

4 Relativized Minimality Effects

LUIGI RIZZI

0 Introduction

Natural language expressions are potentially unbounded in length and depth of embedding, as a consequence of the recursive nature of natural language syntax. Nevertheless, the core of syntactic processes is inherently local, in that such processes are bound to apply within limited structural domains. The search for the relevant locality principles is one of the central topics of generative grammar. In this chapter I would like to report on a subclass of locality effect which has been the focus of intensive research since the late 1980s or so. Such effects are all amenable to the following abstract form: in the configuration:

(1) ... X ... Z ... Y ...

Y cannot be related to X if Z intervenes and Z has certain characteristics in common with X. So, in order to be related to X, Y must be in a minimal configuration with X, where Minimality is relativized to the nature of the structural relation to be established. The major cases illustrating such Relativized Minimality (RM) effects involve the theory of chains: in the general case, a chain cannot be built between X and Y in configuration (1). In the first part of this chapter, I will present Relativized Minimality effects on the different types of chain through a representational formulation of the relevant locality principle, as in Rizzi (1990). I will then focus on A'-chains and refine the approach to deal with complex patterns of Minimality effects induced by different kinds of adverbial modifier. This part will also contain a brief comparison with the derivational approach proposed by Chomsky (1995b). I will conclude with a discussion of Minimality effects in head-phrase interactions.

1 Relativized Minimality

It is a central property of natural languages that elements can be pronounced in positions different from those in which they are interpreted as thematic

arguments, as theta-assigners, as modifiers of various sorts, etc. By “chain” I mean the connection between a displaced element and its traces, down to the relevant interpretive position. I will not take a position here on the question of whether such a connection is established by movement (an operation distinct from the fundamental structure-building operation, which perhaps simply consists in stringing together two elements into a third element, “Merge” in Chomsky’s 1995b sense) or is read off configurations (Rizzi 1986) created by the fundamental structure-building operation uniquely. When I mention movement from now on, I simply use the term metaphorically, to refer to an abstract relation between two structural positions.

On the other hand, following Chomsky’s (1995b) approach, I will assume that traces are silent copies of antecedents, and I will express them as fully specified positions within angled brackets (as in Starke 1997; in other environments I will use more standard symbols for traces such as “t” or “___”). So, (2a) will have a representation like (2b):

- (2) a. How did you solve the problem?
 b. How did you solve the problem <how>?

Now, it is well known that certain structural environments block chain formation; for instance, a *wh*-chain starting from an adverb position fails across an indirect question:

- (3) a. I wonder who could solve the problem in this way.
 b. *How do you wonder who could solve this problem <how>?

It is natural to think of (3b) as illustrating pattern (1): a chain cannot connect *how* and its trace in (3b) because another *wh*-element intervenes in the lower Spec of C.

In order to express the effect precisely, I will adopt the following principle (a simplification and updating of RM in Rizzi 1990):

- (4) Y is in a Minimal Configuration (MC) with X iff
 there is no Z such that
 (i) Z is of the same structural type as X, and
 (ii) Z intervenes between X and Y

So, if we think of Y as the position from which the relevant relation is computed, and X as the target of the computation (i.e., in the case of chain formation, the trace Y seeks for an antecedent X), the two elements are in a minimal configuration when there is no intervening element having certain structural characteristics in common with the target. A proper typology of positions is critical for the system to work; i.e., we must achieve the result that none of the intervening positions blocks the chain in the well-formed (2b), while at least one position does in (3b). The typology must involve at least two irreducible distinctions:

- (5) (i) between heads and phrases and, in the latter class,
 (ii) between positions of arguments (A' -positions) and of non-arguments (A' -positions).

Intervention could be defined hierarchically, in terms of c-command (Z c-commands one but not the other), but it could also be relativized to the kind of relation looked at (particularly if we want to extend the system to deal with locality in ellipsis, on which see below).

We can now define a chain through the notion of minimal configuration (4), as follows:

- (6) (A_1, \dots, A_n) is a chain iff, for $1 \leq i < n$
 (i) $A_i = A_{i+1}$
 (ii) A_i c-commands A_{i+1}
 (iii) A_{i+1} is in a MC with A_i

So, each chain link involves identity (under the copy-theory of traces), c-command and Minimality.

Going back to (3b), chain formation between *how* and \langle how \rangle fails because the two are not in a minimal configuration due to the intervention of *who*, an element of the same structural kind as *how*, an A' -specifier.

Notice that locality as expressed by (4) does not involve c-command. It seems to be the case that prominence (expressed by c-command) and locality (expressed by (4)) are two fundamental and independent configurational notions that natural languages use. Chains require both prominence and locality to hold, but other complex relations dissociate them. For instance, the possibility of interpreting a pronoun as a variable bound by a quantified expression requires c-command, hence (8) is out, but it is totally insensitive to locality, as in (7):

- (7) a. *No candidate* can predict how many people will vote for *him*.
 b. *Every politician* is worried when the press starts attacking *him*.
 c. *Which politician* appointed the journalist who supported *him*?

- (8) *The fact that *no candidate* was elected shows that *he* was inadequate.

The converse case may also exist: certain kinds of ellipsis seem to involve locality but not c-command. E.g., the gapped verb in (9) can be interpreted as identical to the local verb *buy*, not to the non-local verb *sell* (see Koster 1978 for relevant discussion):

- (9) John sells books, Mary buys records and Bill V newspapers.

