Andrea Moro

Istituto Universitario Lingue Moderne (IULM) Milano

ed Università di Venezia

Heads as Antecedents: a Brief History of the ECP (*)

1. A Fundamental Asymmetry

A fundamental empirical asymmetry yields a sharp partition between subjects and objects in English:

(1)a [which photos] do you think (*that) t frightened those boys

b [which boys] do you think (that) these photos frightened t

Although the data are sharply clear, i.e. in a question, only the object can be extracted across an overt complementizer, there is no general agreement about the form in which the principle explaining this phenomenon should be formulated. Since its first analysis in Chomsky (1981) this fundamental asymmetry has been closely related to the *local environment* where the interrogative element starts its movement where by local environment of a certain element we mean that part of the structure which immediately surrounds the element. Given that the origin of a movement yields an empty (i.e. non-phonologically realised) category, represented by a trace *t*, the principle which aims to capture this asymmetry has been called the *Empty Category Principle* (ECP).

A brief history of the debate surrounding the implementation of such a principle can clarify the theoretical issues which this phenomenon raises.

1.2. Chomsky's *Pisa Lectures:* a Binary Principle

The intuitive content of the ECP can be captured by saying that for each trace there must be an element licensing it in a local environment. Locality is embodied in the formulation of the ECP via the notion of "*government*". By relying on the configurational representation of phrases (namely, X'-theory), we can assume that the licensor must govern its trace; technically, they must *m-command* each other, that is they must not be separated by a maximal projection.[1]

Once the structural environment has been defined, *proper government* is stated as a restriction on the class of possible governors. This narrower relation can be performed by only two kinds of elements: either by a lexical head, like a verbal head for example (i.e. V°), or by a maximal projection coindexed with the trace to be governed. Since government by a maximal projection is in fact a form of *binding*, a further restriction must be embodied, namely that the maximal projection *c-commands* the trace to be licensed. This notion captures a narrower structural restriction w.r.t. m-command: in the latter case, the element interrupting "command" is a maximal projection, in the former every branching node is sufficient to block a potential relation. We can paraphrase the first formulation of the ECP (the ECP_1) as follows:

(2) the ECP_1: t must be either (i) governed by a lexical head or
(ii) governed by a maximal projection binding it

Following usual terminology, we will distinguish (i) from (ii) by calling them respectively *head government* and *antecedent government* .

Within the current framework proper government is not the only kind of relation which is restricted by a topological notion: in fact, a trace is submitted to a twofold tension since, *once* it passes the ECP, it must also fulfill a "metric" condition which measures the maximal distance separating it from its first antecedent, and rules out those pairs which stand too far apart. In general, Subjacency condition creates a further structural bound in

the sense that it excludes those pairs which are separated by two or more occurrences of a maximal projection (see Roberts (1988) for a brief history of Subjacency).

Thus, two important principles of grammar make crucial use of the notion of "maximal projection" to define the structural area where they can display their activity, perhaps with some degree of overlapping. Although one can reasonably explore the path toward a unification of the topological conditions which enter into the ECP and Subjacency (see for example Chomsky (1986) and Cinque (1990)) for the sake of simplicity we will maintain the two principles separated.[2]

Within Chomsky's *Pisa Lectures* framework the ECP_1 derived the asymmetry in (1) in a quite straightforward way. We can focus on the relevant fragments related to the extraction of the subject:

(3)a ... [$_{S'}$ that [$_{S}$ t_i ...

b ... [$_{S'}$ [t_i that] [$_{S}$ t_i ...

c ... [$_{S'}$ t_i [$_{S}$ t_i ...

If we extract a subject across an overt complementiser, as in (3a), the structure will not fulfill head government because the complementiser, although governing *t* and being of the proper kind from an X'-theoretical point of view, i.e. X°, is not lexical. The other option, as in (3b), namely antecedent government, will not be available as well because of the failure of the c-command relation between the trace which is Chomsky-adjoined to the complementiser and the trace in subject position: in fact, the former trace is immediately dominated by a branching node which does not contain the element to be antecedent governed. This interrupts the c-comand relation, ruling the sentence out. Nevertheless, it is just antecedent government which yields the only possible strategy to fulfill the ECP_1 in case of subject extraction in English. This is represented in (3c): notice that S does not block proper government because it is not a maximal projection, it is rather the nucleus of the anomalous clausal constituent which, in this system, reaches only the first bar level, namely S'.

The other side of the asymmetry in (1), i.e. extraction of an object across an overt complementiser, is immediately captured by assuming that V° fulfills head government, it being lexical as opposed to the complementiser:

(4) ... [V' [V° frightened] t] ...

Clearly, since this first version of the ECP was formulated, a sharp partition within the class of heads played an essential role, that is the one distinguishing *lexical heads* vs. *non-lexical heads*. Although the empirical content of this partition seems to be *prima facie* quite obvious, as our intuition tends to associate the *[±V, ± N]* elements among the set of heads creating a sort of natural class, once we try to look carefully at such a dichotomy, we cannot but realise that it is just merely postulated.

The aim of this paper is to focus on the role of this postulate in two major approaches, namely Chomsky (1986) and Rizzi (1990), and to explore the possibility of deriving it from more general principles. To do so we can begin with a brief survey of the major trends that stemmed from Chomsky's first formulation.

1.3. Unifying the ECP: a Theoretical Dilemma

Within an updated framework extending the X'-skeleton to non-lexical categories, including the complementizer periphery as in (5), Chomsky's approach had to be slightly reformulated. In *Barriers* (Chomsky (1986)), both the complementizer and the syncretic group of the inflectional elements (basically, tense and agreement features) have been represented as heading autonomous projections according to the general schema. Simplifying somehow the picture, we can represent a sentence in the following way: the inflectional morphemes, namely I°, select a verb phrase. This predicative complex combines with a subject which is structurally speaking the specifier the maximal projection

headed by I°, i.e. IP. This predicative nucleus, is on its turn the complement of another head, the complementizer head C°, which provides the periphery with position for operators like wh-elements and the connections for subordination, like the element *that*. The two heads C° and I° are called "functional heads" as opposed to V°, A°, P° and N° which are the "lexical heads." In general, the latter type is decomposed in a more abstract way as the intersection of two features: ±N, ±V. So that V = [+V,-N], N = [+N, -V], A = [+N, +V] and P = [-N, -V]. Nevertheless, all types of heads project in the same way according to the following general schema:. X, Y, Z being variables ranging on the set of heads, we have (5a-b). The representation of extraction from the subject position is now the one given in (5c) as opposed to (3a-d):

(5)a X' = X° YP

b XP = ZP X'

c ... [$_{CP}$ t_i [$_{C°}$ (that)] [$_{IP}$ t_i I° ...

First of all, the trace of the subject in (5c)is now immediately dominated by a maximal projection, namely IP, which does not dominate the antecedent. If we go back to our first formulation of the ECP, we must conclude that this trace can never satisfy this principle because it will never m-command its antecedent. In fact, the only head which governs it, namely the lower I°, is not lexical, thus it cannot properly govern such a trace. The next step was then to assume that only certain maximal projections can per se interrupt a government relation (i.e. can be *barriers*) and that IP is not among this group.

Moreover, the same extension of X'-theory yields a further problem: when the subject is extracted across an overt complementizer antecedent government is possible, because c-command is now available (as in (5c)). Recall that in the previous schema the problem was solved because the a new branching node was formed interrupting the c-command relation between the trace of the subject and the potential governor.

