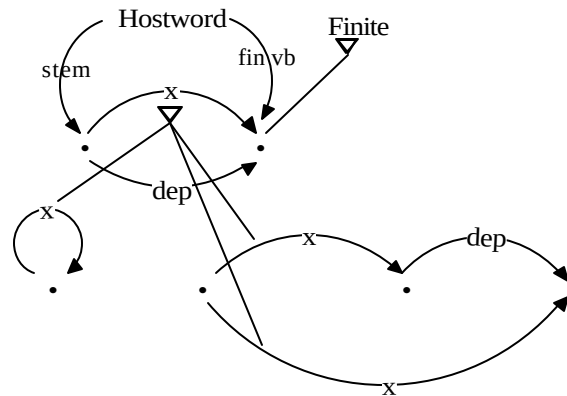


Figure 32

This flexibility in choosing the stem word is combined with the flexibility of clitic climbing which forces a clitic to climb out of a non-finite clause onto verbs such as *želiti*, ‘want’ (ibid:357). For example, (69) contains two clitic pronouns which depend on the infinitive *dati*, ‘give’ but which are attached to the hostword which ‘belongs’ to the finite verb *želim*, ‘I want’.

- (69) Ja mu ga želim dati.
 I to-him it want give
 ‘I want to give it to him.’

Following the model of clitic climbing in French, we can postulate a ‘clitic’ link (labelled ‘1’) between each clitic and the words that are relevant to it, which means not only the stem word from which it takes its position, but also each word in the dependency chain from the finite verb down to the word on which it depends. For example, in (69) *mu* is ‘clitic’ of *ja*, but also of *želim* and *dati*. The structure for this sentence is shown in Figure 33. Once again, to make the diagram comprehensible the clitic-based structure is shown separately from the dependency structure, but it should be remembered that the two structures are in fact one.

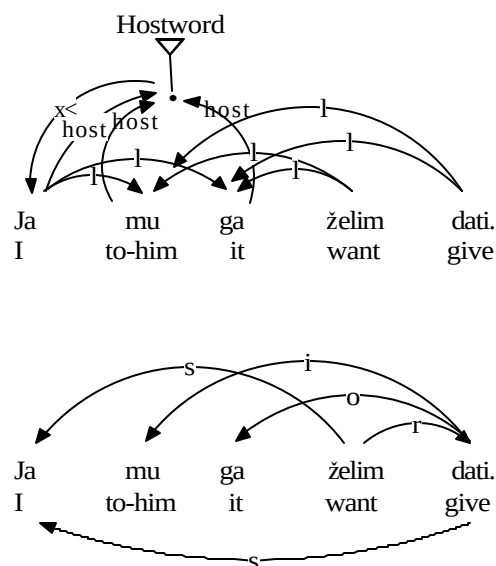


Figure 33

A third possible source of complexity is mismatch between the hostword's extractee ('x<') and its stem. In all the examples given so far they have been the same word, but this is because the extractee happened to have no other words which depended on it and followed it. One assumes that this need not be so; for example, the extractee could be a phrase such as 'people in Belgrade', in which the last word is 'Belgrade' but the phrase's head is 'people'. In this case the stem would obviously have to be the last word rather than the extractee itself. In the absence of facts I shall simply flag this as a possibility.

In summary, then, Serbo-Croat clitics seem to work as follows:

- Each clitic:
 - has a host, which must (of course) be a hostword;
 - is a clitic of the host's stem.
- Each hostword has:
 - a stem, which is a full word;
 - a related finite verb;
 - an extractee, which is in some way connected with focus and stress.
- In the simplest cases, the stem, the finite verb and the extractee are all the same word, which is also the word on which the clitics depend (ibid:354):

(70) Dolazite li često ovamo?

you-come Question often here
 ‘Do you come here often?’

- In more complex cases,
 - the stem may be an ‘extension’ of the finite verb (where ‘extension’ means a dependent, or a dependent of a dependent, and so on), as illustrated in (67-8);
 - the clitic may have climbed up from a lower verb as in (69);
 - possibly the extractee may be different from the stem.

In addition to the complexities that we have illustrated, there is a further challenge: a template of six ‘slots’ for clitics, compared with a mere five for French (Corbett 1992). Presumably these can be handled in much the same way as the clitic clusters of French.

In summary, then, Serbo-Croat clitics can be analysed within the same general framework as French, but with a little extra complexity. The most interesting characteristic of Serbo-Croat is that the hostword itself takes part in syntax by allowing one ‘extractee’. This is equivalent to saying that the clitics occur in second position, because they are required to follow the extractee but no other clause element.

7 Conclusion

We have gradually built up a repertoire of analytical categories for handling clitics, starting with some of the simple clitics of English, then looking in some detail at the quite complex pronoun clitics of French, and finally glancing at the super-flexible system of Serbo-Croat. Each case needed the apparatus for the previous one plus one or two additional relationships. In fact the only difference between the simpler and more complex systems lies in the number of relationships that are specific to clitics. Moreover the analysis of clitics incorporated the same grammatical categories as compound words: words which contain other words, and forms that contain other forms.