We have seen that chains involve both locality and c-command to hold. The fourth case may also exist: simple coreference requires neither:

- (10) The question of whether *John* met Mary worries the people who support *him*.

Going back to chains, they divide into three fundamental kinds, depending on the nature of the displaced element and on the nature of the position it occupies. The first crucial distinction is between X chains and XP chains; then within the latter class, we want to distinguish at least between Argument chains (involved in passive, raising, and in fact any movement to subject position under the VP internal subject hypothesis, as well as object movement to various IP internal positions) and operator chains (involved in questions, relatives, exclamatives, focus movement, etc.). By and large, argument chains involve IP internal positions, while operator chains involve the CP system of the clause, the left periphery, even though the distinction is not absolute (see below). Different descriptive locality conditions hold of the three kinds of chain, head chains being by far the most local, and operator chains the most liberal. The system sketched out here is an attempt to subsume such descriptive conditions under a unified principle. We will now consider the different kinds of chain separately.

2 Head Chains

In general, heads cannot move across other heads: heads can be displaced over significant distances in the tree provided that they move through all the intervening head positions; as soon as one position is skipped, ill-formedness results. This is the Head Movement Constraint of Travis (1984). It is illustrated, for instance, by the fact that only the highest functional verb can move to C in English questions:

- (11) a. They have left.
 b. Have they <have> left?
- (12) a. They could have left.
 b. *Have they could <have> left?
 c. Could they <could> have left?

or by the fact that lexical verbs, unable to reach I in Modern English (as in 13c), cannot be moved to C in interrogatives (as in 13d):

- (13) a. He has often seen Mary.
 b. He *I* often sees Mary.
 c. *He sees often <sees> Mary.
 d. *Sees he *I* often <sees> Mary?

In earlier phases of English, when movement of the lexical V to I was possible, the continuation of the movement to C was also allowed (Roberts 1993b).

Modern French allows (in fact, requires) movement of the lexical V to I (Pollock 1989), and the continuation of this movement to C in interrogatives:¹

- (14) a. Il a souvent vu Marie.
 “He has often seen Marie.”
 b. Il voit souvent <voit> Marie.
 “He sees often Marie.”
 c. Voit-il <voit> souvent <voit> Marie?
 “Sees he often Marie?”

Consider also the fact that in Italian gerundival (Rizzi 1982) and participial (Belletti 1981) clauses the non-finite verb moves to C; but the participle cannot move across the gerundival auxiliary:²

- (15) a. Essendo Mario <essendo> tornato a Milano, ...
 “Having Mario come back to Milan, ...”
 b. Tornato Mario <tornato> a Milano, ...
 “Come back Mario to Milan, ...”
 c. *Tornato Mario essendo <tornato> a Milano ...
 “Come back Mario having to Milan, ...”

The same constraint is also illustrated by the fact that non-finite functional verbs in English can optionally move to the left of *not* (Pollock 1989), but can never cross the infinitive marker *to*:

- (16) a. Not to be
 b. To not be
 c. To be not
 d. *Be to not
 e. *Be not to

Assuming that *to* can occupy at least two functional head positions, higher and lower than *not*, and the latter is the specifier of NegP, *be* can optionally move head to head across *not* (16c), but cannot cross *to*.

The impermeability of *to* to movement is only a special instance of the general fact that particle like elements in the inflectional system have the effect of blocking V-movement. For instance, in Creole languages, in which the inflectional field is typically realized by a rich system of particles designating tense, mood, and aspect, we never find V-movement to C in interrogatives and related constructions. For example, Haitian Creole does not involve movement of the verb to C in interrogatives, contrary to French, etc. Evidently, the particles do not enter into head movement to C, but their presence suffices to block head movement to C of some lower verbal element, as expected under RM. Of particular interest here is the case of mixed systems, pointed out by Guglielmo Cinque. Cinque (1999) observes that languages having both affixes

and particles to express properties of the inflectional field may give rise to configurations like:

(17) ... Aff ... Prt ... V ...

where a structurally higher property is expressed by a verbal affix and a lower property by a particle in the inflectional field. In such cases, the language never resolves the situation by having the V move to the affix by jumping across the particle; rather the language reverts to the insertion of an auxiliary in a position higher than the particle, a semantically empty verb capable of supporting the relevant affix. Evidently, a strong structural constraint is operative to block movement here, forcing the language to complicate the representation in order to achieve morphological well-formedness. A case in point discussed by Cinque is Welsh: tense and agreement suffixes are normally attached to a main verb, but if aspectual particles *wedi* (perfective) or *yn* (progressive) are present, the language reverts to an auxiliary verb (be) to carry tense and agreement suffixes:

- (18) a. Cana i yfory.
sing-Fut-1Sg I tomorrow
"I will sing tomorrow."
b. Bydda i wedi canu erbyn saith o'r gloch.
be-Fut-1Sg I Perf sing by seven o'clock
"I will have sung by seven o'clock."
c. Bydda i 'n canu yfory.
be-Fut-1Sg I Prog sing tomorrow
"I will be singing tomorrow."

All these cases straightforwardly follow from the assumption that chain links must satisfy RM, as formally expressed by our principles (4) and (6).

3 A-Chains

A subject raised to a higher subject position cannot skip an intervening subject position; the banned configuration has been called "Super Raising":

- (19) a. It seems that it is likely that John will win.
b. It seems that John is likely t to win.
c. John seems t to be likely t to win.
d. *John seems that it is likely t to win.

