

Locality theory: Competing models of weak islands¹

M. RITA MANZINI

Abstract

This paper explores two questions, related to Chomsky's (1981) ECP, under which traces must have a proper government relation either to a head, or to an antecedent. The first question concerns the definition of proper head government. Section 1 argues that at least Rizzi's (1990) and Cinque's (1991) referential indices can be dispensed with once the LF syntax of wh-constructions is properly understood. The second question concerns the definition of antecedent government. Section 2 shows that antecedent based conceptions of locality, such as Rizzi's (1990) Relativized Minimality (cf. Chomsky's 1995 Minimal Link Condition) and non-antecedent based notions, such as Chomsky's (1986a) (rigid) Minimality, can be equivalent, at least for phrasal movement, given appropriate assumptions.

Chomsky (1981) proposes that under the ECP, traces must satisfy one of two requirements. Either they have a proper government relation to a head, or they have a proper government relation to an antecedent. This paper addresses two questions related to this proposal, by examining data concerning weak islands, as defined by Cinque (1991). The first question regards the definition of head government. Classical proposals in the literature include Chomsky's (1986a) theta-government, and Rizzi's (1990) or Cinque's (1991) referential indices; my thesis in section 1 will be that referential indices can be dispensed with once the LF syntax of wh-constructions is properly understood. The second question is whether antecedent-government makes use of an antecedent based conception of locality,

1. This article, corresponding to my original presentation at the GLOW Colloquium in Cambridge in the spring of 1990, is written in the same late GB framework as Manzini (1992), with which it overlaps. For a treatment of the same set of problems within the minimalist framework of Chomsky (1995), the reader is referred to Manzini (1997).

such as Rizzi's (1990) Relativized Minimality (cf. Chomsky's 1995 Minimal Link Condition) or of a non-antecedent based notion, such as Chomsky's (1986a) (rigid) Minimality. My answer in section 2 will be that the two are equivalent, at least for phrasal movement, given appropriate assumptions.

1. Against referential indices

As discovered by Huang (1982), the locality behavior of argument wh-movement differs from that of adjunct wh-movement. Thus while extracting an adjunct out of a wh-island gives an ungrammatical result, as in (1), extracting an argument produces a fundamentally grammatical sentence, as in (2):

(1) **How do you wonder [what to repair t_{what} t_{how}]*

(2) *What do you wonder [how to repair t_{what} t_{how}]*

According to Chomsky (1986a), the two classes of elements that are and are not sensitive to wh-islands are characterizable in terms of the notion of theta-government, as defined in (3). Under (3) objects of X are theta-governed, in that they are theta-marked by head X and sisters to it, while adjuncts are not. It is of course non-theta-governed elements that are subject to more stringent locality requirements:

- (3) β theta-governs α iff
 β is a head, β theta-marks α and β is a sister of α

According to Rizzi (1990), the notion of theta-government is empirically inadequate in at least three respects. First, adverbs that obligatorily cooccur with certain V's can be argued to be theta-marked by them, and hence theta-governed; however under extraction they behave exactly like adjuncts, in being sensitive to wh-islands, as illustrated in (4). Similarly, idiom chunk NP's can be argued to be theta-marked by a V, and hence theta-governed; but they again display an adjunct-like behavior with respect to wh-islands, as in (5). And finally, the same can be repeated of measure phrases, as in (6):

(4) **How carefully do you wonder whether he worded the letter?*

(5) **What headway do you wonder whether he made?*

(6) **How may kilos do you wonder whether he weighs?*

What Rizzi (1990) proposes is that the notion of theta-government should be replaced by a notion of referential index, where a referential index is licenced under theta-role assignment, but only if the theta-role is a referential one, as in (7):

- (7) A referential index must be licenced by a referential theta-role

The shift from the notion of theta-government to the notion of referential index allows the behaviors in (4–6) to be correctly predicted on the assumption that none of the wh-phrases involved has a referential index, since none of them is referentially theta-marked, though each of them is theta-marked.

Cinque (1991) extends the notion of referential index to cover contrasts of the type in (8–9) where quantified phrases headed by an existential and a distributive universal quantifier, *qualche* ('some') and *ogni* ('every') respectively, are extracted out of a wh-island, as well as the contrast between (10) and (11), where a D-linked and non-D-linked wh-phrase respectively, in the sense of Pesetsky (1987), are extracted from a wh-island. The first member of the pair is systematically more acceptable than the second, which can be accounted for, if existential quantifiers (such as *qualche*) and D-linked wh-phrases have a referential index, while distributive universal quantifier (such as *ogni*) and non-D-linked wh-phrases do not:

- (8) *Qualche museo, mi chiedo chi visiterà.*
Some museum myself I-ask who will-visit
'Some museum, I wonder who will visit.'
- (9) **Ogni museo, mi chiedo chi visiterà.*
Every museum, myself I-ask who will-visit
'Every museum, I wonder who will visit.'
- (10) *Che macchina ti chiedi come riparare?*
Which car yourself you-ask how to-repair
'Which car do you wonder how to repair?'
- (11) **Che diavolo ti chiedi come riparare?*
What devil yourself you-ask how to-repair
'What the hell do you wonder how to repair?'