This new situation forces a further change in the formulation of the ECP. Since the ECP_1 consists of two distinct conditions, a natural option which immediately suggests it-

self is the one of extending one of the two conditions to cover the data explained by the other.[3] There have been several different approaches to this issue either in one sense or in the other: among all contributions we can single out two influential works which are representative of the two opposite trends, namely Chomsky's *Barrier* (Chomsky (1986)) and Rizzi's *Relativised Minimality* (Rizzi (1990)). The ideas which are developed in these two works will be briefly surveyed in the following subsections. In both cases we will not follow the entire analytical path which led the authors to the two different solutions, we will rather illustrate the final results, focusing on those elements which will be useful to support our thesis.

1.3.1. Chomsky's Notion of "Extended Chain"

One of the most relevant theoretical results of Chomsky's *Barriers* system is that "for a large range of central cases of the the ECP, the principle can be reduced to antecedent government and treated simply as a *chain phenomenon*" (Chomsky (1986) p. 79). Subject extraction in English is included within such a range; the trace in preverbal position is antecedent governed by the trace in spec-CP unless an overt complementiser intereferes in such a relation ruling out the sentence. The interference of an overt complementizer is regarded as a violation of a general principle of *minimality* which requires that the first potential governor is in fact the actual proper governor of a certain trace: thus, since C° is a head and since it governs the trace, the very fact that it does not belong to the class of lexical heads blocks all residual kinds of proper government, in particular proper government by the antecedent in spec-CP. Remember that to allow government of the subject trace in (5) one has to assume that IP is transparent to a government relation. This peculiarity of the IP projection, which is labelled as *defective* in *Barriers,* has its direct genetical antecedent in the older framework where IP was not represented as a maximal projection it being represented by an element without bars, namely *S*.

Although in Chomsky (1986) the inquiry is limited to just a subset of the cases which are potentially explained by the ECP, the process of antecedent government is extremely relevant because it is explicitly considered as potentially extendible to all cases, as originally proposed by "Richard Kayne and others under somewhat different assumptions" (Chomsky (1986), p. 80).

An interesting case for which this proposal has been realised is the following. We have so far concentrated on extraction of the subject, simply assuming that the trace of an object is always in a position to satisfy the ECP by head government, V° being lexical. Chomsky proposes to look at this phenomenon from a rather different perspective.

The fundamental asymmetry in (1) is a case of the so called A'-movement, i.e. the movement of an element to a position where a grammatical function such as subject, object etc. is not defined. The trace in the object position, called "variable", is connected to its antecedent in spec-CP playing the role of an operator (see May (1986)). The modular structure of grammar yields a first problem here. Even if the ECP is satisfied, these two positions are potentially offending Subjacency principle: in fact, in between the trace and its antecedent there are two maximal projections, namely VP and IP. How can we avoid this unwanted result? Chomsky's proposal is the following. Since VP is the complement of a non-lexical head, namely I°, adjunction to it is allowed (see Chomsky (1986), p. 16). This fact which enables extraction from VP inactivates also the potential barrierhood status of IP which can only be a barrier by inheritance, that is by dominating a barrier (see Chomsky (1986), p.88). The conclusion is that the link between spec-CP and the object position does not violate Subjacency. By relying on this independent strategy forcing an A'-chain of the object to undergo adjunction to VP, Chomsky reformulates the way in which the trace of an object fulfills the ECP. The proposal is that this trace is not licensed by the head V°, but rather by the maximal projection adjoined to VP itself, which in fact binds the trace, as an instance of antecedent government.

The next question is what happens w.r.t. the ECP when we face a case of A-movement? A-movement, in contrast with respect to A'-movement is defined as the movement of an element to a position where a grammatical function such as subject, object etc; is defined. The trace left by an instance of A-movement is called "anaphor". We can

reformulate the problem here by asking how an anaphor can satisfy the ECP. Now, there are two major cases of A-movement of an NP (briefly, NP-movement): raising and passive. We can represent instances of such two cases by means of the following formulae (reproducing Chomsky's (1986) examples (169) and (171)):

(6) a $John_i$ [$seem_j$-I°] [$_{VP}$ t_j [$_{IP}$ t_i to be intelligent]]

b $John_i$ [be_j-I°] [$_{VP'}$ t_j [$_{VP}$ killed t_i]]

Again, VP, as a maximal projection is a potential barrier. How is the ECP satisfied then? Notice that one cannot appeal to adjunction to VP, as would be in case the object would be wh-moved, because we would end up with a mixed chain of the kind (A, A', A) which for independent reasons we know to be ill formed. A solution to this problem is suggested by (6)a. Notice first that the subject *John* transmits its agreement features, here represented by the index *i*, to the verb. Now we have the following situation: the trace of the verb governs the trace of the subject, recall that IP is not a barrier to government. If we assume that the features of the verb are shared with the subject (i.e. i=j) we could say that the trace of the verb antecedent governs the subject trace satisfying the ECP: technically, t_j is said to be part of the *Extended Chain* headed by the subject. Although our description is only partial, the essence of Chomsky's theory is rather clear: the notion of "antecedent government" (normally performed by maximal projections) has been extended to include heads. What about (6)b then? The situation is parallel to the one in (6)a: also in this case, adjunction is not available, thus, to fulfill the ECP, one must appeal to other grammatical processes. Chomsky's answer in this case is based on the idea that the trace of the verb *be* creates together with the complement VP "a special case of adjunction structure" (Chomsky (1986) p. 76). Once one regards this structure in such a fashion, then government becomes available by the trace of *be* on the trace of the subject: again, proper government is derived as a sort of antecedent government given that by means of the process of index-sharing the trace of *be* is coindexed with the trace of the object of the verb *killed*; the potential barrier effect of VP is neutralised by the process of adjunction.

A very important remark is to be made at this point: although in these cases agreement is visible at the phonological level, this is not a necessary condition in Chomsky's system. Consider for example a case like:

(7) $John_i$ will be killed t_i

Here, *be* does not manifest any agreement with the subject, nevertheless it is regarded in exactly the same way with respect to the ECP: "there is (indirect) agreement between the subject and each aspectual verb of VP, as a reflex of spec-head agreement" (Chomsky (1986), p. 77).

Although *prima facie* Chomsky's system seems to make a step toward the unification of the complex framework, avoiding any appeal to government by a lexical head (labelled *θ-government*), it seems to me that it creates a rather non-homogeneous situation at least because of the two following reasons. The first reason involves the class of governors which act on the trace of an object . Even if in principle such a situation cannot be excluded, a strongly wellcome property of any version of the ECP is for a trace originated in a certain structural position to be licensed by the same kind of element for *all* kinds of movement: this is not the case here, since while anaphors (involving by definition A-movement) are licensed by heads, variables (involving by definition A'-movement) are licensed by maximal projections. The second reason is directly related to the central question we are addressing here: if overt agreement is not a necessary condition for a head to fulfill the ECP, why isn't this process available on an overt complementiser? More generally, which is the subset of spec-head relations which triggers agreement?

In other words, we see here that the old question of what makes a head lexical can be isomorphically reproduced in *Barriers* system as what makes a head part of a(n extended) chain.

Our contribution to the debate will be that of suggesting an answer to this question. To perform such a task we have to consider a second major trend regarding the updating of the ECP in a system which extends X'-theory to clausal projections. We can anticipate the

results by observing that the solution we are about to propose can also be regarded as a partial synthesis between these two different frameworks.

1.3.2. Rizzi's Reduction to Head Government

An essential push toward a modification of the ECP, as opposed to Chomsky's approach, originates in Rizzi's work as a theoretical disfavour for all principles which are stated in a disjunctive way: in Rizzi's terms, "admitting a disjunctive formulation amounts to admitting that the nature of the generalization is not understood" (Rizzi (1990) p. 76). After exploring the empirical consequences of a conjunctive formulation of the ECP, Rizzi reaches a simpler formulation which is made of just one single proposition. Let's synthesise it here as follows:

(8) the ECP_2: t must be governed by a head of a certain kind within its first projection

A cursory comparison shows that the two theories are poles apart, at least w.r.t. the terminology. In fact, if Chomsky concludes *Barriers* by arguing that the ECP is to be reduced to antecedent government, Rizzi reaches a final formulation where the same principle is rather stated in terms of head government. We will see, however, that the two theories have more than one single point in common.