One could think that the impossibility of (19d) is related to the impossibility of extracting the subject from the tensed clause:

- (20) *John seems (that) t will win.

If A-movement is triggered by Case, one could think that the DP simply moves to the closest case position, so it cannot move further from the embedded subject position in (20) and cannot skip the embedded subject position in (19d).

But in fact, the problem is more general: a pure Case approach is not plausible for cases of languages allowing raising out of tensed clauses, such as the dialect of Turkish discussed in Moore (1998). These languages clearly dissociate (19d) and (20):

- (21) Biz san-a viski-yi iç-ti-k gibi görün-du-k.
 We-Nom you-Dat whiskey-Acc drink-Past-1Pl like appear-Past-1Pl
 "We appeared to you (we) have drunk the whiskey."
- (22) Çok viski iç-ti san-d1-n gibi görün-du-0
 Much whiskey drink-Past-1Sg believe-Past-2Sg like appear-Past-3Sg
 "It appears you believed I have drunk a lot of whiskey."
- (23) *Çok viski iç-ti san-d1-n gibi görün-du-m
 Much whiskey drink-Past-1Sg believe-Past-2Sg like appear-Past-1Sg
 "I appear you believe (I) have drunk a lot of whiskey."

As raising is possible in this variety from the subject position of a finite clause, where the raised DP could receive Case (as in 21a), the locality effect shown by (21c) seems to be independent from attraction to the closest Case position. What this pattern suggests is that intervening subject positions of the kind involved in the satisfaction of the Extended Projection Principle (EPP) block any kind of A-chains, regardless of case considerations, as Moore points out.

This is also shown by the fact that intervening subjects also block A-chains not aiming at another (cased) subject position, but to a different kind of A-position. Consider for instance past participle agreement, triggered by a displaced object in French (e.g. in a *wh*-construction):

- (24) Les voitures qu'il a t' conduites t
 "The cars that he has driven(FP)"

I will assume (following in essence Kayne 1989) that agreement is triggered by the passage of the displaced object through a specifier position of the relevant agreement head (t' in (24)). It has been observed (Kayne 1989a, Déprez 1989) that this agreement is local. For instance, in case of *wh*-extraction of the object, it cannot be triggered on the verb of the higher clause:

- (25) Les voitures qu'il a dit(*es) qu'il a conduit(es)
 "The cars that he has said(FP) that he has driven(FP)"

The structure involving the higher agreement could have a representation like the following:

(26) Les voitures qu'il a t''' dites (t'') qu'il a t' conduites t

(t, t') is a well-formed link of an A-chain, but (t', t''') is not: if t'' is present in the spec of some C-projection, the chain moves from an A-position to an A'-position back to an A-position, a case of improper movement. If t'' is not there, the link (t', t''') crosses a subject position, again a Super Raising violation. All these cases are excluded by RM for X, Z = A specifiers.³

4 A'-Chains: The Asymmetries

Huang (1982) observed that the effect shown in (3) is not fully homogeneous. While adverbial elements strongly resist *wh*-extraction from *wh*-islands, *wh*-arguments are at least marginally extractable, as is shown by near-minimal pairs like the following (only the relevant extraction trace is indicated):

- (27) a. ?Which problem do you wonder how to solve <which problem>?
 b. *How do you wonder which problem to solve <how>?

So, while adverbs fully manifest the expected RM effects for A'-chains, arguments somehow manage to escape, at least in part. Since the mid-1980s, much work has been devoted to the exact structural characterization of the asymmetries, and to the identification of the class of "interveners" determining RM effects in A'-chains. Let us consider these two issues in turn.

If much work in the 1980s treated the contrast in (27) as an asymmetry between arguments and adverbs (or adjuncts), later it became clear that a more accurate characterization should set aside arguments from everything else. An instance of this wider generalization is offered by cases in which part of an argument can be moved out of the DP, apparently in free alternation with movement of the whole DP, as with *combien* extraction in French:

- (28) a. Combien de problèmes sais-tu résoudre ____?
 "How many of problems can you solve?"
 b. Combien sais-tu résoudre [____ de problèmes]?
 "How many can you solve of problems?"

Now, extraction of the whole direct object out of a *wh*-island is marginally acceptable, but extraction of *combien* alone is barred:

- (29) a. ?Combien de problèmes sais-tu comment résoudre ____?
 "How many of problems do you know how to solve?"
 b. *Combien sais-tu comment résoudre [____ de problèmes]?
 "How many do you know how to solve of problems?"

So, here we have an asymmetry between an argument and a proper part of an argument, the latter resisting extraction. Moreover, if pied-piped arguments

are marginally extractable from *wh*-islands, pied-piped predicates are not (Baltin 1992):

- (30) a. How many people do you consider ____ intelligent?
 b. How intelligent do you consider John ____?
- (31) a. ??How many people do you wonder whether I consider ____ intelligent?
 b. *How intelligent do you wonder whether I consider John ____?

In conclusion, the asymmetry appears to be between arguments and everything else: predicates, adverbs and proper subparts of arguments resist extraction from *wh*-islands.

