

Paola Benincà and Cecilia Poletto

## **On some descriptive generalizations in Romance**

### **1. Introduction\***

In this work we intend to present a number of cross-linguistic descriptive generalizations concerning Romance languages and point out their theoretical relevance for syntactic theory. We will make extensive use of dialectal variation, viewing it as a way to shed light on diachronic processes on the one side and on the complexity of syntactic structure on the other.

We will restrict the empirical domain considering in general only some areas of Romance languages and Italian dialects. The linguistic domain that we take into consideration includes three distinct areas of syntactic processes: *wh*-items and questions in general, personal pronouns and negation. The choice of the grammatical topics is due to both practical and theoretical reasons, as these three domains have been - and still are - central to the development of syntactic theory, and have been systematically explored during fieldwork in the last ten years. The geographical area we have chosen is the one whose micro-variation has been more extensively investigated, both with respect to modern and preceding stages (going back to the 13th century).

The aim of this article is not to provide new analyses for a single phenomenon, but to show how cross-linguistic variation can direct our research towards a precise path and narrow down the number of possible analyses of a given phenomenon. As will appear in what follows, descriptive generalizations will be formulated in their strongest form: this does not mean that we are particularly sure that they cannot be falsified if the domain of languages studied is widened. We think that a generalization has an empirical side, which has the

function of a challenge: it provokes further, more detailed, observations and possibly more accurate description. Even if a generalization ends up being falsified, we will have increased our empirical basis and, more generally, our knowledge of how languages work.

In section 2 we examine the pattern of clitic *wh*-elements and illustrate some empirical generalizations that are valid both diachronically and cross-linguistically; in section 3 we do the same with respect to the emergence of pronominal clitics<sup>i</sup>. A comparison between the two evolutionary patterns is presented in section 4, where we isolate some properties common to both *wh*-items and pronouns. In section 5, we present and discuss some empirical generalizations that lead us to analyze *wh- in situ* and *wh*-doubling in Romance as closely related phenomena. Section 6 illustrates the factors that influence the cliticization process of a preverbal negation marker, namely modality and verb movement. Section 7 deals with the common properties of cliticization phenomena. Although the factors favoring cliticization are different for the various classes of elements (*wh*-items, pronouns and negative markers) that undergo the process, it appears that the pattern of cliticization is essentially the same wherever it is manifested. The clitic elements appear in positions where the strong counterpart used to move to; this can either be hypothesized or attested for past stages of a language or directly observed in the present, in closely related dialects or inside the same one as an option with slightly different interpretations.

## **2. The CP layer and *wh*- clitics**

We will first consider a number of cross-linguistic descriptive generalizations concerning the CP layer and precisely the different forms and behavior of *wh*-items. It has been argued that some *wh*-items display clitic-like properties in Romance (cf., among others, Bouchard and Hirschbühler (1986), Obenauer (1994), Friedeman (1997), Munaro (1997),

Poletto (2000)). A closer examination of cross-linguistic patterns of the elements that can or cannot be clitics reveals an interesting picture of the relations between the morphological makeup of *wh*-elements and their syntactic properties.

We adopt here a pre-theoretical, though precise, notion of clitic as an element that has a strong counterpart; contrary to its strong opponent, it has severe limitations on its distribution, and can in general only occur in a fixed position inside the clausal structure. This definition by contrast excludes from our investigation all functional heads that have indeed a fixed position in the clause but do not have a strong counterpart.

Using the tests first formulated by Kayne (1975) for pronouns, we notice that some *wh*-items develop cross-linguistically and diachronically a tonic/clitic pair more easily than others; moreover, some other *wh*-elements never undergo a process of this kind. The facts can be illustrated by means of cross-linguistic descriptive generalizations expressed in the form of implications. Among the *wh*-items that can display clitic properties in Romance languages we observe the following:

(1) a If only one *wh*- behaves like a clitic it is either *what* or *where*.<sup>ii</sup>

b Elements like *who* and *how* can also display clitic-like properties but this is less frequently the case; moreover, the presence of clitic/tonic pairs for *who* and *how* in a language implies that both *where* and *what* also behave as such.

c The *wh*-element corresponding to *why* never behaves as a clitic, and is always expressed by a compound

In what follows, we illustrate each generalization with examples of various languages.

### 2.1. The *wh*-item *what*

Let us begin by illustrating (1a): there are languages that display a single clitic/strong opposition only for the *wh*-item 'what' or 'where' but not for other *wh*-items. Moreover, there

are languages that have a strong/clitic opposition for both of them, but to our knowledge there is no language that displays a clitic form for 'how' and/or 'who' without also displaying a clitic form for both 'what' and 'where' (cf. (1b)). The same seems to be true from the diachronic point of view: the first elements displaying the clitic/strong opposition can be either 'what' or 'where', while the *wh*-items corresponding to 'who' and 'how' can only become clitics if 'what' and 'where' already are such.

Languages like French only have a morphologically distinct clitic form for 'what', namely *que*. The analysis of *que* as a clitic element to explain its distributional properties has a long tradition in the studies on French syntax (we only mention here, among others, Bouchard and Hirschbühler (1986)).

It is well known that *que* cannot be coordinated, modified, used in isolation or stressed and can only occur adjacent to the inflected verb, while its strong opponent *quoi* is restricted to the *in situ* position and in embedded inflected interrogatives the form substituting *que* is *ce que*.

- (2) a \*Que ou qui a-t-il vu?  
 what or who has.he seen?  
 'What or who did he see?'
- b \*Que d'interessant a-t-il dit<sup>iiii</sup>?  
 what interesting has.he said?  
 'Did he say anything interesting?'
- c \*Que?  
 What?
- d Que fais-tu?  
 what do-you?

- 'What are you doing?'
- e Qu'a fait Jean?  
what has done Jean?
- 'What has John done?'
- f \*Que il fait?  
what he does?  
'What is he doing?'
- g \*Il a fait que?  
he has done what?  
'What has he done?'
- g Je ne sais pas \*(ce) qu'il a fait<sup>v</sup>  
I not know not (what) that he has done  
'I do not know what he did'

A strong/clitic opposition concerning *what* seems to be present also in the standard Italian spoken in Central Italy, who use both *che* and *cosa*, *che* being restricted the way *que* is in French, while *cosa* being a full form which can occur in main embedded interrogatives, be used in isolation, coordinated modified and stressed (see Poletto and Pollock (2001) for a detailed presentation of the various classes of Italian speakers). The same type of partition between a clitic and a strong form has been proposed by Ambar et al. (1998) for Portuguese *que* and *o que*, *que* being the clitic form<sup>v</sup>.

## 2.2. The clitic element *where*

Further comparison with other languages and dialects strengthens the generalizations above: in the dialect spoken in S. Michele al Tagliamento (in the Italian region Friuli) two distinct

forms can be used for the *wh*-item 'where' (*do* and *dulà*), one of them (*do*) displays clitic properties:

- (3) a \*Do e quant (a) van-u? S. Michele al T.  
 where and when (*cl*) go.they?  
 'Where and when do they go?'
- b Dulà?/ \*Do?  
 where?
- c Di dulà/\*di do al ven-ja?  
 from where he comes.he?  
 'Where does he come from?'
- d I so-tu zut dulà? / \*I so-tu zut do?  
*cl* are.you gone where?  
 'Where have you gone?'
- (4) a \*Do a van-u?  
 where *cl* go.they?  
 'Where are they going?'
- b Dulà a van-u?  
 where *cl* go.they?  
 'Where are they going?'
- (5) a A mi an domandat dulà ch al era zut  
*cl* to-me have asked where that he was gone  
 'They asked me where he had gone'
- b \*A mi an domandat do ch al era zut  
*cl* to-me have asked where that he was gone

‘They asked me where he had gone’

Examples (3) to (5) exemplify the characteristics that we consider in order to attribute the locative *wh-* *do* the status of clitic element. The clitic *wh*-item *do* 'where' cannot be coordinated, used in isolation, modified and stressed (cf. (3)). Moreover, it has to occur close to the verb (just like French *que*, cf. (4)) and does not occur in embedded clauses, where *wh*-items have to be followed by a complementizer (cf. (5)). None of these restrictions is found with the full form *dulà* or with the element *sé* corresponding to 'what'. In this dialect the *wh*-item *se* 'what' has just one form, the strong one: this *wh-* behaves like the strong *dulà* and has none of the restrictions met by the clitic form *do*.