Before approaching this version of the ECP a first important remark is to be made explicit: although antecedent government disappears from Rizzi's version of the ECP by no means is the empirical content embodied in it wiped out from the theory. It is rather subsumed under a different module. In particular, by borrowing the terminology from Osvaldo Jaeggli work, Rizzi distinguishes the two requirements which constitute the original formulation of the ECP from a conceptual point of view: on the one hand, antecedent government performs the *identification* of the antecedent of a trace; on the other, head government *formally licenses* the trace. From Rizzi's point of view the phenomena to be ex-

plained by the ECP are to be traced back to the presence/absence of a head of a certain kind in a local configuration w.r.t. a trace rather than to the possibility of identifying the content of the trace by means of a chain. That is, the ECP can be reduced to formal licensing. In fact, the identificational component still plays a crucial role within the whole theory: for example, it distinguishes the way in which operators can identify their variables. If a certain element is assigned an index (see Cinque (1990)b for a detailed theoretical treatment of this phenomenon based on the notion of "referentiality"; but against this terminology see also Moro (1991)b) its antecedent can be displaced as far as one wants in a single step (pace, of course, the other principles); on the other hand, if a certain element does not receive an index it could as well be displaced at an unbounded distance but only by means of successive local steps, that is by means of a chain of antecedent government relations.

To better understand Rizzi's version of the ECP we can see how it works in explaining the fundamental asymmetry in (1): a trace in object position is immediately licensed since it is governed by a lexical head (i.e. V°) within its first projection, hence essentially reproducing Chomsky's first system. How does a trace in subject position fulfill the ECP? The only possible candidate according to Rizzi's version of the ECP is C°, as opposed to Chomsky's system where the ECP is satisfied in this case by antecedent government. If the ungrammaticality with an overt complementizer follows here quite immediately, nevertheless the residual question is raised what is the proper governor when subject extraction is succesfull.

Notice that within this system the version of the minimality principle adopted by Chomsky is not to be maintained: since proper government is reduced to head government, there is no need to say that a head can interfere in antecedent government. More generally, Rizzi arrives at the conclusion that a potential governor can only interfere with a government relation of the same kind: that is, a head might only block head government while a maximal projection binding a trace might only block antecedent government. For this reason, Rizzi calls his approach "relativised minimality" as opposed to Chomsky's "rigid minimality".

Again the crucial question has to be formulated: what makes a head a proper governor? We have already implicitly assumed that V° is a proper governor. As usual, this

would derive from the fact that V° is lexical, that is V° belongs to the natural class of *[±V, ±N]* heads. What about C° then? We cannot consistently assume that empty C° belongs to this natural class. Rizzi's answer is based on a wide cluster of empirical data (Rizzi (1990), pp. 51-60) all leading to the same conclusion that in many languages when a subject is extracted across an overt C° it marks its passage through the spec of CP by triggering agreement on C°: Rizzi's proposal is then that Universal Grammar allows such a process in a larger scale than simply those cases which can be overtly detected. In the case of English, the extraction of a subject across an empty C° does realise such a process of abstract agreement while the same phenomenon is not available if C° is overt, like *that*.

To represent the analysis of (1) we can focus on the following segments:

(9)a ... [$_{C'}$ $C°_{+Agri}$ [$_{IP}$ t_i ...

b ... [$_{V'}$ V° t ...

All in all, this amounts to admitting that within the class of propers governors one has to include along with the natural sub class of lexical heads also the element C° (of course only if it contains the +Agr feature). Still, although the advantages coming from a simplified version of the ECP are clear, we have to rely on an heterogeneous catalogue of heads lacking any internal unity; the theory is not so far able to *explain* why a head can perform the task of proper governor, the only thing that it produces is a mere *list* of elements whose common nature seems to elude a welcome empirical generalization. We might also point out that within this system it is not simply $C°_{+Agr}$ which is to be added to the natural class of lexical heads to constitute the class of proper governors: to account for the wh-movement from VP adjoined position in Italian, Rizzi cannot count on V°, as was assumed in previous works (cf. Rizzi (1982)), since for independent reasons he is restricting the area in which a proper governor is active to the first projection. The solution that he offers comes from the assumption of the more articulated clausal structure where I° has been split.[4] The proper governor for the trace of an inverse subject is then T°.[5] Again, this does not seem to clarify the essential characteristic of a proper governor.

It might also be noticed that the inner nature of a head plays also another important role within this framework. As we saw in the two different systems proposed by Chomsky in the *Pisa Lectures* and in *Barriers* , the inner clausal boundary, roughly the sentence without the complementiser periphery, has a particular property of allowing government from outside. Also in Rizzi's system it is in principle assumed that a maximal projection is a barrier to government but the situation is solved in rather different way. To do so Rizzi offers two possibilities (Rizzi (1990) p. 113): either one extends Chomsky's analysis of extraction from VP by allowing adjunction also to IP as a complement of a non-lexical head, i.e. C°, or one follows Cinque's (1990) powerful generalization which we will simply report here: "an XP is a barrier if it is not directly selected by an X° not distinct from *[+V]*" (Cinque (1990), p. 112), assuming, of course, that C° and I° are not distinct from *[+V]*. Although we will not discuss the empirical motivation for assuming one proposal or the other it is sufficient to our purpose to notice that also in this case the notion of "head" appears to be a mere label for a quite more complex epiphenomenon. Within the class of heads different underlying partitions play important roles, at least the one of being a proper governor and that of neutralising the barrierhood status of a maximal projection.[6]

In the following section we will try to suggest a proposal for a partial characterization of the notion of "proper governor" which is based on two different facts: an empirical case which forces a refinement of Rizzi's version of the ECP and the enriched clausal structure as proposed in Chomsky (1989).

2. Generalised Head Governors

Summarising, we have seen how the original dichotomical principle proposed by Chomsky in the *Pisa Lectures* has been reduced to a unified principle in the two immediately possible directions: in *Barriers* the ECP is essentially regarded as a chain phenomenon, i.e. it is reduced to antecedent government; on the other hand, in *Relativised Minimality* the ECP is progressively modified to reach the form of head government. In

both theories the notion of "head" is a central and quite complex one: in particular proper governors are only a subset of the set denoted by such a label.

A unified theory of copular sentences as proposed in Moro (1988) offers an empirical case which allows us to take a step forward toward the answer to the question addressed here. In particular copular sentences provide us with a empirical case to test Rizzi's version of the ECP. Although we will not discuss this case in terms of Chomsky's system, it will be eventually possible to draw the major lines toward a unification of the two approaches.