An additional restriction, more subtle but clearly detectable in many cases, was then brought to light. It was observed that *wh*-arguments are optimally extractable only with a special interpretation, i.e., when the range of the variable is pre-established in discourse, or presupposed (Comorovski 1989, Cinque 1990b). This discourse-linked (D-linked: Pesetsky 1987) or presupposed interpretation is favored or forced by certain lexical choices for the *wh*-operator (e.g. by *which* in English), while certain types of *wh*-phrase like *what the hell*, *what on earth* are incompatible with it (Pesetsky's "aggressively non-D-linked" *wh*-expressions), whence such contrasts as:

- (32) a. ?Which problem do you wonder how to solve ____?
 b. *What the hell do you wonder how to say ____?

The necessity of a pre-established range emerges with clarity also in the interpretation of other *wh*-operators. Consider the following contrast in Italian (and see Frampton 1991 for the discussion of similar pairs in English):

- (33) a. ?Quanti problemi non sai come risolvere ____?
 "How many problems don't you know how to solve?"
 b. *Quanti soldi non sai come guadagnare ____?
 "How much money don't you know how to make?"

While it is natural to assume that there may be a known set of problems of which (33a) can be asked, in general (33b) will not be asked about a known set of objects (say, sums of money pre-established in discourse). On the other hand, even sentences like (33b) can improve if the partitive form of the interrogative DP is used, explicitly stating that the question bears on a specific sum of money:

- (34) ?Quanti dei soldi che ti servono non sai come guadagnare ____?
 "How much of the money that you need don't you know how to make?"

In conclusion, *wh*-extraction from a *wh*-clause is generally barred, as expected under RM. A systematic exception involves D-linked argumental *wh*-phrases, which are marginally extractable. In order to accommodate the exception, it was proposed in Rizzi (1990) that such *wh*-phrases can be related to their traces through a mechanism different from ordinary chain formation as expressed by (6), and as such are not submitted to locality. As the theory must admit a way to relate positions non-locally in order to accommodate long distance binding of a pronoun by a quantified expression (see (7) and the related discussion), the proposal was made that D-linked argumental *wh*-phrases can exploit such a mechanism to be related long distance to their traces. This idea was implemented in Cinque (1990b) and Rizzi (1990) by assuming that only D-linked and argumental (theta-marked) *wh*-phrases could bear (and share with their traces) a referential index capable of ensuring the antecedent-trace connection non-locally. Much of the discussion on this approach was centered on the legitimacy and appropriateness of using referential indices as a technical device to permit long distance dependencies (see Frampton 1991 for a critical discussion, Chomsky 1995b on the possibility of dispensing with indices, and Manzini 1992 for a different view on the asymmetries).

Here I would like to suggest a different implementation, which does not resort to indices and fully exploits the analogy with (7).⁴ Let me first introduce a mechanism to deal with the latter case and, in general, with long distance binding of pronouns from quantified expressions. Following fairly standard assumptions, I will assume that binding holds when the following conditions are met:

- (35) A binds B iff:
- (i) A and B are non-distinct DPs, and
 - (ii) A c-commands B.

It is natural to restrict binding to DPs (possibly to DPs and CPs), the only categories that can enter into referential dependencies. Featural non-distinctness is needed to ensure feature matching between the binder and the bindee (here we cannot require the stronger condition of full structural identity as in (6i): a bound pronoun is not identical to its binder); and c-command must hold (cf. (8)).

Now, suppose that (35) is the grammatical device that can be used as an alternative to “being in a minimal configuration” in chain formation (condition (6iii)). Given the fact that (35) is restricted to DPs, the non-local mechanism is not available with *wh*-dependencies involving adjuncts (27b), predicates (31b), or parts of DPs (29b). What about non-D-linked *wh*-DPs? Following Frampton (1991) I will assume that in such cases “reconstruction” must apply, leaving the bare *wh*-operator in the left periphery, and reconstructing the rest of the DP in situ. If we adopt the core idea of the theory of reconstruction in Chomsky (1995b), exploiting the copy theory of traces, S-structure (36a) is thus converted into LF (36b) by deleting the non-operator material from the left periphery

(an operation which automatically triggers the consequence that the lexical restriction is “exported” from the trace in (36b), if the trace is understood as an identical copy of the antecedent):

- (36) a. Quanti soldi non sai come guadagnare <quanti soldi>
 b. Quanti ____ non sai come guadagnare <quanti> soldi

Therefore, at LF the *wh*-dependency in the non-D-linked interpretation is not a DP dependency, but a bare QP dependency, so that mechanism (35) is not available, locality must be met in the strict form of (6iii), and (36b) is correctly ruled out. Following again Frampton (1991) and Chomsky (1995b), we can account for the obligatoriness of reconstruction here by assuming that the principle of Full Interpretation is strong enough to enforce it: the lexical restriction is not licensed by a left-peripheral mechanism in (36), so it must delete in the syntax of LF.

Now, what about D-linked *wh*-phrases, e.g. in (33a)? Following the logic of the argument, the lexical restriction in D-linked *wh*-elements must be allowed to stay in the left periphery at LF, in order to permit a non-local DP dependency to be established through (35). Why is this legitimate? I continue to assume that D-linked means “pre-established in discourse,” hence topic like; of course, topic interpretation licenses elements in the left periphery (e.g. through the formal device discussed in Rizzi 1997), so we can think that the lexical restriction can remain in the left periphery as it is licensed *qua* topic at LF.