Other languages have a wider set of clitic *wh*-items. In the dialect of Pera di Fassa the *wh*-item *co* 'how' patterns with *che* 'what' and contrasts with *can* 'when' in requiring strict adjacency to the inflected verb:

- (6) a Can vas-to pa? Pera di Fassa  
 when go.you *particle*?  
 ‘When are you leaving?’
- b Can pa tu vas?  
 when *particle* you go?
- (7) a Co l fas-to pa?  
 how it do.you *particle*?  
 ‘How do you do it?’
- b \*Co pa tu l fas?  
 how *particle* you it do?
- (8) a Che compres-to pa?

what buy.you *particle*?

‘What are you buying?’

b \*Che pa tu compre? <sup>vi</sup>

what *particle* you buy?

The last generalization concerns the *wh*-item corresponding to 'why': to our knowledge there are no languages which have a clitic form for 'why', even if in several dialects the morphological makeup of the form is different from the typical Romance pattern formed by the preposition corresponding to 'for' and the item corresponding to 'what', as French *pourquoi*, Italian *perché*, Spanish *porqué*, Paduan *parcossa*. Some Rhaeto-Romance dialects have the form *ciuldì*, which can be decomposed in *what-wants-say*, namely 'what does it mean?'. Although this form is nowadays perceived as a single item, it has developed no clitic counterpart; moreover, a survey of the AIS data shows that there are no Italian dialects that have developed a non compound form for the *wh*-item corresponding to 'why', this strengthens further our claim.

The descriptive generalizations in (1) point towards an implicational scale that has interesting similarities with the one of object clitics, as we are going to show in the next paragraph.

### 3. Implications in the emergence of pronominal clitics

In this section we present and illustrate with examples some descriptive generalizations concerning the presence of clitic forms in the pronominal paradigm across languages. The implications that will emerge suggest hypotheses concerning the diachronic stages of cliticization processes in Romance. The descriptive generalizations we formulate in (10) are obviously to be limited to nominal elements in the Romance languages; first of all, it is well

known that in other languages the set of clitic elements is much wider than in Romance. Slavic languages, for instance, also have clitic auxiliaries: Rivero (1994) and Alexiadou (1995) analyze some adverbial forms as clitics in Greek. At present we cannot make statements concerning the relation between these different areas in which cliticization processes instantiate. Moreover, even in Romance some adverbials are weak elements in the sense of Cardinaletti and Starke (1999), and negative markers can also undergo the process of cliticization (as we will illustrate below). Thus, what the generalizations in (10) describe is not a general prohibition against having clitic adverbials, which is immediately falsified by other language groups and by Romance languages, but a diachronic and cross-linguistic implicational scale concerning nominals, which goes from the direct object to non selected items. With these provisos, we expect that the weaker implication (that non-selected elements can be clitics only if arguments are) to be correct even outside the domain of Romance languages<sup>vii</sup>.

- (10) a        If a Romance language has clitics, it has direct object clitics.
- a'        If a Romance language has dative clitics, it has direct object clitics.
- b        If a Romance language has partitive and/or locative clitics, it has dative clitics:  
(there exist Romance languages with direct and indirect object clitics but without partitives,  
no Romance language having direct object and partitive clitics lacks indirect object clitics).
- c        If a Romance language has subject clitics, it also has direct and indirect object  
clitics.
- d        There is no implication between locative and partitive and between locative,  
partitive and subject clitics.
- e        Adverbial clitic forms for elements that are never selected by a verb are much rarer  
and imply the presence of argument clitics.

It is common opinion that all Romance languages have clitics, but there exist a few that do not have any. Although the fact that there exist Romance languages that do not have a clitic system at all can be at first sight surprising, data are quite clear in this respect. The Rhaeto-Romance dialects of Brigels and Camischollas (in the Grisons region) only have tonic pronouns. The data reported here (from AIS VI 1110; VIII 1597, 1650, 1651, 1667) are the counterpart of standard Italian sentences with clitics (and not with tonic pronouns), as the Italian translations (11a'-d') show:

- (11) a      ša ti vol *al*  
               if you want it
- a'      se *lo* vuoi  
               if it you-want
- b      vus amflayas bec *el*  
               you find not it
- b'      non *lo* trovate  
               not it you-find  
               ‘You do not find it’
- c      yu amfla netur *el*  
               I find nowhere it
- c'      non *lo* trovo in nessun posto  
               not it I-find in no place  
               ‘I cannot find it anywhere’
- d      i an caciau giodor *el*  
               they have chased away him

- d'     *lo* hanno cacciato  
        him they-have chased-away  
        ‘They chased him away’

As the example (11a) shows, the object pronoun occurs at the right of the verb. The examples (11b, c, d, e) show that the pronoun is not in enclitic position, as it can perfectly well be separated from the verb by the postverbal negative morpheme or by other elements. These dialects do not have any other dative, locative or partitive clitic, either proclitic or enclitic (the examples are from AIS VI 1110, 1113; VII 1345, VIII 1638, 1659):

- (12) a     *gi kuai ad el*  
        tell that to him
- a'     *diglielo*  
        tell.to-him it
- b     *dai e a nus*  
        give of-it to us
- b'     *daccene*  
        give.to-us.of-it
- c     *k eu mondi*  
        that I go (there)
- c'     *che ci vada*  
        that there I-go
- d     *koy figesas kun el?*  
        what would-you-do with it?  
        ‘what do you do with it?’

- d'      *cosa ne fareste?*  
           what with-it would you do?<sup>viii</sup>

This peculiar area of Romance can provide further, more detailed, evidence in favor of our generalization, because other dialects only have direct object clitics (cf. (13a)) but no dative, locative or partitive clitics (13b,c), a system of this type is not attested elsewhere in the Romance family. The dialects exemplified here are those of Ardez and Remüs (AIS VI 1110, VIII 1638, 1659):

- (13) a      *se tu il vos*  
           if you it want
- a'      *se lo vuoi*
- b      *di ad el*  
           tell to him
- b'      *digli*  
           tell.him
- c      *k e ia*  
           that I go (there)
- c'      *che io ci vada*  
           that I there go
- d      *t/é fessat cun el?*  
           what do-you do with it?
- d'      *cosa ne farete?*  
           what with-it you-will-do?

As for the other arguments, there are no Romance languages that have partitive clitics but do not have dative clitics, while there exist languages that have dative clitics but no partitive or locative (for instance Spanish or Friulian, cf. Chomsky (1981)) hence the implication in (10b):

- (14) a     gli parlo  
           to-him I-talk  
           'I talk to him'
- b     ne voglio una  
           of-it I-want one  
           'I want one'
- c     ci vado  
           there I-go  
           'I go there'

We have not found other implicational relations concerning the presence of locative and partitive clitics and subject clitics, as there are languages that have subject clitics but no locative or partitive, like some Friulian dialects, and there are languages like standard Italian (see 15 b', c') that have locative and partitive clitics but no subject clitics:

- (15) a     Toni al ven            Friulian  
           Toni he comes  
           'Toni is coming'
- b     Toni al è  
           Toni he is  
           'Toni is there'

- b' Toni c'è
- c Toni al compre doi  
Toni he buys two
- c' Toni *ne* compra due  
'Toni buys two of those'

Moreover, no implication has been found between locative and partitive clitics: some Friulian dialects have a partitive, which appears in limited contexts and no locative, while Spanish has a very restricted usage of the locative, and no partitive clitic:

- (16) a Nd ai vjodut nome doi Clauzetto  
Of-them I-have seen only two
- b O viodi nome doi  
I see only two
- c Juan ha ido (alli) Spanish
- d Hay un muchacho

The lack of implication might be tied to the fact that originally the partitive clitic *ne* is also a locative indicating movement out of a place, so that we can suppose that both locative arguments develop together into clitic elements<sup>ix</sup>. Another argument in favor of the idea that the emergence of clitics is tied to case and that direct objects and indirect objects are the first elements developing a clitic series comes from clitic clusters, which interestingly behave in some languages as compounds belonging to a complex but unique syntactic object. Benincà (1988) noted that, although standard Italian is not in general a language that permits doubling of a dative DP, when the dative clitic is combined with a direct object, clitic doubling of the

dative becomes possible, as if the dative clitic (differently from the object clitic) were not "visible" anymore, in some respects:

- (17) a      \*Gli regalo a Mario il mio violino  
               to-him I-give to Mario the my violin  
               'I give Mario my violin'
- b      Glielo regalo a Mario  
               to-him.it I-give to Mario  
               'I give this to Mario'
- c      \*Glielo regalo a Mario il violino  
               to-him.it I-give to Mario the violin  
               'I give Mario my violin'