This extension of referential indices implies a redefinition of the notion. (7) can remain as a necessary condition on the licensing of referential indices; however some other condition must be added, to the effect that a referential index must be licensed by referential content of some sort, as in (12):

- (12) A referential index must be licensed by referential content

It is then possible to claim that *qualche*-type quantifiers and D-linked wh-phrases have a referential content in the relevant sense of the term, which licences their referential index; but *ogni*-type quantifiers and non-D-linked wh-phrases lack referentiality and hence referential indices.

Here I want to suggest an alternative to both of the conceptions in (7) and (12), which upholds (3) or the variant of it proposed by Manzini (1992). Manzini (1992) defines a notion of K-government so that β K-governs α just in case β is a head and β Case-marks α , as in (13).

- (13) β K-governs α iff β is a head and β Case-marks α

Under (13), an argument wh-trace, as in (2), is K-governed, but an adjunct wh-trace, as in (1), is not. Still, leaving aside the core cases in (1–2), that are indeed accounted for under (13), it remains to be shown that the arguments in favor of Rizzi's (1990) or Cinque's (1991) notion of referential index, as in (4–6) and (8–11), can be handled by (3) and/or (13). Consider Cinque's (1991) evidence. As we just saw, while a phrase headed by the existential quantifier *qualche* can be extracted across a wh-island, as in (8), a phrase headed by the distributive universal quantifier *ogni* cannot, as in (9). Similarly, extraction of a D-linked wh-phrase, in the sense of Pesetsky (1987), produces well-formed results across a wh-island, as in (10). By contrast extraction of a non-D-linked wh-phrase is ill-formed across a wh-island, as seen in (11). Such contrasts are accounted for under Cinque's (1991) theory by assuming that *qualche*-type quantifiers and D-linked wh-phrases can have referential content, but *ogni*-type quantifiers or non-D-linked wh-phrases cannot.

Cinque (1991) furthermore shows that Italian bare quantifiers such as *qualcuno* ('someone') are either associated with a resumptive clitic under Clitic Left Dislocation, as in (14), or else are not associated with a resumptive clitic and hence bind a variable, as in (15). When construed as in (14), *qualcuno* is not sensitive to wh-islands, though it is when construed as in (15), as shown by the contrast between (16) and (17). In order to capture this behavior, Cinque (1991) has recourse to the assumption that bare quantifiers can be used referentially or non-referentially; a bare quantifier used non-referentially will behave as an operator, binding a variable, and will not be able to cross wh-islands:

- (14) *Qualcuno penso che lo troveremo.*
 Someone I-think that him we-will-find
 'Someone I think we will find.'
- (15) *Qualcuno penso che troveremo.*
 Someone I-think that we-will-find
 'Someone I think we will find.'
- (16) *Qualcuno mi chiedo chi lo troverà.*
 Someone myself I-ask who him will-find
 'Someone I wonder who will find.'
- (17) **Qualcuno mi chiedo chi troverà.*
 Someone myself I-ask who will-find
 'Someone I wonder who will find.'

Finally, Cinque (1991) considers scope reconstruction phenomena of the type in (18–19), originally discussed by Longobardi (1991). In (18) the wh-phrase can be interpreted as having narrow scope with respect to the embedded quantifier; in (19), however, which differs from (18) only with respect to the presence of a wh-

island, the narrow scope reading of the *wh*-phrase is impossible. Cinque (1991) explains this by assuming that interaction in scope with another quantifier requires the *wh*-phrase to be non-referential, even though it can be referential when it does not interact with another quantifier:

- (18) *Quanti pazienti pensi che ogni medico visiti?*
 How-many patients you-think that every doctor visits
 'How many patients do you think that every doctor visits?'
 (19) *Quanti pazienti ti chiedi come ogni medico visiti?*
 How-many patients yourself you-ask how every doctor visits
 'How many patients do you wonder how every doctor visits?'

All of the data in (8–11) and (14–19) raise a potential problem for the theory in (3) and/or (13). Indeed all of the phrases moved in (8–11) and (14–19) are apparently theta- and K-governed, so that theta- and K-government appear to be unable to predict any of the relevant contrasts. Consider however the further paradigm in (20–23). Sentences of the type in (20), where no *wh*-island is crossed, are perfectly well-formed in Italian with extraction of *quanti* ('how many') and stranding of the clitic *ne* ('of them'); but sentences of the type in (21), where *quanti* is moved across a *wh*-island, again leaving behind *ne*, are ill-formed. Examples of the type in (22) where *quanti* does not strand the nominal part of its phrase are also well-formed; in this case extraction out of a *wh*-island appears to be acceptable, as in (23):

- (20) *Quanti pensi che ne visiterà?*
 How-many you-think that of-them he-will-visit
 'How many of them do you think he will visit?'
 (21) **Quanti ti chiedi chi ne visiterà?*
 How-many yourself you-ask who of -them will-visit
 'How many of them do you wonder who will visit?'
 (22) *Quanti pazienti pensi che visiterà?*
 How-many patients you-believe that he-will-visit
 'How many patients do you believe that he will visit?'
 (23) *Quanti pazienti ti chiedi chi visiterà?*
 How-many patients yourself you-ask who will-visit
 'How many patients do you wonder who will visit?'