2.1. Resplitting the ECP

The following pair of sentences constitute a quite anomalous departure from a general pattern:

(10) a a picture of the wall was the cause of the riot
b the cause of the riot was a picture of the wall

In general the permutation of the two NPs in a sequence of the kind *NP V NP* yields a very different result: in particular such a permutation preserves here the thematic interpretation of the sentence, as opposed to all other cases involving transitive verbs. This is a rather trivial phenomenon: when a speaker says *Bill loves Mary* says something very different than if he says *Mary loves Bill*. When the verb is the copula this fact does not happen. Nevertheless, this simple fact, which I have elsewhere proposed to call *the anomaly of copular sentences* (Moro 1991)b), challenges the theory of clausal structure in a very sharp and clear way and enables us to formulate precise empirical questions. Why are the two combinations thematically equivalent? Is equivalence maintained w.r.t. the properties determined at s-structure? Leaving a more detailed analysis to cited works, let's assume that the sentences in (10) have the following representation:

(11)a $[_{IP}$ [a picture of the wall]$_i$ $[_{I'}$ $[_{I}$was] $[_{SC}$ t_i the cause of the riot]]]

b $[_{IP}$ [the cause of the riot]$_i$ $[_{I'}$ $[_{I}$was] $[_{SC}$ a picture of the wall t_i]]]

By *SC*, we mean the same kind of constituent that is the complement of verbs like consider or believe in cases like: *John believes a picture of the wall the cause of the riot*. A copular sentence, roughly speaking, can be regarded as a sort of expanded SC. If the theory in Moro (1988) is correct, then the thematic synonymity will follow immediately by the fact that the two sentences do share the same d-structure, which is by definition the level where identical thematic relations receive identical representations. The answer to the second question, instead, will not be immediate: it can be shown that in spite of their surface resemblance the two sentences display very different properties. Although many other major questions can be raised about the representation in (11), to our purpose it will be sufficient to raise only the following one: how is the ECP satisfied? Let's assume Rizzi's version of this principle and follow the two cases separately.

For the first sentence the answer seems to follow immediately: the copula being a verb, thus a V° element, it can perform the task of proper government within its first projection. Moreover, being non-distinct from a *[+V]* element it neutralises the barrierhood status of SC. The tentative answer would then be that the ECP is satisfied by the copula, paralleling the case of object extraction. Although from a structural point of view no exception can be raised here, the assumption that the copula is a lexical head cannot be consistently maintained. Even if from a morphological point of view the copula is indeed a verb, i.e. a V°, it does not share with the other verbs many of the empirical properties which qualify these elements as lexical. For example, it can be shown that the copula is not able to license *ne*-extraction in Italian as opposed to transitive verbs. Take for example the Italian equivalent (roughly its gloss) of the sentence (11b):

(12) la causa della rivolta era una foto del muro

If one tries to extract *del muro* (of-the wall) by means of the clitic *ne* (of-it) the result would be strongly ungrammatical:

(13) * la causa della rivolta$_j$ [ne$_i$ era] [[una foto t$_i$] t$_j$]

It can be proved that there are no structural reasons which prevent such a process but that it is precisely the nature of the copula which does not permit it. In fact, once an element like *ci* (the Italian equivalent of English *there*) is cliticised over it, *ceteris paribus*, *ne*-extraction becomes possible:

(14) pro [ce$_j$ n$_i$'era] [[una foto t$_i$] t$_j$]

Within the analysis offered in Moro (1990) this surprising fact could be explained by assuming a radical change of perspective w.r.t. the so called phenomenon of *there*-insertion. In particular I have proposed to interpret those elements like *ci* and *there* not like semantically null place holders of the subject (in short *expletives*) but rather as predicates themselves originated at d-structure within SC as normal predicates like the one in (11b). Within this framework, then, the clitic *ci*, plays a role parallel to that of V° when it melts with I°: the major difference here is that I° is independently realised by the copula, as comparative grammarians like Meillet claimed following the Aristotelian tradition. Since we would say that the lexical status of an inflected verb is given to it by the V° component, rather than by I°, then we could correspondingly say that the copula has been *lexicalised* by *ci* (see also footnote 6).

If this analysis is true for (10a), it must be also true for (10b): thus, we are forced to find another proper governor. A solution to this puzzle comes to mind once one realises that the structural situation in which the trace must be licensed is equivalent, *mutatis mutandis*, to the one of the preverbal subject: if this analogy proves tenable, then the proper governor for the trace within SC would not be the copula as a verb but rather the agreement

features contained in it. This would entirely parallel the strategy of subject extraction in English across a phonologically null complementiser:

(15) a ... $[_{CP}$ [which picture of the wall]$_i$ C°$_{+Agr\ i}$ $[_{IP}$ t_i ...

b ... $[_{IP}$ [a picture of the wall]$_i$ was$_{+Agr\ i}$ $[_{SC}$ t_i ...

We are left now with the residual problem of explaining why SC is not a barrier to government. Notice that we cannot simply solve this problem by relying on the lexicality of the copula, because we have seen that this would give us inconsistencies w.r.t. the technical use of the term we are adopting here. We have some options here: a first proposal is to say that SC has the same property of IP, i.e. it is defective à la Chomsky, i.e. it can be a barrier only by inheritance. A second possibility is to say that since it is not the complement of a lexical head it allows adjunction, paralleling extraction from VP. A further possibility is to derive this property from the fact that SCs are not projected by a head[7]: see the end of 2.2. for a sketchy proposal along this line.

We can conclude that in both cases Rizzi's version of the ECP yields the right predictions, with the important *proviso* that it is not just the copula as a verb which can perform the role of proper governor but rather the agreement features contained in it.

It is now possible to turn back to the central question addressed here. Let's focus for the sake of simplicity on the sentence in (11b) which we will call *inverse sentence* to distinguish it from the other one where the order of predication is the canonical one, *canonical sentence*. Here we have the following structural situation: there is an NP immediately contained within SC which is governed by a head agreeing with it within its first projection. The legitimate prediction, according to Rizzi's version of the ECP, is that a trace in this position can be licensed, at least as far as proper government is concerned. This prediction is false. Although it can be proved that this is true for all levels of representation we can limit ourselves to support this claim by means of the following simple case:

(16) *which picture of the wall$_i$ do you think the cause of the riot $[_{V'}$was$_i$ $[_{SC}$ t_i t]]

Notice that there is nothing wrong with the structural position of $t_{i:}$ in a canonical sentence the same trace would have been properly licensed. There must be something in the structure of the inverse sentence as a whole which prevents the subject to be moved across the predicate raised to spec-IP. Moreover, one cannot even appeal to the "referential" status of the NP involved as a reason for requiring the more restrictive form of proper government, i.e. antecedent government, instead of head government, as is suggested in Rizzi (1990) for non-referential elements. In fact, the element which is moved is a subject, thus a fully referential element, and in principle it is expected not to require successive cyclic movement (see Cinque (1991) for a detailed discussion).

A reasonable explanation can be given as follows. An inverse sentence is created whenever the predicate is raised to precopular position: to allow such a process, a trace has to be licensed by proper head government. As we have seen, the agreement features contained within the copula are the only one which can perform such a role. The situation is much more perspicuous now: a plausible analysis which comes to mind to interpret the impossibility of moving the subject of an inverse sentence is that there are no proper governors available to license its trace or, equivalently, that the structure yields an the ECP violation. This amounts to saying that for the ECP a *uniqueness condition* holds between proper governors and the traces which are to be licensed: the *ad hoc* character that someone might object to this proposal can be easily overcome once we think about the functional correlate of the ECP. If this device has the role of formally licensing traces, it must necessarily identify some structural relation: since structural relations are not ambiguous in syntax, i.e. a complement is the complement of only one head etc., then one would expect that each trace (thus, each position) has to be identified by one and only one head. Although empirical evidence would be required to support such an assumption it seems to me that it rather follows from a theoretical point of view as a necessary requirement.

Leaving this discussion aside we have a clear consequence: Rizzi's version of the ECP has to be partially reformulated. The new fact which must be accounted for here is that when the role of proper governor is played by the agreement features contained in a head,[8]

we must make it sure that that head licenses exactly one trace.[9] One possibility to capture this requirement is to assume that the spec of the governor must be directly activated in the relevant sense by making the chain of the element pass through it, producing that particular instance of a spec-head relation called "spec-head agreement". This can also be regarded as the counterpart within Rizzi's system of Chomsky's requirement that feature sharing is "not accidental" (Chomsky (1986), p.77). Thus, the formulation of the ECP (disregarding here T°) can be revised as follows:

(17) the ECP_3 : t must be governed by a head

(i) within its first projection if the head is lexical

(ii) by activating its spec if the governor is + Agr.