Locative and partitive clitics also cluster, giving rise to a compound. The following data exemplify the few cases of its functions in Italian and some Veneto dialects:

- (18) a      Ce ne sono due                      Italian  
               *loc part* are two  
               'There are two of them'
- b      Ci sono due ragazzi  
               *loc* are two boys
- c      (Di ragazzi) ce n'è due<sup>x</sup>  
               (of boys) *loc part* is two  
               'There are two (boys)'
- d      \* C'è due ragazzi<sup>xi</sup>  
               *loc* is two boys

- e      Ghe ne zé do                      Venetian  
*loc part* are two
- f      Ghe ne compro do  
*loc part* I-buy two  
 'I will buy two of them'
- g      Te (\*ghe) ne compro do  
 to you (*loc*) *part* buy two  
 'I will buy two of those for you'
- h      Ghi \*(n)'è un tozo                  Coneglianese  
*loc (part)* is a boy  
 'There is a boy'
- i      Ngègghi rivà na fiola              Borgomanerese  
*loc.part.is.loc* arrived a girl  
 'There arrived a girl'

In Italian a locative+partitive cluster *ce ne* is properly connected with a locative and a partitive argument, but the cluster has more features than the sum of its components, as the agreement facts in (18b-d) show. The cluster appears to optionally perform as a subject clitic, triggering agreement. In Veneto dialects the cluster obligatorily appears when the partitive is selected, but the locative disappears if a dative clitic is required. In other dialects, such as Coneglianese (18h), the existential construction with the locative element *ghi* requires also the realization of a partitive clitic *n*. In the dialect of Borgomanero studied in Tortora (1997) this type of locative+partitive cluster surfaces only with unaccusative verbs that have a locative argument. All these cases may suggest that within the positions of clitic elements located in the high portion of the IP layer we have to identify some clitic clusters that are activated

together in the sense that the occurrence of one element of the cluster implies the presence of the other element. This seems to be true for (third person) dative and accusative object clitics, and for locative and partitive. We have not found any cluster formed by three elements.

#### 4. A comparison between pronominal and *wh*-clitics

The comparison between the descriptive generalizations concerning *wh*-items and pronouns leads us to observe the following parallels<sup>xii</sup>:

(19) a in both cases the first elements that give rise to a clitic/strong opposition are those which can be internal<sup>xiii</sup> arguments<sup>xiv</sup>.

b in both cases the clitic corresponding to the external argument is less frequent and implies the presence of direct and indirect object clitics

c non argumental clitic forms are rare and imply the presence of argumental clitics

Concerning (19a), the factors ruling the emergence of a *wh*-clitic element are two: the first factor is connected to the number of semantic features expressed by the *wh*-. Elements like *what* become clitics before elements like *who*, although both can be objects. This leads us to think that inanimacy also plays a role. An example of object *wh*-clitic is the case of French *que*, studied (among others) by Obenauer (1976, 1994). He proposes to connect the special status of *que* to the poverty of semantic features that characterize it. He starts from the assumption that the unmarked value of the feature [+/-animate] is [-animate], which corresponds to the morphological realization of the item *que*. Therefore, in his account the inanimate object is the “least marked” element in the *wh*- series; this poverty is syntactically

encoded as non-expansion of the internal structure of the *wh*-item itself, which is a head and behaves as a clitic. On the other side, it appears evident that also the fact that *que* is a direct object also plays a role in its evolution as a clitic. In modern French, *que* cannot be a subject, as the following contrast illustrates:

- (20) a      \**Que s'est passé?*  
               What itself.is happened?  
               ‘What happened?’
- b      *Que s'est-il passé?*  
               ‘What itself.is.it happened?’

(20a), in which *que* is the subject of the unaccusative verb is ungrammatical. On the contrary, (20b), where a subject clitic appears in enclisis to the inflected verb, is possible. The presence of a subject clitic was not necessary at earlier stages of French, when the language was pro drop; a sentence like (20a) was then possible. However, *que* did not have the typical clitic behavior it displays in modern French, as it was not restricted to main clauses, and could be stressed, modified and found in isolation (cf. Pollock (2001)).

These data show that another factor playing a role in the change has to do with the object nature of *que*. In other words, it seems that when *que* becomes a clitic it is always interpreted as a direct object (cf. (20)): the "prototypical selected element" is the direct object, which is in fact the one which displays clitic properties with the highest frequency.

As for the reason why there is such an implicational scale inside the arguments of the verb and why adverbial clitics are much rarer, we cannot put forth at this point a detailed proposal but can only sketch a possible line of research. The correlations in (19) can be seen as indicating that the development of a given element as a clitic is sensitive to a well-defined

hierarchy, which is ultimately connected to the way thematic roles are encoded in the syntax. It seems that those elements connected to case (in the sense of Kayne (2001), where prepositions are an instance of case) more often develop clitic forms than those that are not. We will not develop this intuition any further, as it lies beyond the scope of this article, which is simply to show how a wide empirical basis can drastically reduce the number of potential explanations for a given phenomenon. We limit ourselves to point out that semantic factors, through the filter of their syntactic realization possibly as case, are at work in the development of clitic forms.

As for the reason why clitic forms emerge, many authors have put forth their analysis, which can be summed up into roughly three major lines of thought: the first and more traditional view is that the origin of clitics has to do with a progressive "erosion" of the phonological component, which at a second stage induces the well-known syntactic phenomena typical of clitics. This approach does not provide any explanation of the process. Another view sees a syntactic process pruning some functional projections internal to the DP - ultimately due to a minimal effort strategy - as the triggering factor for clitic-formation (cf. Halpern (1995), Barbosa (1996), etc). A third one (Jelinek (1996)) conceives cliticization as the answer to a semantic problem concerning the interpretation of non-existential variables. In section 7 we claim that a conspiracy of phonological, syntactic and semantic factors leads to the emergency of clitics in a language. This is in agreement with the new view on the phenomenon put forth by Cardinaletti and Starke (1999), who consider also semantic factors as being essential to the process of weakening, which only at a second stage of evolution becomes syntactically and phonologically relevant.

## **5. *Wh*-doubling and *wh*- *in situ***

Another interesting phenomenon connected to the clitic status of *wh*-items is the one of *wh-in situ*<sup>xv</sup> and *wh*-doubling found in the NIDs: the following generalizations illustrate the connection between the appearance of clitics in the *wh*-paradigm and the possibility of *wh-in situ* on one side and the connection between *wh-in situ* and *wh*-doubling (already noted by Munaro (1997)) on the other:

- (21) a      If *wh-in situ* is found with a single *wh*-item, this *wh*-item corresponds to "what".
- a'     If *wh*-doubling is found with a single *wh*-item, this *wh*-item corresponds to "what".
- b      If a language allows *wh-in situ* cooccurring with SCLI, the only *wh*-items that can be left *in situ* are those that can become clitics<sup>xvi</sup>.
- b'     If a language allows *wh*-doubling cooccurring with SCLI, the only *wh*-items that can be left *in situ* are those that can become clitics.
- c        If a language allows a *wh-in situ* strategy, this is applied to *wh*-phrases only if it applies to *wh*-words.
- c'        If a language allows a *wh*-doubling strategy, this is applied to *wh*-phrases only if it applies to *wh*-words.
- d        *Wh*-doubling in embedded contexts is possible only when the complementizer is not lexicalized.

The first generalization can be illustrated by data reported in Tortora (1997:7). In Borgomanerese only one form of the element corresponding to 'what' can be left *in situ*, all other *wh*-items cannot:

- (22) a.      kus tal [erki?                      Borgomanerese



- (25) a %Ndo va-lo andóe?  
 where goes.he where?  
 ‘Where is he going?’
- b %Ci e-to visto ci?  
 who have.you seen who?  
 ‘Whom did you see?’

The generalization in (21b) can be illustrated on the basis of Bellunese data, reported from Munaro (1997), (1999). In modern Bellunese the class of *wh*-items that occur obligatorily *in situ* are precisely those that display clitic properties cross-linguistically:

- (26) a \*Ché a-tu fat? (Munaro 1997: 3.62) Tignes d’Alpago  
 what have.you done?  
 ‘What have you done?’
- b A-tu fat ché?  
 have.you done what?
- (27) a \*Chi laore-lo?  
 who works.he?  
 ‘Who is working?’
- b E-lo chi che laora?  
 is.he who that works?
- (28) a Va-lo andè?  
 goes.he where?  
 ‘Where is he going?’
- b ??Andè va-lo?