It may be helpful to summarise the categories that are needed in these analyses specifically for clitics:

- Two general word classes: Clitic and Hostword.
- One or more of the following relationships:
 - host: a clitic’s host is the hostword that contains it;
 - stem: a hostword’s stem is the non-clitic on which the clitics ‘lean’;

- clitic: a clitic is a 'clitic' of the stem word, or of the word that it depends on, or of a word in the dependency chain between these two;
- finite verb: a hostword's 'finite verb' is the finite verb to which it 'belongs';
- extension: a word is an extension of the hostword's finite verb if it is subordinate to this verb in a dependency chain.

All these relationships are quite easy to learn on the basis of positive evidence, so there is no need to assume that they are part of an innate UG. Some languages have all of them; some (presumably) have none; and others have some but not all. This is as predicted by WG's strongly non-innatist approach to learning.

This range of categories and relationships may best be seen as an analytical framework rather than a theory. However it could be developed into a theory by providing some causal connections. Why do some languages have clitics? Because they provide a useful combination of the compositional freedom of syntax and the easy processing of morphology. Therefore clitics combine syntax with morphology: they are words, with ordinary syntactic dependency links to other words, but they are also part of the (rigidly ordered) morphological structure of a larger word. Given that syntax and morphology exist, clitics are the intersection of their properties that is to be expected if languages evolve under the guidance of functional pressures without any in-built formal constraints whatever.

References

- Anderson, S. (1993). Wackernagel's Revenge - clitics, morphology and the syntax of 2nd position. *Language* 69. 68-98.
- Bonet, E. (1995). Feature structure of romance clitics. *Natural Language & Linguistic Theory* 13. 607-647.
- Borsley, R. and Rivero, M. (1994). Clitic auxiliaries and incorporation in Polish. *Natural Language & Linguistic Theory* 12. 373-422.
- Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.
- Corbett, G. (1992). Serbo-Croatian. In Bright, W.(ed.), *International Encyclopedia of Linguistics*. Oxford: Oxford University Press. 3.422-3.425.
- Creider, C. and Hudson, R. (1999). Inflectional Morphology in Word Grammar. *Lingua* 107. 163-187.
- Grévisse, M. (1986). *Le Bon Usage*. Paris: Duculot.
- Grohmann, K. K. (2000). Towards a syntactic understanding of prosodically reduced pronouns. *Theoretical Linguistics* 26. 175-210.
- Hawkins, R. and Towell, R. (2001). *French Grammar and Usage*. London: Arnold.
- Hudson, R. (1984). *Word Grammar*. Oxford: Blackwell.
- Hudson, R. (1990). *English Word Grammar*. Oxford: Blackwell.

- Hudson, R. (1998). *English Grammar*. London: Routledge.
- Hudson, R. (2000). *I amn't. *Language* 76. 297-323.
- Hudson, R. (2000). Discontinuity. *Traitement Automatique Des Langues*. 41. 15-56.
- Hudson, R. (2000). Language as a cognitive network. In Simonsen, H. G. & Endresen, R. T.(eds.), *A Cognitive Approach to the Verb Morphological and Constructional Perspectives*. Berlin: Mouton de Gruyter.
- Marantz, A. (1988). Clitics, morphological merger and the mapping to phonological structure. In Hammond, M. & Noonan, M.(eds.), *Theoretical Morphology. Approaches in Modern Linguistics*. San Diego: Academic Press. 253-270.
- Miller, P. H. and Sag, I. (1997). French clitic movement without clitics or movement. *Natural Language & Linguistic Theory* 15. 573-639.
- Rudin, C., Kramer, C., Billings, L., and Baerman, M. (1999). Macedonian and Bulgarian li questions: Beyond syntax. *Natural Language & Linguistic Theory* 17. 541-586.
- Sadock, J. (1991). *Autolexical Syntax: A theory of parallel grammatical representations*. Chicago: University of Chicago Press.
- Sag, I. and Godard, D. (1994). Extraction of de-phrases from the French NP. In Anon., *Proceedings of NELS 24, University of Massachusetts at Amherst*. Amherst: GLSA. 519-541.
- Spencer, A. (1991). *Morphological Theory*. Oxford: Blackwell.
- Sproat, R. (1988). Bracketing paradoxes, cliticization and other topics: the mapping between syntactic and phonological structure. In Everaert, M., Evers, A., Huybregts, R., & Trommelen, M.(eds.), *Morphology and Modularity: in Honour of Henk Schultink*. Dordrecht: Foris. 339-360.
- Terzi, A. (1999). Clitic combinations, their hosts and their ordering. *Natural Language & Linguistic Theory* 17. 85-121.
- Volino, M. (1990). Word Grammar, Unification and the Syntax of Italian Clitics. PhD. University of Edinburgh.
- Wilder, C. and Cavar, D. (1994). Long head movement - verb movement and cliticization in Croatian. *Lingua* 93. 1-58.
- Zwicky, A. (1977). *On clitics*. Bloomington, Indiana: Indiana University Linguistics Club.
- Zwicky, A. (1987). Suppressing the Z's. *Journal of Linguistics* 23. 133-148.
- Zwicky, A. and Pullum, G. (1983). Cliticization versus inflection: English *n't*. *Language* 59. 502-513.