In conclusion, the selective sensitivity to *wh*-islands can be reduced to the formal distinction between DP and non-DP dependencies at LF. Among the *wh*-dependencies, only D-linked phrases allow DP dependencies at LF, which can exploit the non-local connecting device (35), and survive across a *wh*-island. Adjunct, predicative, and bare *wh*- (extracted from DP) dependencies are not DP dependencies, so they must obey RM; non-D-linked *wh*-dependencies like (36) are DP dependencies at S-structure, but not at LF, the level where RM applies, so that they are expected to obey strict locality under this approach.^{5, 6}

5 A'-Chains: The Interveners

It was observed in the early 1990s that the class of possible interveners triggering minimality effects is not coextensive with the class of target positions, but significantly wider. For instance, asymmetries of the kind illustrated by (29) are also determined by an intervening negation:

- (37) a. Combien de problèmes ne sais-tu pas résoudre ____?
 “How many of problems can’t you solve?”
 b. *Combien ne sais-tu pas résoudre [____ de problèmes]?
 “How many can’t you solve of problems?”

This is a particular case of the selective islands induced by negation, which non-arguments are generally sensitive to (Ross 1983 and, for recent critical discussion, Szabolcsi and Zwarts 1997, Kuno and Takami 1997, and the references cited there).

Certain kinds of quantificational adverb expressing the frequency of an action, or the intensity of a state, determine similar asymmetries in French (Obenauer 1983, 1994):

- (38) a. Combien de livres a-t-il beaucoup consultés ____?
 "How many of books has he a lot consulted?"
 b. *Combien a-t-il beaucoup consulté ____ de livres?
 "How many has he a lot consulted of books?"
 c. Combien de films a-t-elle peu aimés ____
 "How many films did she little like?"
 d. *Combien a-t-elle peu aimé [____ de films]?
 "How many did she little like of films?"

Given the width of the blocking effects, it seemed natural to define the structural typology relevant for RM in a very broad way. As *wh*, negation, and quantificational adverbs like *beaucoup*, *peu*, *souvent*, etc. plausibly have in common the fact of occupying an A'-specifier position, the formulation of the principle in Rizzi (1990) referred to the sole distinction between A- and A'-specifiers to express the typology of elements triggering minimality effects at the XP level.

On the other hand, the plausibility of such a purely geometric approach was immediately threatened by the observation that non-quantificational adverbs do not induce a similar Minimality effect (Obenauer 1983, 1994, Laenzlinger 1996): compare (36) with the following:

- (39) a. Combien de livres a-t-il attentivement consultés ____?
 "How many of books did he carefully consult?"
 b. Combien a-t-il attentivement consulté [____ de livres]?
 "How many did he carefully consult of books?"

Now, it appears unlikely that the position of a non-quantificational adverb like *attentivement* differs enough in tree geometry from the position of *beaucoup*, *peu*, etc. to allow us to maintain a purely geometric approach. This is even more so in view of recent advances on the study of adverbial positions *qua* specifiers of particular functional heads (Cinque 1999, Laenzlinger 1996) and on the opportunity of eliminating the distinction between specifier and adjoined position (Kayne 1994). If these approaches are on the right track, then the theory of phrase structure has no resources to differentiate the position of these adverbs uniquely in terms of configurations.

Clearly, a more selective characterization of the Minimality inducing factor is needed. An extremely selective characterization is in fact provided in the reinterpretation of RM effects in Chomsky (1995b). According to Chomsky,

movement is triggered by feature attraction: a head endowed with a certain feature attracts a phrase bearing the same specification to its immediate structural domain (say, in traditional X-bar terminology, to occupy its specifier position). This attract operation is phrased in such a way that only the closest “attractee” can move:

(40) **Minimal Link Condition:**

K attracts a only if there is no b, b closer to K than a, such that K attracts b. (Chomsky 1995b: 311)

(40) differs from (4) in two major respects. First, it is a principle that operates on derivations, not a well-formedness condition on representations, so it requires one to take the “movement metaphor” literally, whereas (4) does not. Moreover, it makes the Minimality effect sensitive to identity of features; i.e., in the derivation of (41) the higher C endowed with the *wh*-feature (element K in (40)) fails to attract the *wh*-phrase *how* from the embedded clause due to the intervention of a closer attractee, the *wh*-phrase in the embedded Spec of C:

(41) C you wonder [which problem C [to solve how]]

Here I will not address the important distinction between derivational and representational approaches to locality (but see n. 6), and will focus on the different selectivity of the two approaches. It is clear that (40) suffers from the opposite problem with respect to (4): it is too selective to account for cases of Minimality such as (37b) and (38b–d), in which the intervener bears a feature different (formally and interpretively) from the one of the attractor. If we want to maintain that a genuine generalization underlies (29b) and these other examples we cannot adopt this approach to RM effects.⁷

What seems to be needed is an intermediate position between the geometric approach and the one based on identity of features. Features determining chain formation seem to cluster into natural classes, such that Minimality effects are determined within classes, but not across them.

6 Classes of Features

In accordance with much current work, I will assume that specifiers must be licensed by the sharing of certain features with the respective heads. One can think of this licensing as a special need that certain features have (they must be “checked” in a local environment, as in Chomsky’s 1995b system), or a subcase of head-XP selection, mutual selection with respect to certain formal features. In this family of approaches, it is natural to assume that the typology of specifier positions is determined by the licensing features. So, the typology of positions reduces to the clustering of features into natural classes.