- where goes.he?
- (29) a Se ciame-lo comè?  
 himself.calls.he how?  
 ‘What is his name?’
- b ??Come se ciame-lo?  
 how himself calls.he?
- (30) a In che botega a-tu comprà sta borsa?  
 in which shop have.you bought this bag?  
 ‘In which shop did you buy this bag?’
- b \*A-tu comprà sta borsa in che botega?  
 have.you bought this bag in which shop?

All *wh*-phrases are excluded from the *in situ* position, but the crucial datum showing that the distinction is not the one between *wh*-phrases and *wh*-words is the following:

- (31) a Parché sié-o vegnesti incói?  
 why are.you come today?  
 ‘Why did you come today?’
- b \*Sié-o vegnesti incói parché?  
 are.you come today why?

Hence, the property cutting across elements that can be left *in situ* and elements that cannot is the same that underlies the process of clitic-formation. Moreover, it seems to underlie also the process of clitic doubling, as the second generalization suggests. *Wh*-doubling is limited to the *wh*-items that have clitic forms: in the dialect of Monno, (cf. Munaro (1999)) only the *wh*-

items corresponding to 'what', 'where', 'who' can be doubled:

- (32) a      ché fè-t majà què?  
               what do.you eat what  
               ‘What are you eating?’
- b      ch'è-l chi che vè al to post?  
               what is.it who that goes to the your place?  
               ‘Who is going in your place?’
- c      ngo è-l ndat ngont?  
               where is.he gone where?  
               ‘Where has he gone?’

This is parallel to the behavior of pronominal clitics, which only display doubling when one of the two elements is a clitic. We will analyze this property in section 7.

The generalizations in (21c) and (21c') express the empirical observation that, to our knowledge, there is no variety of Romance that admits *wh- in situ* or *wh*-doubling with *wh*-phrases without admitting it with the subset of *wh*-words that can display clitic properties.

The generalization in (21d) can be illustrated with data of the dialects of Illasi and Monno, where *wh*-doubling in embedded clauses is marginally possible only when the complementizer is not present; the following data also provide an insight of what the relation between doubling and cliticization of *wh*-items could be:

- (33) a      I m à domandà cossa ho fato stamatina                      Bellunese  
               they to-me asked what I-have done this morning  
               ‘They asked me what I did this morning’

- b I m à domandà cossa che ho fato stamatina  
they to-me asked what that I-have done this morning
- c I m à domandà sa ho fato stamatina  
they to-me asked what I-have done this morning
- d \* I m à domandà sa che ho fato stamatina  
they to-me asked what that I-have this morning
- e ? I m à domandà sa ho fato ché stamatina  
they to-me asked what I-have done what this morning
- f À-lo fato ché stamatina?  
has.he done what this morning?  
'What did he do this morning?'
- g \*Cossa à-lo fato ché stamatina?  
what has.he done what this morning

In this dialect there are two possible elements for 'what' which occur at the left of the clause, *sa* and *cossa*: *cossa* can occur either with a lexical complementizer or without a lexical complementizer, as shown in (33a) and (33b). On the contrary, *sa* can only occur when no complementizer is present, (cf. the ungrammaticality of (33d)). Doubling is only possible with the form *sa*, and not with the *cossa* form (cf. the contrast in (24)), both in main and in embedded clauses. In other words, the form that tolerates doubling is the same that does not tolerate a complementizer; hence the form occurring in the highest position in the CP when doubling occurs is precisely a clitic. This in turn shows that the generalization in (22d) concerning the ungrammaticality of clitic doubling in embedded contexts when a complementizer is present has to be derived from the fact that clitic *wh*-items cannot cooccur with a lexical complementizer, contrary to non-clitic *wh*-items. Whatever the explanation for

this turns out to be, (22d) depends on (34):

(34) Clitic *wh*-items are not compatible with a complementizer<sup>xviii</sup>

We will not try to propose an analysis for (34) but notice that there is a striking parallel between the emergence pattern of *wh*-in-situ and the pattern of emergence of *wh*-doubling: the generalizations (22a-a') to (22c-c') suggest that both diachronically and cross-linguistically the development of *wh*- *in situ* and *wh*-doubling follow the same path that is typical of the emergence of a clitic pattern: the first element is the inanimate direct object, followed by other arguments; only at the last stage (represented by languages like spoken French and some Western Lombard dialects) *wh*- *in situ* is generalized to all elements, including complex phrases.

The fact that *wh*- *in situ* is tied to *wh*-doubling on one side and to the emergence of *wh*-clitics on the other has to be taken into account in any theory aiming to explain the *wh*- *in situ* phenomenon in general and its variational and diachronic pattern. The descriptive generalizations above clearly point to a direction that excludes analyses of *wh*- *in situ* which do not encode any relation between the phenomenon and the emergence of a clitic series for at least some *wh*s.

This immediately excludes analyses as the one usually found in the literature for languages like French which views *wh*- *in situ* structures as "pure IPs" where the *wh*-item is left in its argumental position and moves to SpecC only at LF or is not moved at all and the sequence is interpreted as an interrogative by pragmatic strategies.

It also excludes more refined analyses as the one proposed by Rooryck and Cheng (2000), where the *wh*- *in situ* strategy is attributed to the presence of an intonational morpheme in C° realized on the lowest element of the clause, namely the *wh*-item itself.

The only set of solutions compatible with the descriptive generalizations above is the one that takes into account the link between *wh*-clitics and *wh- in situ* structures, which might lead to postulate that in the languages that only have *wh- in situ* but not *wh*-doubling there is a "null doublet" occupying the same position the overt clitic fills in doubling structures.

## 6. Negation and the clitic pattern

An apparent counterexample to the idea that clitics correspond to elements that can be arguments is negation, which is an adverbial element and nevertheless has apparently developed a clitic status in many Romance languages (cf. Pollock's (1989) proposal for French *ne* and Belletti's (1990) proposal for Italian *non*). Before taking into account the pattern of cliticization of the negative marker in Romance and discussing the properties it shares with the development of *wh*-clitics, we provide a brief sketch of the distribution of sentential negation across Romance languages, which we need for the comparison between negation and *wh*-items.

Romance languages have by and large three systems for negating a clause:

- (35) a      preverbal negation (Neg1)  
       b      pre and postverbal negation (Neg 1/2)  
       c      postverbal negation (Neg2)

The languages that only have a preverbal negative marker are the most conservative ones; those that have a pre and a postverbal negative marker have innovated creating a clitic-strong pattern in which the preverbal element is clitic and the postverbal element is not; those languages that only have postverbal negation have lost their preverbal morpheme entirely and

have maintained the strong postverbal negative marker. Anticipating the conclusion we will reach in section 7, we point out that, on a par with *wh*-doubling and pronominal clitic doubling, negation displays doubling when one of the two elements (the higher one) is a clitic form.

This type of evolution is quite common across languages and has been first described by Otto Jespersen (1917), so that the progressive loss of a preverbal negative marker in favor of a postverbal element is known as "Jespersen's cycle". Notice that the term "cycle" can be misleading, since the change has a beginning and an end, and, as far as we know, there are no languages that have undergone the opposite process, developing a preverbal negative morpheme starting from a postverbal one. This is a curious fact in itself, which we will try to explain in section 7. As for now, we will focus on those languages that have a pre and a postverbal morpheme, showing that the emergence of the complex pattern in (35b) is sensitive to a series of factors including modality and sentence type. We will adopt the theory proposed by Zanuttini (1997) for Romance, who distinguishes between two positions for preverbal negation (a clitic and an independent negative head, see also Cinque (1999: 21) and three positions for postverbal negation (whose specifiers are occupied by negative elements). We will use these five positions to account for dialectal data, showing which might be the factors that influence Jespersen's cycle, the various steps of evolution of postverbal negation and finally what negation has in common with *wh*-clitic and pronominal clitics.

Among the languages that have a pre and a postverbal negative marker, there are syntactic contexts in which only one out of two elements is possible. We are now going to illustrate some of these cases.

### 6.1. Negation and verb movement

The first factor that plays a role in the emergence of a postverbal negative marker seems to be sentence type: in some Veneto dialects, interrogative structures show a different negative pattern with respect to declarative clauses.