Under the theory of Cinque (1991) it can of course be assumed that the bare quantifier *quanti* is not referential, thus predicting the ill-formedness of (21). In this case however a solution is also available under (3) or (13). Given usual assumptions, the DP *quanti-ne* is Case-marked and hence K-governed by the embedded V, as well as theta-marked and hence theta-governed. On the other hand, *quanti* presumably corresponds to the D head of DP. If so, we only need to say that

Case- and theta- marking always is to a maximal projection, to predict that a D head is not K- and theta-governed respectively. Hence the fact that *quanti* is subject to wh-islands, as in (21), is correctly predicted without recourse to referential indices.

Let us now consider the examples in (22–23) more closely. While they both are well-formed, they do not necessarily have the same interpretation. Suppose then that the two options available in the syntax, in (20) and (22–23), are available at LF as well. In other words, either the whole DP, or only its head D, can be in operator position. This means that there are two LF's corresponding to examples of the form in (22–23). The first LF straightforwardly corresponds to the surface configuration, as in (24); the second LF corresponds to a configuration of the type in (20–21) where the quantifier head has been raised out of its DP and the NP predicate reconstructed in the trace position, as in (25):

- (24) a. [*How many patients*: x] you think that he will visit x
 b. [*How many patients*: x] you wonder who will visit x
 (25) a. [*How many*: x] you think that he will visit x patients
 b. [*How many*: x] you wonder who will visit x patients

In the LF's in (24) the DP is of course Case- and theta-marked and hence K- and theta-governed by the embedded verb. Consider on the other hand the LF's in (25). D is not K- or theta-governed by the embedded verb, on the assumption that Case- and theta-marking are always to a maximal projection. Therefore we predict that an example like (22) can be associated with both the interpretations in (24a) and in (25a), while an example of the type in (23) can only be associated with the interpretation in (24b), since (25b) is blocked by the wh-island. If this is correct, the hypothesis that the two different LF's in (24) and (25) are possible is also confirmed; for a different range of interpretations for (22) and (23) is now derived from this hypothesis.

With this background we can return to the original problems in (8–11) and (14–19). Consider first (8–9). The contrast between (8) and (9) can be explained without any appeal to referentiality, if we assume that an *ogni*-type quantifier, as in (9), is compatible only with an LF at which D raising/ NP-reconstruction has taken place. For if so, D is not K- or theta-governed, hence extraction across a wh-island, as in (9) is ungrammatical. On the other hand, if a *qualche*-type quantifier is compatible with an LF at which the whole DP is in operator position, the grammaticality of (8) also follows. For, the whole DP is K- or theta-governed and hence it can extract across a wh-island.

Similarly, consider (10–11). The claim of the present theory must be that the only interpretation for the non-D-linked wh-phrase in (11) corresponds to an LF characterized by D in operator position. If so, the impossibility of crossing a wh-island, as in (11), reduces to the fact that D is not K- or theta-governed. On the other hand, if we assume that the LF under which the whole DP is in operator

position is also possible for the D-linked wh-phrase in (10), then no wh-island violation is predicted to arise, because DP is K- or theta-governed.

As for (14–17), no stipulation about the same quantifier being either referential or non-referential is needed under the approach being explored here. Instead we can simply assume that a quantifier such as *qualcuno* can move either as a head or as a maximal projection. If it moves as a maximal projection DP, it will be K- or theta-governed and able to cross wh-islands, as in (16). If it moves as a head D, it will not be K- or theta-governed, hence it will not be able to cross wh-islands, as in (17). Finally, with respect to (18–19) we can assume that scope interaction between quantifiers requires an LF with a D-operator. Because of the lack of K- or theta-government on its trace, this LF is possible only in the absence of a wh-island. This means that scope interaction between the two quantifiers is predicted to be possible in (18), but not in (19), correctly.

Crucially, the theory proposed here is no more complex than Cinque's (1991). The option of raising either D or DP must be available in all theories, given examples of the type in (20–23). What the present theory furthermore says is that there are two possible interpretations in examples of the type in (22–23), which correspond to the two LF's in (24) and (25). Via the notion of referential index, the relevant interpretations are identified by Cinque (1991) as referential and non-referential respectively. However, since referential content is not independently characterized by Cinque (1991), referential index theories have no explanatory advantage over the one proposed here.

The potential problem posed by Rizzi's (1990) selected adverbs, idiom chunks and measure phrases, as in (4–6), remains still to be taken into consideration. Consider first measure phrases, as in (6). All that is required under Rizzi's (1990) theory to explain the behavior of measure phrases is the assumption that measure phrases are not assigned a referential theta-role, so that they do not have a referential index. Their movement patterns follow then correctly. By contrast the present theory apparently predicts examples of the type in (6) to be well-formed, on a par with examples of the type in (23). One hypothesis that we can pursue in this respect is that (6) is ill-formed to the extent that it is not only incompatible with a reconstructed LF of the type of (25) for syntactic reasons, but also incompatible with a non-reconstructed LF of the type of (24) for semantic reasons. Indeed a parallel assumption is necessitated by the conception of referential indices of Cinque (1991), as opposed to Rizzi's (1990). Under (12), as opposed to (7), it must be a semantic property of measure phrases that they are not associated with referential indices, so that (6) is blocked.