What question and negative elements have in common with adverbs like *beaucoup*, *peu*, *souvent* is their quantificational character (Rizzi 1990, Laenzlinger 1996). As in the core case A'-positions are scope positions of quantificational elements, one could think of restricting minimality effects in the A'-system to such scope positions, thus correctly excluding cases like (39b) from the domain of the principle. On the other hand, one cannot simply exclude non-quantificational adverbs from RM effects without further qualifications: intervening adverbs of all kinds generally determine a Minimality effect on certain kinds of adverb movement, a fact often pointed out in the relevant literature (now reviewed in detail in Cinque 1999, a study on which the following discussion is based). For instance, Koster (1978) observed that, given the hierarchy of adverbs in IP internal position in Dutch, shown by the relative order in the embedded clause (42a) (vs. the ill-formedness of (42b)), the lower epistemic adverb "probably" cannot be moved in a V-2 construction across the higher evaluative adverb "unfortunately," as in (43b); if no other adverb intervenes, the epistemic adverb can be moved to the first position in V-2, as in (43c):⁸

- (42) a. Het is zo dat hij helaas waarschijnlijk ziek is.
 "It is so that he unfortunately probably sick is."
 b. *Het is zo dat hij waarschijnlijk helaas ziek is.
 "It is so that he probably unfortunately sick is."
- (43) a. Helaas is hij ____ waarschijnlijk ziek.
 "Unfortunately is he probably sick."
 b. *Waarschijnlijk is hij helaas ____ ziek.
 "Probably is he unfortunately sick."
 c. Waarschijnlijk is hij ____ ziek.
 "Probably is he sick."

This preservation of the ordering under movement is clearly reminiscent of RM (and in fact it is explained by a very similar principle in Koster's account, as Cinque points out).

We observe the same effect with adverb preposing in Italian, except that, of course, preposing does not trigger V-2 in this language: a lower adverb like *rapidamente* can be fronted to pre-IP position, as in (44a), but not across a higher epistemic adverb, as in (44c):

- (44) a. I tecnici hanno (probabilmente) risolto rapidamente il problema.
 "The technicians have probably resolved rapidly the problem."
 b. Rapidamente, i tecnici hanno risolto ____ il problema.
 "Rapidly, the technicians have resolved the problem."
 c. *Rapidamente, i tecnici hanno probabilmente risolto ____ il problema.
 "Rapidly, the technicians have probably resolved the problem."

On the other hand, it is not the case that movement of an adverb across another adverb is barred in general. The Minimality effect is observed when the landing site is another adverbial position. If the landing site is different in nature, the Minimality effect may not arise. This is clearly shown by a contrast observed in French by Schlyter (1974) and mentioned in Cinque (1999): simple fronting of a lower adverb across *probablement* is banned, as in the Italian example (see (45b)), but focalization of the lower adverb in a cleft construction across the higher adverb is fine (see (45c)):

- (45) a. Il a probablement travaillé énergiquement.
 “He has probably worked energetically.”
 b. *Énergiquement, il a probablement travaillé.
 “Energetically, he has probably worked.”
 c. C’est énergiquement qu’il a probablement travaillé.
 “It is energetically that he has probably worked.”

The same effect is found in Italian: (44c) clearly improves if the adverb is focalized:

- (46) RAPIDAMENTE i tecnici hanno probabilmente risolto il problema (non lentamente).
 “RAPIDLY the technicians have probably solved the problem (not slowly).”

Cinque quotes an observation by Moltmann to the effect that a preposed adverb can cross a higher adverb in a V-2 construction if the adverb is focused, as in the following German example:

- (47) SEHR OFT hat Karl Marie wahrscheinlich gesehen
 “VERY OFTEN has Karl Marie probably seen”

In conclusion, the descriptive statement capturing RM effects with adverbs is quite complex. *Wh*-movement of elements different from D-linked arguments is sensitive to the intervention of *wh*, negation, and quantificational adverbs, but not of non-quantificational adverbs. Adverb movement on the other hand is sensitive to the intervention of any adverb, quantificational or not; but if the moved adverb targets a position different from an adverbial position, say a focus position, then intervening adverbs determine no blocking effect.

In the spirit of the licensing approach mentioned above, adverb positions typically sit in specifier positions of heads possessing certain licensing features. I will call such features “modifier features,” and adopt the hierarchy of the relevant heads arrived at in Cinque (1999): evaluative adverbs like *unfortunately* occupy the specifier position of a higher head than epistemic adverbs like *probably*, which in turn are higher than amount/frequency adverbs,

manner adverbs, etc. Moreover, adverbs may be moved from the IP system to the CP system in cases like (43a–c), (44b), etc. Following Haegeman (1998) I will assume that in such cases the initial adverb “sets the scene” for the following proposition (see also Cinque (1999) on this notion); technically, this can be expressed through the assumption that adverb movement targets the specifier of a modifier head in the CP system where the adverb is interpreted as setting the scene for the event expressed by the following IP.⁹

Now, some modifier features such as amount, extent, frequency, etc. are quantificational features, too, so the Spec positions occupied by adverbs of amount, frequency, etc. have the dual status of modifier and quantificational positions. We end up with (at least) the following two classes of specifier-licensing features in the A'-system:

(48) Quantificational: *wh*, Neg, amount/frequency, . . .

(49) Modifier: evaluative, epistemic, amount/frequency, manner, . . .