Most Veneto dialects normally only have a preverbal negative marker and insert a postverbal negative morpheme only when a presuppositional meaning is possible<sup>xix</sup>:

- (36) a      No l vien      Paduan  
                  not he comes  
                  “He is not coming”
- b      No l vien miga  
                  not he comes not

The presuppositions activated by the postverbal negative particle in these languages is similar to those related to *mica* in Italian (see Cinque (1977)). In these languages, preverbal negation is impossible in main interrogative clauses if subject clitic inversion applies: for example a sentence like (36a) cannot have a corresponding interrogative with SCLI:

- 37) \*No vien-lo?  
                  not comes.he?  
                  ‘Isn’t he coming?’

A sentence like (37) becomes grammatical if a postverbal negation is added:

- 38) No vien-lo miga?  
                  not comes.he not?

‘Isn’t he coming?’

In this case, the preverbal negative marker becomes optional, giving rise to a structure which only has a postverbal negative marker, a possibility which is excluded in declaratives:

39) a Vien-lo miga?

comes.he not?

‘Isn’t he coming?’

b \*El vien miga

he comes not

Before analyzing these data, let us consider another Veneto dialect, where a preverbal non-clitic negative head blocks V to C. The case is represented by the dialect of S. Anna, a Veneto dialect spoken south of Venice<sup>xx</sup> In this dialect negation is usually expressed by a pre and a postverbal morpheme, and both elements are obligatory, contrary to Paduan:

(40) Ne l vien \*(mina) S. Anna

not he comes not

‘He is not coming’

In a negative interrogative clause - which displays SCLI, on a par with non-negative interrogative - the preverbal negative morpheme is obligatorily deleted:

(41) Vien-lo mina?

comes.he not?

‘Isn’t he coming?’

- (42) \*Ne vien-lo mina?  
not comes.he not?

What in Paduan is simply a possibility, in the S. Anna dialect becomes obligatory. This pattern is widely attested: postverbal negation becomes the only negative marker when the verb has moved to the Comp domain, as it does in main interrogative clauses (as SCLI indicates).

Hence, in a dialect like the one of S. Anna preverbal negation is incompatible with V to C movement (cf. Poletto (2000) for arguments showing that SCLI corresponds to syntactic movement). This contrasts with French-type languages, where the preverbal negative morpheme has no effect on the interrogative structure with SCLI. Therefore, the preverbal negative morpheme, which cooccurs with the postverbal negative marker *mina* in the S. Anna dialect, must be different from French *ne*. This difference can be explained within Zanuttini's (1997) framework: she proposes that preverbal negative markers are of two types: independent heads and clitic elements. Independent heads block verb movement, clitics do not. As main interrogative clauses have obligatory V to C movement, the independent head *ne* blocking verb movement cannot be present in main interrogatives. According to what we said so far, Paduan has both the “French option” of using a preverbal clitic, which does not interfere with verb movement, and the “S. Anna option” of using an independent head, which cannot surface when V to C applies<sup>xxi</sup>.

Our analysis assumes, then, that independent heads can also cooccur with a postverbal negative marker, and this might lead us to a better understanding of the way postverbal negation arises: in the first stage negation is only preverbal and is probably to be analyzed as an independent head (cf. Zanuttini (1997) for arguments in favor of this hypothesis); in the

second, a postverbal negative element is inserted in contexts of presupposition, as Cinque shows for Italian *mica*, but, if we are on the right track in our analysis of the S. Anna data, even in languages that normally have both a pre and a postverbal negative marker the preverbal negative marker can maintain its properties of independent head. In other words, the number of stages a language undergoes within the Jespersen's cycle must be more than the ones we have roughly sketched in (35). Moreover, even within the same language stage there are contexts in which postverbal negation is favored for independent reasons (in our case, verb movement combined with the  $X^{\circ}$  status of preverbal negation)<sup>xxii</sup>.

There also exist contexts where the presence of a preverbal negative marker is favored even in languages that normally have only postverbal negation, and this seems to be tied to the second factor influencing the complex pattern of pre and postverbal negation, namely modality: in some Northern Lombard dialects, which display only postverbal negation with indicative verbs, preverbal negation occurs at least in modal contexts as subjunctive and conditional (cf. Vai (1996) and p.c.).

The preverbal negative marker surfaces only when the verb is a subjunctive or a conditional, while when the verb is in its indicative form, preverbal negation is not attested, and the only morpheme is a postverbal one. Note that while in the languages examined above there is a ban against the occurrence of a preverbal negative marker, which is the usual form for sentential negation, here there is a requirement which forces the presence of a preverbal negative marker, that does not occur in other contexts. Hence, the loss of a preverbal negative marker is also influenced by modality, a fact that should not be surprising because it matches what we find in other languages (such as Greek, for instance) in which negation also displays sensitivity to mood. What is more interesting is that subjunctive, which probably raises higher than indicative (cf. Poletto (2000)), requires a preverbal negative morpheme, while indicative does not, and not vice versa.

## 6.2 Mood and Negation

Another typical context in which negation changes its form according to modality are imperative clauses. Negative imperatives often display a peculiar pattern different from all other contexts: the imperative form is substituted by a suppletive form when combined with preverbal negation. Benincà (1992) and Zanuttini (1997) note that this is true only for those languages that have only a preverbal negative marker, while postverbal negation is perfectly compatible with a true imperative form:

(43) a Mangia! Standard Italian

eat-imperative

‘Eat!’

b \*Non mangia!

not eat-imperative

(44) Non mangiare!

not eat-infinitive

Don't eat!’

(45) a Bùgia! Piedmontese

Move

‘Move!’

b Bùgia nen!

move not

‘Don't move!’

While the standard Italian sentence in (43b) combining preverbal negation with a true imperative form is excluded, in Piedmontese there is no restriction to the combination of a true imperative form with postverbal negation, as the grammaticality of (45b) shows.

However, there also exist some cases of postverbal negation that is not compatible with a true imperative form (Emilian *mia*, Rhaeto-Romance *buca*: cf. AIS VIII, 1647): the following example is from Emilian:

- (46) Movrat mia!                      Albinea (Emilian)  
       Move-infinit.yourself not!  
       Don't move!

Here the infinitive substitutes for the true imperative even if the negative marker is postverbal<sup>xxiii</sup>.

Zanutini's analysis of this asymmetry between pre and postverbal negation does not explain why cases like (46) should exist. She proposes that preverbal negation requires the presence of a modal projection, whose features cannot be checked by the defective true imperative form. Therefore, a suppletive form able to check the modal feature is used. On the contrary, postverbal negation, being located lower in the structural tree, is not sensitive to mood at all, and it is compatible with a true imperative form. How can cases like the ones in (46) be analyzed within such a framework? The generalization can be maintained in a weaker form: preverbal negation is incompatible with imperatives. Therefore, we would like to preserve Zanutini's intuition that in general preverbal negation is sensitive to mood while postverbal negation is not, given that the majority of languages and dialects confirm her generalization. But how can we integrate somewhat exceptional cases like (46) into her account? We propose that some languages, although they only display a lexical postverbal negative marker, have a

phonetically null preverbal negation imposing the same requirement its phonetically realized counterpart imposes in language with preverbal negation. If this is correct, we are facing once again a case in which negation is not what it seems to be simply looking at the position of the visible negative marker. In other words, in some languages that already seem to have undergone the whole process which transforms preverbal negation into postverbal negation still retain some of the typical features of preverbal negation, leading to the idea that it is still somehow present even where we do not see it. This view agrees in general terms with what we proposed above for the S. Anna dialect, where a preverbal negative marker, which looks at first sight like French clitic *ne*, is not clitic at all, although it obligatorily cooccurs with a postverbal negative element.

In a diachronic perspective, we can better understand this state of affairs: some dialects that used to have a preverbal negation maintained that preverbal NegP active, even when they stopped inserting a morpheme there. This can be thought of as a stage in the diachronic process, which is recognized and becomes visible in those contexts in which preverbal negation clearly differs from postverbal negation, as imperative clauses.