Consider then idiom chunks, as in (5). The fact that the idiom chunk in (5) moves according to the pattern of adjuncts rather than according to the pattern of direct objects follows under Rizzi's (1990) theory if idiom chunks do not have a referential index. Notice however that all of the idiom chunks considered by Rizzi (1990) are of the form verb-bare nominal. Suppose then that in 'make headway' we take the DP projection to be missing altogether in 'headway'. Suppose further-

more that while a DP can be an argument, in the sense of Chomsky (1981), a bare NP never is. If NP is not an argument, it is not theta-marked under Chomsky's (1981; 1986b) Theta-Criterion; hence under Chomsky's (1981; 1986b) Visibility it is not Case-marked. This means that an idiom chunk is not K- or theta-governed and hence its sensitivity to *wh*-islands follows, as in (5).

Similarly, consider selected adverbs, as in (4). The suggestion of Rizzi (1990) is that selected adverbs must satisfy antecedent-government because they are not referentially theta-marked and hence do not have referential indices. However adverbs arguably are not arguments and hence do not satisfy Visibility/ the Theta-Criterion. If so, adverbs are never K- or theta-governed, and the impossibility for them to cross *wh*-islands is straightforwardly predicted under the present theory as well. Indeed though Rizzi (1990) uses selected adverbs and idiom chunks to argue against the notion of theta-government of Chomsky (1986a), the argument does not really go through.

Having now reviewed the empirical evidence relating to Rizzi's (1990) and Cinque's (1991) notion of referential indices, the general conceptual issue of the interpretations corresponding to the postulated LF's in (24) and (25) remains to be taken into consideration. Consider the semantic reconstructions of Rizzi's (1990) and Cinque's (1991) theory provided by Frampton (1990b) or Szabolcsi and Zwarts (1990). According to such reconstructions of the theory, the class of dependencies that can be successfully formed across a *wh*-island is characterized by the presence of an operator-variable structure with the variable ranging over individuals, or in other words of an individual variable, in the sense of Heim (1987). In Frampton (1990b), two representations of the type in (24–25) are in fact motivated on semantic grounds for sentences involving extraction of *how many*, where (24) represents the individual variable reading, (25) the other possible reading. Once given (24–25), it is obvious that the present discussion goes through as before.

In short, given the availability of both DP-raising and D-raising in the syntax, we can assume that the same possibility is open at LF. The resulting LF's account for the systematic ambiguity between individual variable readings, corresponding to DP variables, and non-individual readings, corresponding to D variables. Crucially a DP can be K- or theta-governed, while a D variable cannot. This accounts for the fact that individual readings remain available across *wh*-islands, but non-individual readings disappear.

2. Rigid and Relativized Minimality

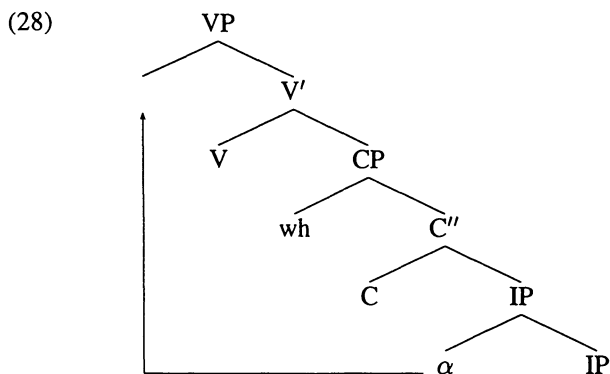
So far I have been assuming that some antecedent-government condition holding of non K- or theta-governed elements explains their locality behavior; but I have not provided any definition for this notion. In the literature, essentially two definitions are available. One, (rigid) Minimality, is provided by Chomsky

(1986a). The second corresponds to another major theoretical proposal put forward by Rizzi (1990), namely Relativized Minimality, and its core is accepted by Chomsky (1995). In what follows I will consider them in turn.

According to Chomsky (1986a) antecedent-government is defined, as in (27), in terms of the notion of barrier in (26). To be precise, the definition of barrier in (26), or minimality barrier, is only one of two definitions of barrier that enter into antecedent-government; (26) however is sufficient for the purposes of the present discussion. The definition in (26) itself is slightly recast with respect to Chomsky's (1986a); in particular no exceptional status is accorded to IP, and the phrasal projection XP, rather than X', is conceived of as creating a barrier. But the central idea behind (rigid) Minimality remains unchanged, namely that a head defines a barrier for its complements, though not for its Spec(s) nor for itself:

- (26) β is a barrier for α iff
 β is an XP, β dominates α and α is not the head or Spec of β
- (27) β governs α iff
 there is no barrier for α that excludes β