In other words, we might say that if the governor is a head containing agreement, we cannot just check the "first projection" of such a head: this is only a necessary but not a sufficient condition.

This refinement has two theoretical consequences: first, it distinguishes lexical heads from the one containing +Agr w.r.t. the structural environment involved in proper government; second, it splits the ECP_2 in a disjunctive formulation wiping out the efforts which led Rizzi to assume the simpler form. Since we agree with Rizzi's methodological assumptions regarding disjunctive principles as essentially non-adequate from an explanatory point of view, the challenging question is added to the major one whether the ECP_3 can be restated in a simpler (monadic) form.

2.2. Heads as Antecedents: the role of AGR°

From an abstract point of view there are two options now: again, either we extend one of the two conditions of the ECP_3 or we radically rethink the ECP in a completely different

way. I would like to suggest here that it is indeed possible to preserve the spirit of the ECP as it has evolved up to the ECP_3 by following the second strategy.

Let's assume a more articulated representation of the clausal structure as it has been proposed by Chomsky (1988), see also footnote 4 here. Two new assumptions are required with respect to the *Barriers* system: first, agreement heads its own projection; second, agreement is not exclusively related to preverbal subject. By combining these two different statements we obtain the following temptative (partial) representation of the clausal structure (for simplicity, IP = AGR_2P; for TP, i.e.Tense Phrase, see again footnote 4):

(18) ... $AGR°_1$ [$_{IP}$ ... $AGR°_2$ [$_{TP}$ ... $AGR°_3$ [$_{VP}$ ...

A detailed discussion of this structure would take us too far: we can limit ouself to a weaker position, namely we can show that for all positions we have indicated here by AGR° there are cases where agreement can be overtly detected.

The first AGR°, which has been discussed at lenght in *Relativised Minimality* (Rizzi (1990) p. 51-60), is the one that we can find in languages like French where the complementiser *que* becomes *qui* when its spec is hosting the trace of the subject. The second case is verbal agreement with preverbal NP: it is generally assumed that it can be occupied only by the subject or by an expletive but, if we are right in assuming Moro's (1988) unified theory of copular sentences, then it can also be triggered by a predicate, of course when it is realised by an NP. The third case is the one where the object triggers agreement on the verb: although it is less frequent in most *familiar* languages it is nevertheless well known. For example, R. de la Grasserie (1900)) studied this phenomenon naming it "objective conjugation" and Hjelmslev said that "c'est, en effet, un de ces phénomènes grammaticaux qui semblent pouvoir naître partout" (Hjelmslev (1928), p. 144). It can be found for example in Basque, or in the group of Uralic languages. Hjelmslev goes a step forward by noticing that "l'accord qui s'est ainsi établi entre l'objet et le verbe transitif est analogue à celui qui s'établit, en d'autres langues, entre un substantif et son adjectif-épithète à l'égard du cas, du nombre, ou de la détermination" (Hjelmslev (1928), p. 146).

Hjelmslev's observation that agreement with the object has nothing special, it being just an instance of an independent phenomenon, suggests a wider generalization: all instances of agreement are in fact instances of the *same* agreement. This fits into our representation in (18): all occurrences of AGR° are identical. The differences which obviously show up at the surface level must be traced back to the interaction of independent principles. The AGR° elements would combine in the course of derivation with the other heads yielding *complex molecules* which may end up looking very different, according to the configurations allowed in each particular language by Universal Grammar. To borrow an image from biology one could regard these occurrences of AGR° as body cells which although potentially identical w.r.t. their genetical endowment differentiate themselves by expressing characteristic properties as determined by the local environment. So, for example, according to Chomsky (1988) when AGR° is adjacent to T° it assigns the nominative case, when it is adjacent to V° it assignes accusative and so on and so forth.

Of course, however suggestive, this metaphor does not tell us whether (18) is a true representation of the clausal structure. A well formed question would rather be the one of determining for which values of X is the following segment possible:

(19) ... AGR° XP ...

This would be a major task which we cannot pursue here, nevertheless we can conclude that the representation in (18) is essentially correct at least as far as the two assumptions which have been made about the distribution of AGR° prove tenable.[10]

By maintaining this reasoning as a guideline we can now go back to our main proposal, namely that of reducing the ECP_3 to its second condition. The intuitive idea which I would like to develop here is that the ECP is *always* associated with a form of agreement, whether or not it is overtly realised.[11] We can temptatively assume the following informal version:

(20) the ECP_4: when an element is moved it must locally discharge its features

Let's see first how we can implement this idea within the system. Technically, we can stipulate that a trace "discharges" its features if and only if the head which governs it agrees with it. Since typically this phenomenon is realised under a spec-head relation, we can assume that an element discharges its features if and only if it activates a spec-head agreement on a head governing it. More accurately, if we assume that agreement always heads its own projection, we should refine this approach by saying that an element discharges its features if and only if the head governing it can be raised to a local AGR° head whose spec can be activated by the chain of the same element. Let's see how the system works in the case of the fundamental asymmetry (the AGR° here is $AGR°_3$ of the abstract schema in (18)):

(21) a ... $[_{CP}\ t_i\ \ AGR°_i\ [_{IP}\ t_i\ \ ...$

b ... $[_{AGRP}\ t_i\ \ V°_j\text{-}AGR°_i\ [_{VP}\ ...\ t_j\ \ t_i\ \ ...$

In both cases the trace to be licensed is the rightmost t_i . Let's consider the two cases separately.

The first case reproduces Rizzi's theory: for the sake of clarity we will simply assume that C° "contains" AGR° rather than explore the possibility of representing this phenomenon as a case of head-to-head movement. Notice that we have to restrict government relation because otherwise AGR° adjacent to TP might act as a proper governor in English according to the ECP_4; in other words, we must not allow verbal agreement in English to satisfy the ECP. This is a long standing problem and one which in the late seventies led to the formulation of "parameter" (namely the *pro-drop* parameter, see Chomsky (1981) and Rizzi (1982)). Although it would be very tempting now to trace it back to the idea that the subject is generated in spec-VP (see footnote (5), here) and the fact that in English the inflectional morphemes are lowered at s-structure (as shown in Pollock (1989)), this reduction does not seem immediately tenable. "Unfortunately", we have a language like French where V° raises to I° at s-structure but still the complex molecule which is formed is inactive w.r.t. the property of being a proper governor of the trace of the subject. So, we

can simply limit ourself here to the assumption that government involved in the ECP is in fact to be intended à la Rizzi as government within the first projection.

The second case, where the object activates AGR° adjacent to VP, is particularly interesting if we assume the hypothesis that the subject is generated within VP. If this is true, why couldn't the subject be raised to this AGR°? Metaphorically, how can the subject and the object recognise *their* AGR° (recall we are assuming all AGR° to be identical)? In fact, we might assume, they don't. Suppose that the subject is raised from VP to trigger the spec of the AGR° which has the same VP as a complement; then, in order for V° to raise to T° (or alternatively for T° to lower) it would be impossible because another head will stand in the middle. In other words, this would lead to a Head Movement Constraint (HMC) violation (cfr. Travis (1984)), requiring heads not to skip any head along their trajectory. This is a welcome result which supports our point of view: it is the general design of the structure which *decides* the combination of AGR°s. If this proves tenable it is not necessary to require for AGR° to be local in the ECP_4: this would follow from the overall design of grammar.