Postverbal negation is also sensitive to the presence of Neg-words inside the clause. In general, as originally noted by Rizzi (1982), in languages with preverbal negation, the preverbal negative marker is obligatory if the sentence contains a postverbal Neg-word; in languages with postverbal negation, the negative marker is not obligatory in this case. However, the concept of postverbal negation is in itself misleading; as has been shown by Zanuttini (1997), the data of NIDS provide evidence for at least three postverbal positions for negation: one position corresponds to the presuppositional negation (cf. Italian *mica*), and is located higher than the adverb 'already' in Cinque's (1999) hierarchy of lower adverbials; a second position is located lower than 'already' but higher than 'no longer'; a third position is located lower than 'no longer' and 'always' and is usually focalized. We report here

Zanuttini's (1997) scheme for possible positions of postverbal negation

(47) [NegP1 non [TP2 [NegP2 pa [ TP2 [NegP3 nen [ Asp perf. [Asp gen/progr. [NegP4 no]]]]]]]]]

The three distinct postverbal negative markers in (47) correspond to three different etymological types, the highest element usually derives from a word meaning "small quantity" (Italian *mica*, Emilian *brisa*, French *pas*, Milanese *minga*, Rhaeto-Romance *minne*, Polesano *mina*, Lombard *mia*) the second from the N-word corresponding to 'nothing' (Piedmontese *nen*, Rhaeto-Romance *nia*), the third is the same morpheme used as pro-negative sentence (Lombard *no*).

These three types of postverbal negation obey different constraints with respect to the cooccurrence with an N-word. In some dialects both the *mica*-type and the *niente*-type can cooccur with N-words, while the lowest postverbal negation *no* does not tolerate any type of N-word:

(48) A ne l'è mina vignù nisun                      Loreo (Rovigo)

*cl not cl is not come nobody*

'Nobody came'

(49) A l'à nen vist gnun                              Piedmontese (Zanuttini )

*cl cl has not seen nobody*

'He saw nobody'

(50) \*A l'à vist no nisun                              Milanese

*cl cl has seen not nobody*

'He saw nobody'

Although it is probably not possible to identify each position with a distinct etymological type, clearly each type corresponds to a structurally contiguous series of negative markers. In this view, belonging to a given etymological type does not constitute sufficient evidence to place the negative marker in a single position, but at least provides evidence for isolating a small set of possible contiguous positions among which the negative marker selects its own.

If we are on the right track, the three distinct etymological types might also correspond to three distinct possible ways of negating a clause: taking Cinque's (1999) hierarchy of functional projections, Zanuttini (1997) proposes that a presuppositional negation is located immediately above the Anterior Tense projection and a non presuppositional negation is located immediately above a Perfective Aspect projection. However, French *pas* occurs in a position that should always give rise to a presupposition, although this is not the case. Hence, the NegP higher than Anterior Tense is not necessarily bound to be presuppositional.

Whatever the original functions of these elements, when they become a negative marker, they lose their peculiarity, but apparently preserve the position they had in connection with their function.

The etymological origin of these negative markers will possibly result in matching their position in the clause only in part; the connection is, first of all, a diachronic one; when they resulted in substituting the function of the preverbal negative marker, their original function weakened, and it is surprising to find their position to still correspond to the one they used to occupy before losing their original meaning. It could also be found that in some cases they moved upwards. About their position we only point out the following: the type of negative marker deriving from the negative direct object "nothing", like Piedmontese *nen*, occurs immediately above Terminative Aspect in Cinque's hierarchy (or Perfective Aspect in Zanuttini's approach); negative markers originally indicating "small quantity", which add a

presuppositional content to negation, are located above perfective Aspect. For the moment we leave this argument to future research.

Another aspect that points towards a connection between the functional properties and the internal makeup of the negative marker reflected in etymology concerns the possibility for a negative marker to become an "expletive" negation. As Portner and Zanuttini (1996) show, the so-called "expletive negation" in exclamative contexts is not expletive at all, but has a specific function: it widens the scale of implicature of the exclamative clause. A potentially interesting generalization connected to this concerns the type of negative marker that can act as expletive negation: it has been proposed that only preverbal negative markers can perform this function. This is clearly not true, as German *nicht*, which is located quite low in the structure of the clause, can be found in exclamative clauses with expletive value. However, the generalization is not entirely incorrect but has to be made precise along the following lines: the negative markers that can function as expletive negation in exclamative contexts are those that contain a negative morpheme. Given that in general preverbal negative markers are made up precisely by a negative morpheme, they all have the relevant property that enables them to behave as expletive negation. Nevertheless, reformulating the generalization has consequences on postverbal negative markers. It predicts that among the postverbal negative markers, only those containing a negative morpheme, like German *nicht*, but not those which do not contain any (for instance French *pas*) can be used as expletive negation. Again, this descriptive generalization makes reference to the internal form of the negative marker and not to a specific position, which is probably only indirectly connected to the morphological makeup.

After having recognized some interesting properties that characterize the various types of negative elements, and the way they are localized in the structure in relation to their form and original value, we can observe how the general developmental path of negation, starting from

a very high position in the clause and ending up in a low position, shows interesting similarities to *wh- in situ* cases<sup>xxiv</sup>. In particular, the intermediate stage with a pre and a postverbal negative marker is similar to *wh*-doubling cases; moreover, the higher *wh*-in doubling structures and the preverbal negative marker in pre and postverbal cases are similar in their feature composition: in both cases the higher element encodes only part of the information encoded in the lower element. In both cases there are good reasons to believe that this is due to the clitic nature of the higher element. The same is obviously true for pronominal clitics, where the clitic is only a partial copy of the information contained in the full DP.

Postulating preverbal empty negative elements as we have done leads us to another interesting comparison with clitics, which across the Romance languages also alternate with an empty category, namely *pro*. The case of subject clitics is probably the best known case of this type: some languages have (different types of) subject clitics, others have *pro*. The same would be true for preverbal negation, which probably also has a null counterpart in some languages.

But pronominal clitics do not seem to follow the same diachronic path that negation and *wh*-display, namely a pattern which is starting from of a high position which is progressively lost through a mechanism of doubling of a clitic and a full form. However, what has been noticed in the evolutionary path of negation and *wh*-items might be helpful for reinterpreting an old idea proposed originally by Antinucci and Marcantonio (1980) who viewed the position of clitics as marking the position where direct and indirect objects used to appear in Latin. Assuming Kayne's (1994) hypothesis that in SOV languages the object moves to a position higher than the verb, we might interpret the change from SOV to SVO that has occurred in the early development of Romance from Latin as related to the loss of object movement to a very high position in the clausal structure. Hence, the emergence of object clitics would also be, on a par with negation and *wh*-items, related to a change in the movement pattern of objects,

which used to raise to a very high position and have lost this property by creating a clitic counterpart which still moves high in the structural tree. In what follows we will try to develop this intuition on the rising of clitic series as related to the progressive loss of syntactic movement.

## 7. Some theoretical observations

In the preceding sections we have shown that the development of clitic forms for *wh*-items, pronouns and negative markers is influenced by a number of factors. These factors vary according with the type of element that is undergoing the process. Pronominal elements become clitics following a well-defined hierarchy: indirect object clitics imply the presence of direct object clitics, and both direct and indirect object clitics are present in languages that have subject clitics. *Wh*-items are also sensitive to the same hierarchy (as the ban against subject *que* in modern French shows), but their development as clitic forms is also influenced by the number of semantic features the *wh*-element encode: the poorer the semantics of an element is, the easier it becomes a clitic. Moreover, clitic *wh*-forms are banned from embedded contexts, while strong forms are not. Hence, clitic *wh*-forms are sensitive to (some form of) verb movement to the C domain. Verb movement also influences negation: preverbal negative markers which are independent heads are not allowed in V to C contexts, as shown by the distribution of preverbal and postverbal negative markers in main interrogatives. Negation is in turn also sensitive to mood, as preverbal forms are preferred in some usages of subjunctive, while being banned in indicative clauses. Clitic formation is thus sensitive to both semantic (as the intrinsic semantic poverty of some *wh*-items which are more frequently clitics testifies) and syntactic factors (as verb movement).

More generally, it is possible to characterize all cases of cliticization along the following

lines: the first common property characterizing the process is well known. Clitic forms generally display only part of the semantic features that strong forms encode. In a parallel way, it seems that the morphological and phonological reduction that the elements undergo leads to an  $X^{\circ}$  containing only part of the functional information and no lexical information of the full form (see footnote 3). Again, it is well known that third person pronominal clitics are very similar to determiners, while the preverbal negative marker that becomes a clitic contains only the -n- morpheme indicating that an element is negative and does not derive from a form indicating “small quantity” or from a negative quantifier like “nothing” as postverbal negation does (see section 6). In other words, preverbal negation does not encode any quantificational feature, while postverbal negation apparently does. The same seems true also for *wh*-items: in Friulian the clitic form *do* loses precisely the “locative” indicator “là” (where) that the strong form *dulà* still retains; *que* has been noticed (cf. Obenauer (1976)) to be identical to the complementizer, a purely functional element. Therefore, the morphological makeup of clitic elements preserves the functional information, lacking the lexical part, which is probably located lower in the internal structure of the element itself.