Objects, which are typically K-/ theta-governed, do not need to satisfy antecedent-government, hence do not give rise to wh-island violations. Adjuncts, on the other hand, that do not satisfy theta- or K-government, and must therefore satisfy antecedent-government, do give rise to wh-island violations. The relevant structure for the violation in (1) is provided in (28). Since I am assuming, with Frampton (1990a), Lasnik and Saito (1992) among others, that IP does not have exceptional status, the adjunct is shown in the IP-adjoined position α . Of course IP in (28) is not a barrier for α , since it does not dominate it in Chomsky's (1986a) sense. Hence if α moves to the Spec of CP position, no barrier is crossed. If the Spec of CP position is filled, however, CP must be crossed. Now, CP is a barrier for α in (28), since it dominates α , and α is not in its head or Spec. Thus the crossing of the CP barrier violates antecedent-government, correctly predicting the ill-formedness of (1):



Consider now Rizzi's (1990) theory. This is based on the idea that antecedent-government is crucially constrained by Relativized Minimality, as in (29); informally, antecedent-government is blocked whenever a potential antecedent-governor is crossed. The notion of typical potential antecedent-governor, which enters into (29), is defined as in (30); informally, an A'-Spec is a potential antecedent governor for an A'-dependency, an A-Spec is a potential antecedent-governor for an A-dependency, and a head is a potential antecedent-governor for a head-dependency:

- (29) X antecedent-governs Y only if there is no Z such that
- (i) Z is a typical potential antecedent-governor for Y, and
 - (ii) Z c-commands Y and does not c-command X
- (30) Z is a typical potential antecedent-governor for Y,
- $$Y \text{ in an } \left\{ \begin{array}{l} \text{A-chain iff Z is an A Spec} \\ \text{A'-chain iff Z is an A' Spec} \\ \text{head-chain iff Z is a head} \end{array} \right\} \text{ c-commanding Y}$$

Let us consider how Relativized Minimality works. Consider first wh-island violations, as in (1-2). In the terms of Rizzi (1990), the object wh-phrase in (2) is assigned a referential theta-role, which licences a referential index. Hence it need not satisfy antecedent-government. In (1), on the other hand, the adjunct wh-phrase is not assigned a referential index because it is not assigned a referential theta-role; hence it must satisfy antecedent-government. Now suppose the adjunct in (1) has reached the IP-adjoined position α in (28). In the presence of a wh-phrase in the Spec of CP, its next possible landing site is adjoined to the matrix VP. However, movement from the IP-internal to the VP-adjoined position is blocked by Relativized Minimality, precisely because a typical potential antecedent-governor intervenes between the two positions in the form of the wh-phrase in the Spec of CP. Thus the ungrammaticality of (1) is correctly predicted.

Crucially, though the theories of antecedent-government in Chomsky (1986a) and Rizzi (1990) are empirically equivalent with respect to (1-2), two further types of islands are introduced in support of Relativized Minimality by Rizzi (1990), namely inner islands in the sense of Ross (1984), and pseudo-opacity in the sense of Obenauer (1976; 1984). They belong to the same family as wh-islands in that adjuncts are sensitive to them, but objects are not; hence the designation of weak islands introduced by Cinque (1991). Under the simplest hypothesis the component of the theory involved in predicting them must of course be the same.

Consider first pseudo-opacity. The relevant data are of the type in (31-34). Both of the extractions in (31) and (32) are well-formed in French, where in (32) the quantifier *combien* moves stranding the rest of its DP, while in (31) the whole DP moves. Consider however the contrast in (33-34). (34), where only *combien* moves across the quantificational adverb *beaucoup*, is ill-formed; (33), where the whole DP moves across the adverb, is well-formed:

- (31) *Combien de livres a-t-il consulté?*
How-much of books has-he consulted
'How many books has he consulted?'
- (32) *Combien a-t-il consulté de livres?*
How-much has-he consulted of books
'How many books has he consulted?'
- (33) *Combien de livres a-t-il beaucoup consulté?*
How-much of books has-he much consulted
'How many books has he much consulted?'
- (34) **Combien a-t-il beaucoup consulté de livres?*
How-much has-he much consulted of books
'How many books has he much consulted?'

In present terms, movement of the D head *combien* out of its DP is predicted to give rise to weak islands effects, because it is not K- or theta-governed. Thus we predict movement of *combien* across a wh-island in (36) to be ill-formed, and movement of the whole DP to be well-formed, as in (35). The question is whether the contrast between (33) and (34) can similarly be accounted for:

- (35) *Combien de problèmes sais-tu comment résoudre?*
How-much of problems know-you how to-solve
'How many problems do you know how to solve?'
- (36) **Combien sais-tu comment résoudre de problèmes?*
How-much know-you how to-solve of problems
'How many problems do you know how to solve?'

Under Rizzi's (1990) Relativized Minimality an island is created in (34) by the quantificational adverb *beaucoup*, on the assumption that it is in the Spec of VP. If so, *beaucoup* counts as a potential A'-antecedent for the wh-trace, exactly as the wh-phrase in the Spec of CP in (36); hence the ill-formedness of (34) and (36) follows, since movement takes place across a potential antecedent-governor. What is more, Relativized Minimality correctly predicts the contrast between (37) and (38), since in (38) an adjunct, *comment*, is extracted across the floated quantifier *beaucoup*, a potential antecedent-governor, while in (37) no potential antecedent-governor is crossed:

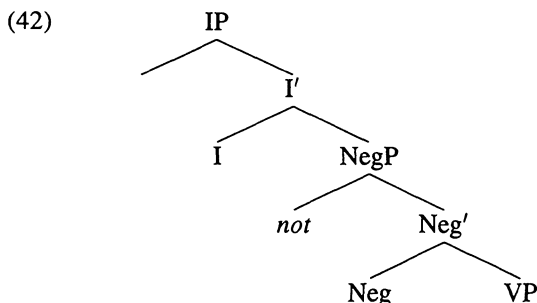
- (37) *Comment a-t-il résolu beaucoup de problèmes?*
How has-he solved much of problems
'How did he solve many problems?'
- (38) **Comment a-t-il beaucoup résolu de problèmes?*
How has-he much solved of problems
'How did he solve many problems?'