If we now look at the situation from a non-technical point of view we can see the advantages that it offers. First, we can now better understand the main issue concerning the nature of proper governors: all heads are equal with respect to this property, their ability to perform the role of proper governor is to be related to the environment in which they happen to occur, that is to the possibility of combining with a local AGR°. Thus, the heterogeneous catalogue containing all lexical heads plus AGR° (disregarding once more T°) can be entirely avoided in favour of a different way of looking at things. The burden of proper government has shifted from heads to structures, provided that a theory of the distribution of AGR° would be explicitly given. This seems to fit very naturally in the spirit which originated the ECP: recall that since its first formulation, the ECP has been intended as a device to ensure recoverability of traces, both with respect to their structural position and to the identification of their content. More generally, "it is not unreasonable that Universal Grammar should require that the presence of an empty category be signalled in some manner by elements that are overtly present" (Chomsky (1981), p. 251): we propose

here that this *signal* is AGR°, of course combined at s-structure with other elements as required by the interaction of independent principles. The old notion of "lexical head" as a proper governor was not wrong, it was only hiding the more abstract underlying reality which can now be brought to light as an off side effect of the general development of the theory. The weird addition of +Agr to the list of lexical heads can now be absorbed within a more abstract framework. Again, an image borrowed from natural sciences might be useful to understand the situation: it is as if AGR° were the *particle* which mediates the *action at distance* the ECP is about.

This point of view suggests also that the incompatibility of Rizzi's theory with Chomsky's one might be not radical if not only terminological. We have already seen how our proposal refines Rizzi's version of the ECP, we can now see which is the relevance of the same proposal w.r.t. Chomsky's theory.

Recall that *Barriers* concludes by suggesting that the ECP is to be regarded as a chain phenomenon, i.e. it can be reduced to antecedent government. We have seen that if we look more carefully at this conclusion we see that the notion of "head" still plays a crucial role. All instances of A-movement are in fact allowed (as far as the ECP is concerned) by a head: raising, by agreement of the matrix verb with the chain headed by the subject, and passive, by reinterpreting the modal as an adjunct to the VP with the auxiliary assumption that agreement might be abstract, i.e. non-overtly realised. This situation contrasts with A'-movement where the role of proper governor is performed by maximal projections.

If we adopt the hypothesis proposed here, Chomsky's analysis of A-movement can be reinterpreted and extended in a rather natural way. The trace of an object is licensed in all types of movement by a head which combines with the AGR° taking VP as a complement: there is no need to say that wh-movement is licensed by antecedent government if we can assume that AGR° is always present, even if it is not overtly realised. The trace of a subject in A' movement would also be licensed by a head containing the feature of the moved element, namely AGR°-in-Comp, as in the case of A-movement.

If this analysis is correct, then Rizzi's and Chomsky's theories become compatible since if it is true that the ECP can be traced back to the presence/absence of a head of a certain kind (as in Rizzi's proposal) it is also true that this head must contain some property

of antecedence triggered by a chain (as in Chomsky's proposal). This seems to interpret the original spirit of the ECP as expressed in the *Pisa Lectures*. If for a trace to be licensed we must have a *signal* of some sort, then a head *containing* AGR° in the proper configuration seems to be a plausible candidate to perform such a role. A proper governor is a head which can perform the role of antecedent of an empty category. Of course, one has to admit that agreement might not be overtly realised, but this is a common assumption as in the case of *Barriers* or *Relativised Minimality* proposals.

A final remark about the issue concerning the barrierhood status of maximal projections: we have seen that this is a rather murky question which has not been treated in a uniform way. Chomsky (1986) proposes to treat it in a complex manner involving adjunction processes and the notion of "defective system" for IP. Cinque (1990)b proposes a unified theory by saying that an XP is a barrier if it is not selected by a head not distinct from *[+V],* C° and I° belonging to this class. Rizzi (1990) leaves the alternative between these two options open. An alternative hypothesis is suggested by facts analysed here.

Suppose one assumes that the "governing area" of a head is derived by minimality conditions, i.e. every head governs everything it m-commands unless a closer head intervenes. Then a head would not in general be able to govern inside its complement unless some process has occurred to the head of it. Of course, for this process to be plausible, it must in principle affect also the higher head. A natural candidate is head-to-head movement: as a first approximation we can assume that when a head absorbs the head of its complement, the new complex head retains the governing properties of the complement, in particular the maximal projection which separates the new head from the spec of the complement becomes *invisible* to government relation. An essentially analogous proposal based on entirely different empirical issues has been made by Baker (1988: 64) and it is called the *Government Transparency Corollary*. Although we will not explore all the differences between these two devices, it should be noticed that if Baker's idea was explicitly limited to lexical heads, here we would try not to rely on the nature of the head involved, assuming that the observed differences can be traced back to the syntactic environment where they occur.

If this approach proves tenable, all our cases would be immediately solved with two potential exceptions, i.e. government of the trace of the subject by AGR° adjacent to IP (the AGR-in Comp à la Rizzi) and government of the traces contained in SC. Since we are assuming that barrierhood is due to minimality induced by the presence of a head, being SC a non-projected constituent, the problem is reduced to IP. To solve this puzzle one can extend a proposal suggested by Chomsky (1988) and, independently, by Law (1991) and Manzini (forthcoming). The idea to be developed is that V°-to-I° movement, although blocked in languages like English at s-structure, is indeed always active at LF. This process would be required by the intrinsic nature of Tense which can be regarded as an operator quantifying over a range of values. If this proves tenable, then one could also tentatively assume that this same characteristic forces movement of the complex V°-I° to an empty C°, a position which by definition can host operators: from this position the trace of the subject would be able to fulfill the ECP_4 because IP would not be a barrier to government any longer.[12] Of course, if this position is occupied by an overt complementizer this process cannot take place and the structure is ruled out. Notice that this presupposes that the ECP is a property of the LF level: since this hypothesis interacts in a rather unclear way with other cases we are not discussing here, we will not develop further this topic and simply extend Chomsky's idea by assuming that SC is defective on a par with IP.

At this point the major question addressed in the introductory remark has been answered: the empirical content of the ECP predicting the fundamental asymmetry is preserved by the new formulation and the underlying nature of proper governors has come to light. We have concluded that AGR° is the *particle* which mediates the *action at distance* the ECP is about.

3. Concluding remarks

The ECP has originally been stated in the *Pisa Lectures* as a binary principle. Several reasons, both empirical and theoretical, conspire in favour of a reformulation of this principle. A natural simplification suggesting itself is to extend one of the two conditions to cover the facts explained by the other. We have then surveyed two major and asymmetric proposals: *Barriers* system, where the ECP is essentially reduced (or reducible) to antecedent government and *Relativised Minimality*, where the ECP is traced back to the other branch of the original version, i.e. head government. We have analysed how the notion of "head" still plays an important role in both frameworks respectively via the notion of "extended chain" and the one of "lexical head".

In particular we have tested Rizzi's version of the ECP and asked whether the hetherogeneous catalogue of proper governors including *[+V]* heads, plus AGR°, plus T° could be avoided in favour of a generalised notion of "proper governor" based on some underlying property which is shared at a more abstract level by all members of such a list.[13]

A proposal has been made by forcing Rizzi's version of the ECP to account for some empirical facts related to the syntax of copular sentences. We first reformulated the ECP by adopting the more articulated clausal structure which has been proposed by Chomsky (1988) essentially based on two assumptions: (i) agreement heads its own projection; (ii) agreement is not only related to clausal subject. The property which is shared by members of the set of proper governors naturally comes to light if one assumes that a head governing a trace is a proper governor for that trace only if it can be raised to an AGR° which enters into spec-head agreement with the chain of the element which is moved. The locality condition restricting the structural environment where this process can take place would be derived from the general design of grammar.