We formalize the observation that clitic forms only encode the functional portion of their strong counterpart by proposing that clitics forms are the overt counterpart of the mechanism of feature movement at LF proposed by Chomsky (1995): pied piping of the whole category is not necessary, only the feature(s) that has to be checked moves to the target projection.

The clitic form checks the semantic features of an element in a high functional projection without pied piping the whole complex category: pronominal clitics check agreement feature in a high Agr projection without pied piping the whole  $DP^{xxv}$ . The same is true for *wh*-doubling structures<sup>xxvi</sup>, where the clitic checks the interrogative features of the sentence initial projection in the interrogative CP without moving the whole *wh*-, located in a lower position.

Our account also derives the observation that diachronically clitics seem to develop when

movement of the independent category gets lost: if the connection originally noted by Antinucci and Marcantonio (1980) between the development of pronominal clitics and the change from SOV to SVO from Latin to Romance has to be interpreted as we propose in section 6.2: clitic forms develop when movement of an independent element to a high functional projection is lost. This appears quite evident in the case of *wh*-doubling and *wh- in situ*, which substitute for *wh*-movement to the clause-initial position and is also transparent in Jespersen's cycle, where an independent preverbal negative head is substituted by a preverbal clitic form cooccurring with a postverbal negative marker.

If the relation between the rising of SVO and the development of clitics is to be analyzed as a loss of object movement to a very high Agreement projection, the interpretation of clitics as a case of more economical, non-pied piped version of movement due to feature checking also applies to pronouns. Negation also checks a negative feature in a high NegP projection, without pied piping the whole XP, which remains in a postverbal position. Evidently, the rising of clitic forms constitutes the first step towards the loss of movement to the high functional projection: negation in spoken French, Piedmontese, Lombard and in some Rhaeto-Romance dialects has undergone the whole of Jespersen's cycle, as postverbal negation is the only element occurring in these languages. The same seems to be true for *wh- in situ* of the French and Northern Lombard type: *wh- in situ* is possible with all *wh*-elements, showing that this is not more related to any clitic phonetically realized or null form and that in these structures *wh*-movement to the initial CP projection has been completely lost. As for the reason why Romance languages have maintained pronominal clitics retaining just postverbal objects, although the evolution of a clitic pattern is quite ancient, there must be some independent acquisitional evidence that agreement projections are strong in Romance, hence must be checked in overt syntax by the "smallest possible element" namely a clitic.

The idea that clitics are connected to the loss of movement provides some interesting insight

into some of the facts noted in the previous sections. In section 3 we noticed that adverbial clitic forms other than negation (i.e. corresponding to lower adverbs like *più* 'anymore', *già* 'already', *mai* 'never', *sempre* 'always', or higher adverbs like *forse* 'perhaps', *fortunatamente* 'luckily', *francamente* 'honestly', etc.) are virtually absent from Romance languages: if clitics are developed as a sort of more economical strategy to check a high projection, it is straightforward that elements that do not move do not develop clitic forms. According to Cinque's (1999) analysis of adverbs in Romance, these elements do not move to check any functional features located in higher projections. The only movement that adverbs can undergo in Romance is focalization, which is clearly a different kind of movement with respect to the checking of some functional features. Hence, if adverbials do not move in Romance, they are not expected to develop clitic forms.

## References

- AIS. *Sprach- und Sachatlas Italiens und der Südschweiz*. Rengien, Zofingen 1928-40.
- Alexiadou, A. (1995). Word Order Alternation in Modern Greek. Paper presented at the Fifth Colloquium on Generative Grammar, La Coruña, Spain.
- Antinucci, F. & A. Marcantonio (1980). I meccanismi del mutamento diacronico: il cambiamento d'ordine dei pronomi clitici in italiano. *Rivista di Grammatica Generativa*, 5:3-50.
- Ambar M., H. Obenauer, I. Pereira & R. Veloso (1998). The Internal Structure of *wh*-phrases in Portuguese and French. Paper presented at the round table *Structures formelles du langage* (UPRESA 7023), Paris.
- ASIS. *Atlante sintattico dell'Italia Settentrionale*. Unpublished material. CNR. Department of Linguistics, University of Padua.
- Barbosa, P. (1996). A New Look at the Null Subject Parameter. In J. Costa *et alii*, eds., *Proceedings of ConSOLE 3*, Leiden University.
- Belletti, A. (1990). *Generalized Verb Movement*. Rosenberg and Sellier, Turin.
- Belletti, A. and L. Rizzi, (1988). Psych-verbs and Theta Theory. *Natural Language and Linguistic Theory*, 6:291-352.
- Benincà, P. (1988). L'ordine degli elementi della frase e le costruzioni marcate. In L. Renzi, ed., *Grande Grammatica Italiana di Consultazione*, Vol. 1, pages 129-194. Il Mulino, Bologna.
- Benincà, P. (1992). Geolinguistica e sintassi. In G. Ruffino (ed.) *Atlanti linguistici italiani e romanzi*, pages 29-4. CSFLS, Palermo.
- Bouchard D. & P. Hirschbühler (1986). French *QUOI* and its clitic allomorph *QUE*. In C. Neidle and R. A. Nuñez Cedena, eds., *Studies in Romance Languages*, pages 39-60. Foris, Dordrecht.
- Cardinaletti, A. and M. Starke (1999). The Typology of Structural Deficiency: On the Three Grammatical Classes. In H. van Riemsdijk, ed., *Clitics in the Languages of Europe*, vol. 8 of *Empirical Approaches to Language Typology*, pages 145-234. Mouton de Gruyter, Berlin.
- Cheng, L.S & J. Rooryck (2000). Licensing *Wh*-in-Situ. *Syntax*, 3.1:1-19.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.
- Chomsky, N. (1995). *The Minimalist Program*, The MIT Press, Cambridge, Mass.
- Cinque, G. (1976). "Mica". *Annali della Facoltà di Lettere e Filosofia dell'Universtà di Padova*, 1:101-112.
- Cinque G. (1999). *Adverbs and Functional Heads*. Oxford University Press, New York.
- Friedemann M.-A. (1997). *Sujets syntaxiques, positions, inversion et pro*, Peter Lang, Berlin.
- Dobrovie Sorin, C. (1994). *The Syntax of Romanian: Comparative Studies in Romance*. Mouton de Gruyter, Berlin.
- Etxepare R. & M. Uribe-Etxebaria (2000). On the properties of *in situ Wh*-questions in Spanish. Talk delivered at the Workshop *Minimal Elements of Linguistic Variation*, Paris.
- Halpern, A. (1995). On the Placement and Morphology of Clitics. CSLI Publications, Stanford.
- Jelinek, E. (1996). Definiteness and Second Position Clitics in Straits Salish. In A. Halpern & A. M. Zwicky (eds.), *Approaching Second: Second Position Clitics and Related Phenomena*, pages 271-298. CSLI Publications, Stanford.
- Jespersen, O. (1917). *Negation in English and Other Languages*. Host, Copenhagen.
- Kayne, R. (1975). *French Syntax*. MIT Press, Cambridge.
- Kayne, R.S. (1994). *The Antisymmetry of Syntax*. The MIT Press, Cambridge, Mass.
- Kayne, R.S. (2001). Prepositions as Probes. Unpublished manuscript, New York University.