By contrast, Chomsky's (1986a) Minimality does not yield any prediction with respect to either (34) or (38). In both cases the lower VP is a barrier for the non-theta-governed/ non K-governed *wh*-phrase *combien* or *comment*, assuming the latter is generated VP-internally. But in both cases the VP barrier can be circumvented by simply adjoining to it. Of course adjunction to VP cannot circumvent Relativized Minimality, since the Spec of VP still intervenes between any VP-internal position and the VP-adjoined one.

Consider then inner islands. (40) differs from (39) only in the presence vs. absence of a sentential negation. (40) is ill-formed in the interpretation under which the adjunct is extracted from the embedded sentence, as indicated, though it is well-formed in the interpretation, irrelevant here, under which the adjunct is associated with the matrix verb. (39) is well-formed under both interpretations. Inner islands are precisely the islands created by the negation in sentences of the type of (40). Crucially, no inner island effects arise with objects, as in (41):

- (39) *Why do you think John left?*
 (40) **Why don't you think [Peter left t]*
 (41) *What don't you think Peter did?*

Rizzi (1990) assumes that the negation element is generated in an A'-Spec. This can be taken to be the Spec of VP; alternatively, if a sentential negation defines a NegP projection of its own, the negation can be taken to fill the Spec position of NegP. If so, the resulting phrase structure configurations are of the type in (42), under Pollock's (1989) assumption that NegP is generated above VP and under IP:



Given structures of the type in (42), the extraction of the adjunct in (40) must at some point cross *not*, a potential antecedent-governor for the *wh*-phrase in virtue of its A'-Spec position. Hence Rizzi's (1990) Relativized Minimality does not hold, and ill-formedness is correctly predicted to arise. By contrast, as before, no island is defined in (40) under Chomsky's (1986a) Minimality; like the VP barrier, the NegP barrier can be escaped by adjoining to it. Hence no predictions as to the ungrammaticality of (40) follow under Chomsky's (1986a) theory.

In summary, Minimality and Relativized Minimality are equally successful when applied to standard *wh*-island violations; but Relativized Minimality goes on to predict two other types of island which remain unexplained under Minimality. Let us then consider again *wh*-islands, the one type of Relativized Minimality island that Minimality derives so far. Data of the type in (1) can be accounted for under Minimality, by having recourse to the fact that the adjunct *wh*-phrase can move neither through the Spec of CP, which is already filled, nor through a CP-adjoined position, since following Chomsky (1986a) adjunction to CP is impossible. Obviously, the analogous solution for inner and pseudo-opacity islands would be to say that the adjunct *wh*-phrase not only cannot move through the Spec of NegP or the Spec of VP for the obvious reason that it is already filled, but that it cannot move through a NegP-adjoined or VP-adjoined position either.

Suppose then we take the step of prohibiting adjunction to maximal projections in general. Consider first the consequences of this hypothesis for VP. If movement cannot proceed through adjunction to VP, the Spec of VP remains the only VP-internal escape hatch. This predicts that if the Spec of VP is filled, so that a *wh*-phrase cannot move through it, a violation of antecedent-government is produced, correctly accounting for (38). However a potential problem also arises. For, while an infinite number of adjunction sites are available at any given VP, there is just one Spec of VP position. Thus no two different *wh*-phrases can be extracted through the same Spec of VP, though infinite ones can be extracted through VP-adjoined positions.

Now, let us assume with Rizzi (1990) that in the absence of antecedent-government requirements, argument *wh*-phrases can move in a single step to their final landing site. If so, they never need to use the Spec of CP, thus accounting for the well-formedness of (2), and they also never need use the Spec of VP, thus accounting for the well-formedness of (33). Consider on the other hand adjunct *wh*-phrases. Any two such phrases adjoining to the same VP, must pass through the same Spec of CP; and movement of more than one of them through the same Spec of CP is blocked in any theory. Thus prohibiting adjunction to VP, and allowing only the Spec of VP to function as an escape hatch, appears to have no undesirable empirical consequences.

Much the same can be repeated for NegP. The non-availability of multiple adjunction to NegP is irrelevant for argument *wh*-phrases, on the assumption that in the absence of minimality requirements these can move in one step to their final landing site. As for adjunct *wh*-phrases, the movement of two or more such phrases at the same time is always blocked by the existence of a single escape hatch from CP; hence it appears to be irrelevant again that NegP allows only one such phrase to move through, as under the no-adjunction hypothesis, rather than several of them.