The crucial step now is to assume that RM only holds within classes of feature, but not across them. So, the quantificational amount/frequency adverb *beaucoup* will determine a Minimality effect on the quantificational chain aiming at a *wh*-position in (38b), while the pure modifier manner adverb *attentivement* will have no effect on the *wh*-chain in (39b). On the other hand, the chains involved in adverb movement in (43b) and (44c), aiming at a modifier position (the “scene setting position” in the CP system), will undergo Minimality effects triggered from any modifier position, i.e., any intervening adverbial position, no matter whether quantificational or not.¹⁰

Concerning the observed improvement under focus movement, the adverb clearly does not aim at a modifier position in (45c), (46), and (47), but at a focus position, identified by a focus feature distinct from modifier features. The minimality effect is then correctly excluded (we will not take a position here on whether focus belongs to the class of quantificational features or uniquely to another subclass of discourse-related or pragmatic A'-features).

7 Local Actions on Phrases by Heads

In this final section, I will try to show that the scope of our locality principle (4) is not limited to chain formation: it also constrains the local actions that heads can perform on neighboring phrases. Consider a simple syntactic action that a head can be responsible for, such as the licensing of a case feature on a DP. Even if we do not consider the possibility of LF checking (Chomsky 1995b) and just look at surface environments, case licensing remains possible in a small class of local configurations. A head is capable of licensing a case feature on its specifier (50), on its complement (51), and on the specifier of its complement (52):

- (50) a. [He has left]
 b. [His book] is nice
- (51) a. Bill [saw him]
 b. Bill works [with him]
- (52) a. John [believes [him to be a nice guy]]
 b. John [considers [him a nice guy]]
 c. [For [him to do that]] would be a mistake
 d. [With [him sick]], the team is in trouble

but not on the specifier of a higher head or on the complement of a complement. Consider first the specifier of a higher head:

- (53) a. *[A man to [be t in the garden]] is unlikely
 b. *[A man to [come t]] is unlikely
 c. *[John to [call]] would be unlikely
 d. *[This conclusion to be arrived [at t]] is surprising
 (meaning: That this conclusion will be arrived at is surprising)

In (53a, b, c) the verb is unable to assign whatever case feature it possesses (possibly partitive in a, b: Belletti 1988, Lasnik 1992, as in *There is a man in the garden. There came a man*, and accusative in c, as in *John called Bill*) to the DP moved to the specifier of I to satisfy the EPP. More generally, a verb does not determine the case of its subject in VP-external position, which is determined by the properties of the inflectional system; an apparent exception is the case of quirky subjects, but quirky case is plausibly assigned in VP-internal position and then licensed in the inflectional system in a position of structural case assignment. In (53d) the preposition *at* cannot assign case to a DP moved out of its projection.

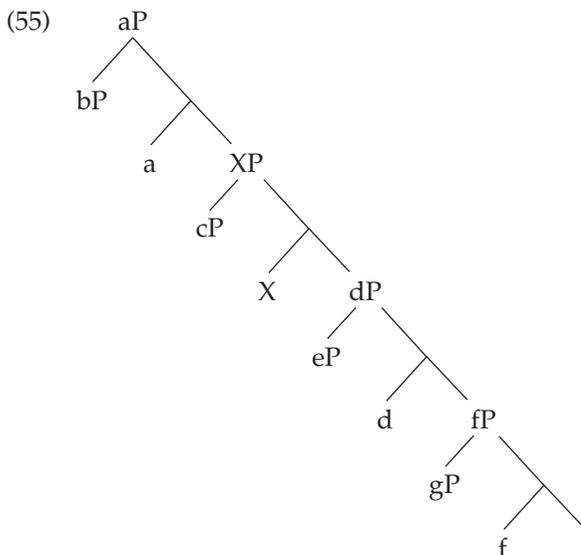
The following cases illustrate the fact that a head cannot license the case on the complement of a complement:

- (54) a. John discussed [(about) it]
 b. John believes [that he/*him is sick]
 c. *John tries [C him to win]
 d. *John wonders [where C him to go]

In (54a) when the preposition is present the verb cannot determine the case on the prepositional object; the preposition does (as is shown by languages in which verbs and prepositions assign morphologically distinct cases). A higher verb cannot determine the case of a lower IP specifier across an intervening C layer, so that the embedded subject in such configurations can only receive whatever case I assigns (nominative in (54b), none or null case (Chomsky and Lasnik's chapter in Chomsky 1995b) in (54c–d)). In general, a higher verb

can never license a case on the embedded subject if an (overt or covert) C intervenes.

The following scheme summarizes the possible surface configurations in which a head can perform case-licensing actions (aP, bP, etc. = arbitrary phrases):



X can license case on its specifier cP, on its complement dP and on the specifier of its complement eP, but not on the specifier of its maximal projection bP or on the complement of its complement fP or on the specifier of the complement of its complement gP, etc. This state of affairs suggests a simple formal generalization: a head cannot license a case across an intervening head, either upward or downward. In fact, in (55) a intervenes between X and bP, d intervenes between X and fP, gP, etc., but no other head intervenes between X and cP, dP, eP. Of course, this is the generalization we expect to hold, if case licensing is constrained by RM. In order to achieve this result formally, some ingenuity is needed, as the case licensing head is sometimes lower and sometimes higher than the DP.

We can reason in the following way: case-licensing procedures typically involve minimal c-command, that is to say the conjunction of c-command and Minimality, along the following lines:

- (56) Case Feature K is licensed on (X, YP) only if YP is in a MC with X and X c-commands/is c-commanded by YP.