Under this proposal the two different theories proposed in *Barriers* and *Relativised Minimality* are synthesised by the unifying idea that proper government is *always* performed by a head (as in *Relativised Minimality*) provided that it embodies in itself some property of antecedence (as in *Barriers)*, it being the possibility for a head to contain the features of the moved element. Paraphrasing Chomsky's first formulation in the *Pisa*

Lectures we can assume that it is not unreasonable that Universal Grammar requires that the presence of the origin of a chain be signalled in some manner by elements that are overtly present: it is AGR° which turns out to be the *particle* which mediates the *action at distance* the ECP is about..

If our analysis will prove tenable, then the property that is shared among the set of proper governors would come to light in a quite natural way. Of course, this analysis can only be considered as a possible point of departure of an accurate empirical exploration, evaluating the advantages that it offers within the whole system of grammar, as commonly required by any scientific inquiry.[14]

Footnotes:

(*) This work can only be conceived as a small comment to a major debate whose source is mainly related to the work of Noam Chomsky, Guglielmo Cinque, and Luigi Rizzi. I would like to thank them here for many discussions about this topic. A special thank to Giorgio Graffi, Michael Hegarty, Richard Kayne, Giuseppe Longobardi, Alec Marantz, Alessandra Tomaselli and Phil Branigan for checking my English.

(1) For an accurate and up-to-date terminological reference see Haegeman (1991). For c-command and m-command in particular, see Haegeman (1991): 122-125.

(2) In Moro (1991)b, evidence has been provided that the ECP and Subjacency cannot be completely subsumed under a more general definition of barrier. The empirical basis of this conclusion is also briefly reported here in footnote (6).

(3) Of course, these two possibilities do not exhaust the entire range of options. In principle one can also explore a conjunctive formulation of the ECP, namely one in which both head government and antecedent government must be satisfied to license a trace. Along this lines, see Stowell (1981) and works cited in Rizzi (1990), p. 30. John Frampton (forthcoming), Rita Manzini (forthcoming) offer alternative analyses where the ECP is maintained as a binary principle. Interestingly, these two alternative approaches try to avoid the employment of a crucial notion which is introduced by Rizzi (1990) and refined in Cinque (1990), namely *referentiality*. For a discussion about this concept see also Kroch (1989) and Moro (1991a), the latter relying heavily on Geach (1962).

(4) We will soon directly approach this new proposal here in section 2.2.: the idea that the syncretic features contained within INFL°, namely Tense and Agreement, project independently has been proposed by Pollock (1989) developing some original ideas concerning the inflectional system across languages discussed in Emonds (1985) and the notion of "agreement" adopted by Kayne (1987). Instead of (1)a we should have (1)b:

(1)a ... $[_{IP}$ ... $[_{I'}$ $I°_{[\pm tense,\ \pm agreement]}$ ...

b ... $[_{AGRP}$ ... $[_{AGR'}$ AGR° $[_{TP}$ ... $[_{T'}$ T° ...

This proposal has undergone a major refinement in Chomsky (1989) and Belletti (1991). Moro (1988) arrived to the same conclusion reached by Pollock while analysing the syntax of copular sentences.

(5) Recent works inspired substantially by Koopman-Sportiche (1988), have normalised the X'-theory w.r.t. the relationship between arguments and lexical heads. It has

been proposed that the subject is generated within VP, in a promiment position w.r.t. the object. The essential idea is that the subject is generated in spec-VP. If this theory proves tenable, then one can explore the possibility that wh-movement of the subject in pro-drop languages does not start from a VP adjoined position, but rather from spec-VP. Then, Rizzi's proposal that T° is the proper governor for the variable left by the subject can be entirely avoided.

(6) It can be proved that these two properties are not *per se* necessarily embodied in a verb. For the sake of simplicity, let's consider two crucial cases (only relevant traces are represented here):

(1)a * [di quale muro]$_i$ credi che [la causa della rivolta]$_j$ fu [$_{SC}$ [$_{NP}$ una foto t_i] t_j]
of which wall pro think that the cause of-the riot was a picture

b [quale foto del muro]$_i$ credi che fu [$_{SC}$ [$_{NP}$ t_i] la causa della rivolta]
which picture of the wall pro think that was the cause of-the riot
"which picture of the wall do you think was the cause of the riot"

Taking this sort of paradigm as the major empirical fact to rely on, it was suggested that while the ECP is to be related to Agr, Subjacency is rather to be related to the selectional capability of a head according to some structural condition which have been explicitly indicated in Moro (1991)b. Since in general verbs embody these two properties simultaneously, this disjunction would go unnoticed unless copular sentences are analysed. To counterchek this hypothesis see discussion about the contrast between (13)-(14) in this paper.

Cinque's generalization is then to be slightly reformulated by spelling out the inner content of the [+V] features which solve barrierhood.

(7) In Moro (1988) SC was assumed to be a projection of AGR°. In fact this assumption was essentially related to the idea that the copula is the spell out of Tense and that by

small clause the predicative nucleus of a copular sentence was denoted: if the basic intuition supporting this version of the *Split INFL Hypothesis* is to be maintained (namely, that a predicative NP can occupy the position of subject of predication) some technical refinements are now required because AGR° is currently employed to indicate a morphological phenomenon which has nothing to do with predication.

In fact, we have empirical cases where agreement is unselectively triggered by preverbal NP disregarding its role in predicative linking like English (1a-b) (see Moro (1991) for an account of cross linguistic variations related to the pro-drop parameter) and cases where no agreement is required within a small clause, like Italian (1c):

(1)a the cause of the riot is/*are them

b they are/*is the cause of the riot

c ritengo [$_{SC}$ le ragazze bionde il vero desiderio di Andrea]

believe the blonde girls-fem.pl. the real desire-mas.sing. of Andrea

" I believe blonde girls Andrea's real desire"

Notice that (1c) is not to be related to *pro* licensing, since its equivalent with *pro* instead of the subject *le ragazze bionde* would be ungrammatical.

(8) Although we will not explore this topic in detail here, we may notice that the idea that the proper governor for the subject within the small clause is not the verb as a lexical element but rather the agreement features contained in it can be extended to those small clauses which are complements of lexical verbs. In Italian, for example, from a structure like (1a) we can derive (1b):

(1)a Gianni ha ritenuto [$_{SC}$ queste foto il peggior regalo]

"Gianni has believed these pictures the worst gift"

b Quali foto$_i$ Gianni ha [$_{AGRP}$ t$_i$ ritenut-e$_i$ [$_{SC}$ t$_i$ il peggior regalo]]

Which pictures Gianni has considered-fem.plur. the worst gift-mas.sing.

A major consequence of this approach would be that extraction from an embedded inflected CP is not subject to the ECP, at least as far as the trace in spec-CP is concerned, because in this case agreement with the matrix verb would be entirely ungrammatical. This implies that the ECP is not to be applied to every single step of a chain, that is to intermediate traces but rather to its very origin or equivalently that the ECP does not regard traces but rather chains. Since this hypothesis would induce a major departure w.r.t. current theory, we will leave this possibility aside here.

(9) This argument raises a natural question. The uniqueness condition says that a head can properly govern at most one trace: this interacts with the assumption that SC are not projected constituents. In fact, the combination of these two independent assumptions leads to the conclusion that small clauses can never be *empty* at s-structure. This prediction seems to be *prima facie* too strong: consider for example a case of double object construction *à la Kayne* (see Kayne (1987)) like:

(1)a Mary$_i$ was given [$_{SC}$ t_i a book]

b Which book$_j$ was Mary$_i$ given [$_{SC}$ t_i t_j]

Prima facie, this example seems to falsify the hypothesis of uniqueness. However, it seems to me that this conclusion is not entirely legitimate. In fact, there are at least two options to preserve uniqueness condition which I will simply indicate here. The first one is that *give* can be lexically decomposed along the lines suggested by Larson (1988): if this analysis proves tenable, then the costituent which is complement of *give* is rather to be considered as a "pseudo-SC" projected by a lexical head. This head would now be a plausible candidate to license (via its trace) the trace of the object. There is however a second option which anticipates some of the result offered here in section 2.2.. The idea is that if we assume the multiple AGR° hypothesis, then a plausible proper governor for *Mary*

would be AGR° associated with the copula while (abstract) AGR° associated with *give* would be the proper governor for the object. We will leave this option open here.