On the other hand, prohibiting adjunction to every maximal projection means prohibiting adjunction to IP as well. Since in this case the Spec position is systematically filled by a subject, the impossibility of adjunction means that movement

must take place directly from the Spec of VP position to the Spec of CP position, crossing IP. Thus if IP is a barrier, as under the definition of Minimality that we have adopted so far, movement across IP becomes impossible in all cases; or else some stipulation concerning IP must be reintroduced into the theory, essentially the reverse of Chomsky's (1986a) stipulation, namely that adjunction is only possible to IP. Obviously the first result is unacceptable, while the proposed alternative is at least undesirable.

It appears then that the strong hypothesis that adjunction to maximal projections is in general impossible cannot be maintained. Consider however a weaker hypothesis, namely that adjunction is possible only to those maximal projections which do not have an A'-Spec. This prevents adjunction to NegP and VP, while allowing adjunction to IP, as desired. This hypothesis, however, raises a conceptual problem of its own. If it is correct that adjunction and A'-Specs are in complementary distribution, then the simplest grammar is one in which they are identified. Since CP must have a (unique) A'-Spec, rather than (multiple) adjunctions, and the same can be true of VP and NegP, this apparently means that the simplest grammar is one under which IP also has a (unique) A'-Spec.

Suppose, in fact, we say that for each maximal projection there is at most one A'-position, as in (43), disregarding the question whether this is to be conceived as a Spec position or as an adjoined one:

- (43) Each maximal projection is associated with at most one A'-position

The question as to whether IP obeys (43) can largely be answered as for VP and NegP. Since the presence of a single Spec of CP position limits the number of non-argument wh-extractions to one per sentence in any case, the availability of one landing site at any other maximal projection does not have any empirical consequences. Thus as far as wh-movement is concerned, the hypothesis in (43) that there is at most one A'-position per maximal projection, can be maintained. Under (43) pseudo-opacity and inner islands are now straightforwardly derived on exactly the same grounds as wh-islands. Thus a quantificational adverb or negation in the A'-Spec of VP or NegP prevents a wh-phrase from passing through the same position. This forces the crossing of a barrier and hence a minimality violation.

Since, on the other hand, we have limited our discussion so far to wh-movement, the question arises whether multiple adjunctions, as opposed to a single A'-Spec, must be admitted for other types of movement. An obvious candidate is multiple adjunction to IP of quantified expressions, as in LF's of the type in (44), taken from May (1985). May (1985), in fact, basing himself on Guéron (1981), explicitly suggests that a syntactic principle of the type in (43) holds. According to May (1985), multiple adjunction to IP, as in (44), is unnecessary. Instead (44) is argued to be equivalent to (45), where one quantified DP is adjoined to IP, and the second one is adjoined to the first:

- (44) [_{IP} *Every spy*_i [_{IP} *some Russian*_j [_{IP} *t_i suspects t_j*]]]

- (45) [IP [DP *Every spy*_i [DP *some Russian*_j]] [IP *t_i suspects t_j*]]

What is crucial is that in May's (1985) terms, in both (44) and (45) the two operators have the same absolute scope, since they have the same c-command domains, or equivalently they commute in relative scope, since they c-command each other. Technically, May (1985) defines a sigma-sequence as in (46), where *O* stands for an operator. Under (46), it is obvious that the two quantified DP's form a sigma-sequence in both (44) and (45). May's (1985) Scope Principle then simply says that members of sigma-sequences are free to take on any type of relative scope relation, as in (47); under (47), (44) and (45) give rise to exactly the same range of interpretations, as desired:

- (46) *S* is a sigma-sequence iff
for all *O_i* and *O_j* in *S*, *O_i* c-commands *O_j* and *O_j* c-commands *O_i*
- (47) *Scope Principle*
If *O_i* and *O_j* are members of a sigma-sequence,
O_i and *O_j* are free in scope with respect to one another

Exactly as two quantifiers adjoined to one another can fill the unique *A'*-position associated with IP in (45), we expect that two *wh*-phrases that commute in scope can be adjoined to one another and can therefore cooccur in the same Spec of CP. Since clusters of *wh*-phrases are in fact overtly excluded in English, as in (48), this prediction appears to be incorrect. However, in English multiple *wh*-phrases in the Spec of CP must be admitted at LF in cases of *wh*-in-situ at S-structure. In other words, something like (48) must in fact be the LF corresponding to the well-formed (49). Thus the prediction of the theory is once more seen to be correct:

- (48) **What who saw?*
(49) *Who saw what?*

Furthermore, we can expect that given that structures of the type in (48) must be available at LF, and *wh*-movement is in general available in the syntax, the impossibility of (48) occurring in the syntax in English corresponds to a parametric choice, rather than to a universal property of natural languages. Again this appears to be confirmed by the facts, since examples of the type in (48) are found in other languages, for instance in Slavic languages, according to Rudin (1988).