- Munaro N. (1997). *Proprietà strutturali e distribuzionali dei sintagmi interrogativi in alcuni dialetti italiani settentrionali*. PhD thesis, Dipartimento di Linguistica, University of Padua.
- Munaro, N. (1999). *Sintagmi interrogativi nei dialetti italiani settentrionali*, Unipress, Padua.
- Obenauer, H.-G. (1976). *Etudes de syntaxe interrogative du français. 'Quoi', 'combien' et le complémentateur*, Niemeyer, Tubingen.
- Munaro, N. & C. Poletto (1998). On two types of *wh-* in situ. Unpublished manuscript, University of Venice and ISTC-CNR.
- Obenauer, H.-G. (1994). *Aspects de la syntaxe A-barre. Effets d'intervention et mouvements des quantifieurs*. Thèse de Doctorat d'Etat, Université de Paris VIII.
- Poletto, C. (2000). *The higher functional field in the Northern Italian dialects*. Oxford University Press, New York.
- Poletto, C. & J.-Y. Pollock (2001). On the left periphery of Romance interrogatives. Unpublished manuscript, Universities of Padua and Amiens.
- Pollock, J.-Y. (1989). Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry*, 20, 365-424.
- Pollock, J.-Y. (2001). Subject clitics, Subject Clitic Inversion and Complex Inversion: Generalizing Remnant Movement to the Comp Area. Unpublished manuscript, Amiens.
- Portner, P. and R. Zanuttini (1996). The Syntax and Semantics of Scalar Negation: Evidence from Paduan. In K. Kusumoto, ed., *Proceedings of NELS XXVI*, pages 257-271. GLSA, University of Massachusetts, Amherst.
- Rivero, M.L. (1994). Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory*, 12.1:63-120.
- Rizzi, L. (1982). *Issues in Italian Syntax*. Foris, Dordrecht.
- Sportiche, D. (1996). Clitic Constructions. In L. Zaring, and J. Rooryck, eds., *Phrase Structure and the Lexicon*. Kluwer, Dordrecht.
- Tortora, Ch. (1997). *The Syntax and Semantics of the Locative Expletive*. Ph.D. thesis, University of Delaware.
- Tortora, Ch. (2002). On Prepositions and Spatial inalienable Possession. Talk delivered at LSRL XXXII.
- Vai, M. (1996). Per una storia della negazione in milanese in comparazione con altre varietà altoitaliane. *ACME*, 49.1:57-98.
- Zanuttini R. (1997). *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford University Press, New York.



## \*otes

\* Our heartfelt thanks go to the editors of the volume, Guglielmo Cinque and Richie Kayne, for their comments and suggestions on the first draft of this paper. We are also grateful to the participants in the ASIS weekly seminars for insightful discussions and data.

When not otherwise indicated, the source of the examples is the ASIS corpus.

For the concerns of the Italian academy, Cecilia Poletto takes responsibility over sections 1-4 and Paola Benincà over sections 5-7.

i

In this work we only examine head clitics, leaving weak pronouns as defined by Cardinaletti and Starke (1999) aside.

ii

This is true in general for Northern Italian dialects, but is also the case in standard French, standard Portuguese and Southern Italian dialects, although our empirical basis with respect to Southern Italian dialects is far more limited than for the dialects of Northern Italy.

iii

Apparently clitic *que* is different from object clitics in languages such as some NIDs, where it is possible to cliticize the pronoun to a preposition (see (i)), hence the contrast between (i) and (ii):

(i) El vien drio-ghe

He comes behind.him

(ii) \*avec que s'est-il blessé?

with what has-he hurt himself?

iv

Note that cliticization is not a simple matter of adjacency, as *que* is ungrammatical in embedded stylistic inversion interrogatives, where it is adjacent to the verb:

(i) \*Je ne sais pas qu'a fait Jean

I not know not what has done Jean

For an analysis of this effect see Poletto and Pollock (2001).

v

As Richard Kayne pointed out to us, most of the restrictions on *que* also apply to *que diable*, thus suggesting that they are not due to the clitic status of *que* but to the "semantic poverty" of this item. However, the fact that *que diable* can violate the constraint on adjacency leads us to think that *que* is indeed a clitic element and that its semantic poverty has probably contributed to its development as a clitic element. In Old and Middle French *que* was not a clitic, as cases of non-adjacency are attested; some speakers probably still have that type of grammar. As for the constraints on its *in situ* counterpart *quoi*, we cannot even begin to do justice to its complex syntax here; we limit ourselves to refer to Poletto and Pollock (2001) for the idea that the opposition between *que* and *quoi* cannot simply be described in terms of strong versus clitic.

vi

Some speakers find these sentences acceptable with a pitch intonation on the *wh*-item. This changes the interpretation of the question, which becomes of the type: "tell me exactly how you do it, or what you buy".

vii

On some interesting topics, such as possessives, adnominal genitives, object pronouns in Brazilian Portuguese, etc., we do not have data or descriptions detailed enough.

There could still be the doubt that this dialect does not lack clitics but is in fact like Borgomanerese (see Tortora (1997), (2002)), which has clitic pronouns appearing in enclisis after the inflected verb or a special class of adverbs and prepositions. The evidence in favor of our interpretation is probably not compelling; however we point out that in Borgomanerese the pronoun is phonologically reduced and fused with its host (a process typical of enclisis), while this is not the case in the dialects we analyze here. Moreover, the order in clitic clusters is Dative -Accusative and not the opposite.

ix

While direct and indirect object clitics and subject clitics still preserve in their morphological makeup at least person (frequently also gender and/or number information), locative arguments - apart from the very special clitic for 'out of a place', which only survives in formal Italian and French - reduce to just one clitic, completely void of any feature with respect to their strong counterpart: they encode neither any directional information nor any information concerning the reference to the speaker or to the hearer. We will come back to this in section 7.

x

The presence of the left dislocated element *di ragazzi*, of which *ne* becomes a copy, has no consequence.

xi

The sentence is only acceptable in the colloquial registers of some regional varieties of Italian.

xii

The two sets of descriptive generalizations also display one major differences concerning datives, which can be clitics in the personal pronominal series, but never are in the interrogative series.

xiii

As for *wh*-items, the fulfilling of this requirement is not enough; the internal argument has to be inanimate, while this is not the case for pronouns. This difference is clearly tied to the fact that animacy is not morphologically encoded in pronouns while it is in *wh*-items.

xiv

We are aware of the fact that we are simplifying the relation between argumentality and case, if we think for example of cases like inalienable, benefactive and ethic dative clitics, which are not arguments but still have the form of a dative clitic. These generalizations should be seen in a diachronic perspective: the clitic form arises in correspondence with the argumental usages and is then extended to the non-argumental cases. A topic related to this concerns the number and type of features that two elements must share in order to be affected by the extension of the pronominal clitic form.

For some tests distinguishing argumental and non-argumental *ci* in Italian see Benincà (1988).

xv

Here we use *wh- in situ* as a pre-theoretical term, and do not imply any analysis of the phenomenon, which could be handled as a case of covert movement to the CP domain; if much recent work on *wh- in situ* (cf. Pollock Munaro and Poletto (1998), Etcheparre and Uribe-Exebarria (2000)) are on the right track, the low position of the *wh- in situ* is not a true *in situ* one but a low position in the CP domain.

xvi

Munaro and Poletto (1998) show that *wh- in situ* is not a unitary phenomenon and that it is important to distinguish between *wh- in situ* with SCLI and *wh- in situ* without SCLI.

xvii

The symbol % marks the difference among different classes of speakers.

xviii

This is probably due to the fact that *wh*-clitics have to cliticize onto a verbal form (see Poletto and Pollock (2001)) and are then "dragged along" by the verbal complex moving to the Comp domain. As R. Kayne pointed out to us, the incompatibility of *wh*-clitics and the complementizer contrasts a bit mysteriously with the close combination of subject clitics and complementizers. On the other hand there is an interesting similarity between *wh*-clitics and one specific type of subject clitic, namely *a*: they can never be in enclisis to a verb.

xix

The presuppositional value is similar but not completely parallel to that of standard Italian *mica*, analyzed in Cinque (1976).

xx The data have been collected through fieldwork; they are not available in the ASIS corpus.

xxi

Italian can be interpreted in the same way, given that the data are the same.

xxii

At the first stage of evolution the presence of a Neg-word like *niente* 'nothing' as the internal argument favors the loss of preverbal negation: cf. Vai (1996) for Old Milanese, but cf. also the following standard Italian examples:

i) Sarà niente

it-will-be nothing

'It is nothing serious'

ii) Sembra niente

it-seems nothing

'It does not seem serious'

iii) L'ha ridotta a niente

her.he-has reduced to nothing

'He destroyed her'

In the same contexts French *ne* can be omitted.

xxiii

There are also interesting instances of negative imperative in Emilian dialects in which the negation *briza*, normally postverbal, appears preposed and the imperative is substituted with the infinitive (cf. AIS VIII, 1647: *briza movrat* "don't move *infin.* yourself").

xxiv

This parallel development between negation and *wh-* was first pointed out to us by Massimo Vai, p.c.

xxv

Cf. Sportiche (1996) for a similar idea that clitics check Agr features that are checked by scrambled DPs in languages like German. Note however that the technical execution of Sportiche (1996) is quite different from ours.

xxvi

On the view that *wh- in situ* is the null counterpart of *wh-*doubling as Pollock and Poletto (2001) propose, *wh- in situ* - at least when it is restricted to *wh-*items that can become clitics - is totally equivalent to *wh-*doubling with a null clitic