(10) Potentially, there are various strategies to approach the problem represented by (19). One interesting possibility is to assume a methodological approach which has already been fruitfully applied within generative grammar, as in the case of Binding theory: instead of providing a constructive theory which indicates the possible combinations, one can explore the opposite way, i.e. one can assume that AGR° can freely occur everywhere and that all ungrammatical combinations are to be ruled out by the interaction of independent principles.

(11) As Giorgio Graffi noticed (p.c.), this assumption which reduces proper government to agreement seems to have a rather sharp analogy with a conclusion reached within a very distant framework: "On conçoit que ce que nous venons d'appeler rection pure est identique à ce qu'on appelle généralement accord ou concordance" (Hjelmslev (1928), p.141). Furthermore, the analogy seems to be even stronger if one considers that objective conjugation is included within the set of phenomena called *rection pure*. We will not develop this analogy here.

(12) Notice that V_I°-to-C° does not *per se* solve the barrierhood status of IP. Were it true, then all Verb Second constructs would permit extraction from a lower subject in spec-IP, which is false.

Apart from this case, the barrierhood status of IP seems to be even more pervasive. In Moro (1991)b the following contrast has been noticed:

(1)a there aren't many girls / there are few girls

b aren't there many girls / are there few girls

Although in both cases is the quantifier *many* in the scope of negation only in the first case the sentence can be paraphrased by substituting *n't ...many* by *few*. At the moment I do not see any better explanation for this fact (and analogous argument can be constructed in Aux-to Comp costructions in Italian) than assuming that IP still is a barrier (by inheritance).

The idea that I° raises to C° has been independently formulated by Law (1991). The analogy which appears here between the two proposals is only superficial, though. In particular no commitment is made here about obligatory deletion of complementiser at LF as a consequence of the principle of Full Interpretation.

(13) That lexical elements do not behave homogeneously w.r.t. government relation is indeed a rather wide accepted assumption; a very interesting testimony from an historical point of view comes again from Hjelmslev (1928): he points out that "dans l'école de Basra [...] on apprend que la faculté d'etre régi est bien inherénte à tout mot, mais la faculté de régir est plus restreinte: elle ne se trouve pas dans le nom, mais seulement dans le verbe et dans la particule" (Hjelmslev (1928) p.158-159).

(14) We started our path by saying that the ECP aims to capture the fundamental asymmetry in (1) but it must be said that along with this core case the ECP has been successfully employed to other major cases. Although we will not discuss them here it might be useful to notice that the ECP is also employed to account for the following phenomena: movement at the LF level, wh-movement in relatives, adjunct extraction and movement of X° projections, namely head-movement. If most cases might be covered by the verson of the ECP we suggested in this paper (but see also footnote (8) in this paper), the following two observation should be spelled out here.

First, it is clear that head movement cannot per se be covered by the same principle (nor by Rizzi's (1990) approach) given that heads do not trigger agreement by definition, it being mediated by a structural relation of spec-head agreement which is only accessible to maximal projection. To account for head-movement, it seems that antecedent government ramains as the only possible candidate, at least within current frameworks.

Second, to account for the possibility of interpreting the universal quantifier in subject position of an embeddd clause, the ECP should be formulated in such a way to tolerate "very short" movements without proper head governors. Take for example, a sentence like (1):

(1) Someone thinks that everyone should leave

Here, the quantified subject *everyone* must move to receive an interpretation but cannot escape the clause. Since the simple possibility of deleting *that* at LF will spoil the whole system, we must conclude that *everyone* has been adjoined to IP, and that antecedent government relation is again necessary and sufficient in this case to rule the sentence in.

References

Baker, M.C. (1988) *Incorporation*, The University of Chicago Press.

Belletti, A. (1991) *Generalised Verb Movement*, Rosemberg & Sellier, Torino.

Chao, W. - Horrocks, J. eds. (forthcoming) *Levels, principles and Processes,* Mouton de Guiter, The Hague

Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky , N. (1986) *Barriers*, MIT Press, Cambridge, Mass.

Chomsky, N. (1988) *Some Notes on the Economy of Derivation and Representation,* ms, MIT

Cinque, G. (1990)b *Types of A' Dependencies*, MIT Press, Cambridge, Mass.

Emonds, J. (1985) *A Unified Theory of Sytactic Categories*, Foris, Dordrecht.

Frampton, J. (forthcoming) *The Fine Structure of Wh-movement and the Proper Formulation of the ECP*, in

Chao, W. - Horrocks, J. ed. (forthcoming).

Geach, P. (1962) *Reference and Generality*, Cornell University Press, Ithaca.

Giorgi, A.-Longobardi, G. (1991) *The Syntax of Noun Phrases*, Cambridge University Press, Cambridge,

UK.

Grasserie, R. de la (1900) *Conjugaison pronominale*, J. Maisonneuve Libraire Editeur.

Haegeman, L. (1991) *Introduction to Government and Binding Theory*, Blackwell, Oxford.

Hjelmslev, L. (1928) *Principes de Grammaire Generale*, Copenhagen.

Kayne, R. (1987) *Connectedness and Binary Brancing*, Foris, Dordrecht.

Kroch, A. (1989) *Amount Quantification, Referentiality and Wh-movement*, ms. University of

Pennsylvania

Larson, R. (1988) *On the Double Object Construction*, *LI* 19, 3, pp. 335-39

Law, P. (1991) *Verb Movement, Expletive Replacement and Head Government,* talk given at XIV GLOW
Conference, Leiden, The Netherlands.

Longobardi, G. (1980) *Connectedness, complementi circostanziali e Soggiacenza*, *Rivista di Grammatica*
Generativa 1980, pp. 141-185.

Longobardi, G. (1989) *Symmetry Principles in Syntax*, Clesp, Padova.

Manzini, R. (forthcoming) *Locality Theory: Toward a Unification*, in Chao, W. - Horrocks, J. eds.
(forthcoming).

May, R. (1986) *Logical Form*, MIT Press, Cambridge, Mass.

Moro, A. (1988) *Per una teoria unificata delle frasi copulari*, *Rivista di Grammatica Generativa,* 13, pp.
81-110.

Moro, A. (1990) *There-Raising: Principles Across Levels,* talk given at XIII Glow Conference, St. John's
College, Cambridge, UK.

Moro, A. (1991)a *The Raising of Predicates: Copula, Expletives and Existence*, XV MITWPL, MIT,
Cambridge Mass.

Moro, A. (1991)b *The Anomaly of Copular Sentences*, University of Venice Working Papers, 8.

Moro, A. (forthcoming) *There as a Raised Predicate*, in Chao, W. - Horrocks, J. eds. , Foris, Dordrecht.

Pollock, J.-Y. (1989) *Verb Movement, UG and the Structure of IP*, *LI*, 20, 365 -424

Rizzi, L. (1982) *Issues in Italian Syntax*, Foris, Dordrecht.

Rizzi, L. (1990) *Relativised Minimality*, MIT Press, Cambridge Mass.

Roberts, I. (1988) *From Rules to Constrains*, *Lingua e Stile*, XXIII, 3.

Stowell, T. (1981), *Origin of Phrase Structure*, Ph. D. Dissertation, MIT, Cambridge, Massachussets.

Travis, L. (1984) *Parameters and Effects of Word Order Variation*, Ph.D. Dissertation, MIT, Cambridge

Massachussets.