In short, Relativized Minimality is equivalent to (rigid) Minimality, with respect to pseudo-opacity and inner islands, once the restriction in (43) is imposed on the grammar. However what precedes only takes into account arguments for Relativized Minimality from phrasal movement, as presented by Rizzi (1990), while a number of arguments has recently surfaced in the literature in favor of Relativized Minimality for head-movement. In particular, Baker and Hale (1990) propose that a lexical head can move only to the nearest lexical head, and a functional head to the nearest functional head. This state of affairs can be captured by Relativized

Minimality if the notion of potential antecedent governor is revised so that a potential antecedent-governor for a lexical head is another lexical head, while a potential antecedent-governor for a functional head is another functional head. The fact that Baker and Hale (1990) seek to explain is that D-incorporation, i.e. cliticization, can take place from an indirect object or subject position, but not N-incorporation. Here we will only notice that the introduction of the lexical/functional distinction in addition to the head/phrase and A/A' distinctions has an undesirable conceptual drawback. For, on the one hand, not all properties of heads can be relevant to Relativized Minimality; but on the other hand, there appears to be no principled reason why the lexical/functional property should be, and not others.

A different extension of Relativized Minimality for head-movement, which does not necessarily incur in the same problem, is suggested by Roberts (1994). The latter's idea is that the A/A' distinction is relevant not only for phrasal movement but also for head-movement. Thus an A-head (Agr, lexical heads) is a potential antecedent-governor for another A-head, an A'-head (C, Neg) for another A'-head. The crucial empirical evidence comes from Long Head Movement patterns, of the type presented by Lema and Rivero (1990). Notice that the problem in this case is providing a general characterization for what constitutes an A- or an A'-head. But quite clearly Relativized Minimality violations involving heads cannot be made to fall under rigid Minimality by simply imposing (43). Thus head-movement offers a potentially decisive ground for the testing of minimality notions.

*Dipartimento di Linguistica
Università di Firenze
Piazza Brunelleschi
50121 Firenze (Italy)
rmanzini@cesit1.unifi.it*

*Department of Linguistics
University College London
Gower Street
London WC1E 6BT
rita@ling.ucl.ac.uk*

References

- Baker, Mark and Kenneth Hale (1990). Relativized Minimality and pronoun incorporation. *Linguistic Inquiry* 21: 289–297.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- (1986a). *Barriers*. Cambridge, Mass: MIT Press.
- (1986b). *Knowledge of Language*. New York: Praeger.
- (1995). *The Minimalist Program*. Cambridge, Mass: MIT Press.
- Cinque, Guglielmo (1991). *Types of A'-Dependencies*. Cambridge, Mass: MIT Press.
- Frampton, John (1990a). Parasitic gaps and the theory of wh-chains. *Linguistic Inquiry* 21: 49–77.
- (1990b). The fine structure of wh-movement and the proper formulation of the ECP. Paper delivered at the 13th GLOW Colloquium (Cambridge).
- Guéron, Jacqueline (1981). Logical operators, complete constituents, and extraction transformations. In *Levels of Syntactic Representation*, Robert May and Jan Koster (eds.). Dordrecht: Foris.
- Heim, Irene (1987). Where does the definiteness restriction apply? In *The Representation of (In)definiteness*, Eric Reuland and Alice ter Meulen (eds.). Cambridge, Mass: MIT Press.

- Huang, James (1982). Logical relations in Chinese and the theory of grammar. Ph.D. Dissertation. MIT, Cambridge, Mass.
- Lasnik, Howard and Mamoru Saito (1992). *Move- α* . Cambridge, Mass: MIT Press.
- Lema, José and Maria-Luisa Rivero (1990). Long head movement: ECP vs. HMC. In *Proceedings of NELS 20*. Amherst: GSLA.
- Longobardi, Giuseppe (1991). Extraction from NP and the proper notion of head government. In *The Syntax of Noun Phrases*, Alessandra Giorgi and Giuseppe Longobardi (eds.), 57–112. Cambridge: Cambridge University Press.
- Manzini, Maria Rita (1992). *Locality*. Cambridge, Mass: MIT Press.
- (1997). A minimalist theory of weak islands. In *The Limits of Syntax* (Syntax and Semantics, vol. 29), Peter Culicover and Louise McNally (eds.), 185–209. New York: Academic Press.
- May, Robert (1985). *Logical Form*. Cambridge, Mass: MIT Press.
- Obenauer, Hans (1976). *Etudes de syntaxe interrogative du Français*. Tübingen: Niemeyer.
- (1984). On the identification of empty categories. *The Linguistic Review* 4: 153–202.
- Pesetsky, David (1987). Wh-in-situ: Movement and unselective binding. In *The Representation of (In)definiteness*, Eric Reuland and Alice ter Meulen (eds.), 98–129. Cambridge, Mass: MIT Press.
- Pollock, Jean-Yves (1989). Verb-movement, universal grammar and the structure of IP. *Linguistic Inquiry* 20: 365–424.
- Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge, Mass: MIT Press.
- Rudin, Catherine (1988). On multiple questions and multiple wh fronting. *Natural Language and Linguistic Theory* 6: 445–501.
- Roberts, Ian (1994). *Verbs and Diachronic Syntax*. Dordrecht: Kluwer
- Ross, John Robert (1984). Inner islands. *Proceedings of BLS 10*. University of California, Berkeley: Berkeley Linguistic Society.
- Szabolcsi, Anna and Frans Zwarts (1990). Semantic properties of composed functions and the distribution of wh-phrases. In *Proceedings of the 7th Amsterdam Colloquium*. Amsterdam: ILLC.