(56) states that the minimal configuration, as defined in (4), is always computed from the phrase YP, which is looking for the case-licensing head X, so that any intervening head in the path from the phrase to the head (either upward or downward) will induce a Minimality effect, as desired. On the other hand, the

required direction of c-command between the two elements may be flipped around depending on the specific case feature licensed. Some cases require that the head minimally c-commands the phrase: objective case in many languages, the case assigned by prepositions, case-assigning complementizers, etc. Other cases require that the phrase minimally c-commands the licensing head: nominative case in many languages, accusative under object shift and, in general, OV order, the case assigned by postpositions, genitive in languages like English, etc. Following traditional terminology, we may continue to call “agreement” the configuration in which the phrase minimally c-commands the head, and “government” the configuration in which the head minimally c-commands the phrase, but this terminology should not obscure the fact that these two configurations are not primitive theoretical entities: they are made of the same elementary ingredients, minimality and c-command, and differentiated by the simple flipping around of the direction of the c-command relation. So, the direction of c-command is a simple parameter differentiating, e.g., accusative assignment in English and in OV languages, etc.¹¹

This approach naturally extends to the licensing action of heads on traces. Consider a standard subject object asymmetry like the following:

- (57) a. Who do you think that Bill likes t?
 b. Who do you think that Bill believes t to be innocent?
 c. *Who do you think that t believes John to be innocent?
 d. *Who would you prefer for t to win the race?

(57a–b) illustrate the fact that a verb provides an adequate licensing environment for a trace (in a head-complement or head/specifier-of-the-complement environment), a fact that various versions of the Empty Category Principle have attempted to capture. (57c–d) show that a complementizer (of this sort) does not license a trace; moreover, the higher verbs *think*, *prefer* evidently are too far away from t to have any beneficial effect on it. Similarly, a raising trace is well formed in Italian in a local environment with the raising verb (58a), but not if the infinitival complementizer *di* intervenes (*di* is possible with *seem* in the control construction (58c): see Kayne 1984: ch. 5, Rizzi 1982: ch. 3):

- (58) a. Gianni sembra t essere stanco.
 “John seems t to be tired.”
 b. *Gianni sembra di t essere stanco.
 “John seems ‘di’ t to be tired.”
 c. Mi sembra di PRO essere stanco.
 “It seems to me ‘di’ PRO to be tired.”

So there seems to be a licensing requirement on traces such that they must be minimally c-commanded by a head of a certain kind, i.e., a V counts and a C normally does not (but it may in special circumstances, e.g. through the mechanism illustrated in Rizzi 1990). Again, c-command must be minimal

in that another head cannot intervene between the licensing head and the licensed trace:¹²

(59) X licenses t only if X c-commands t in a LC.

It is worth noticing, as a concluding remark, that from the perspective developed here the possibility of a direct relation between a head and the specifier of its complement naturally falls out without the need to stipulate an independent government relation: this possibility is generated by the simple composition of the elementary relations of c-command and Minimality, and there is no conceptual gain in trying to rule it out. So, in a sense, this system meets the desideratum expressed in Chomsky (1995b) of eliminating head government from the inventory of primitive structural relations, while at the same time preserving the option of a direct action of a head on the specifier of its complement, a possibility which receives strong empirical support from considerations of case licensing and trace licensing.

NOTES

- 1 Mainland Scandinavian Languages disallow the equivalent of (13c) (in non V-2 environments) and still permit (13d): evidently in these languages, contrary to Modern English, there is no absolute ban against moving lexical verbs to I, but only a blocking due to other derivational possibilities, say through Chomsky's (1995b) principle Procrastinate. When no other possibilities exist, e.g. in questions, nothing blocks movement of the lexical verb to C via I. In Modern English, on the other hand, I is specialized for functional verbs (a fact which may in turn be related to the existence of a special class of functional verbs, the modals), hence it disallows movement of the lexical verb, direct movement of the latter to C being banned by the Head Movement Constraint, subsumed under RM.
- 2 In fairy-tale registers of Italian we have structures with proposed

participles across the finite inflectional system:

- (i) Tornato che fu a Milano, . . .
"Come back that he was to Milan, . . ."

But this construction is more akin to remnant VP topicalization in Germanic (den Besten and Webelhuth 1989) than to genuine head movement, as is suggested by the fact that the participle appears in an XP position, to the left of the C *che*, and also by the fact that pied-piping of the verbal complement is only somewhat cumbersome, not categorically excluded:

- (ii) ?Tornato a Milano che fu, . . .
"Come back to Milan that he was, . . ."

I suspect that cases of so-called "long head-movement" to the left periphery often reported in the literature could be analyzed as more

or less disguised instances of remnant VP topicalization.

- 3 I will not address here the problems raised by the postulation of Spec Agr positions higher than the basic position of the subject and to which the object moves for case or other reasons. Clearly, a version of RM simply referring to A-positions is insufficient for such cases. See Chomsky (1995b), Bobaljik and Jonas (1996), Collins and Thráinsson (1993), Haegeman (1993b), and the references cited there for different technical solutions which are expressible into the representational approach adopted here.
- 4 The analogy is only partial. Both D-linked A'-chains and pronominal binding by quantified expressions are insensitive to "weak islands" (intervention of *wh*, negation, quantificational adverbs, etc.); on the other hand, D-linked A'-chains are sensitive to strong islands (extraction out of a relative, and other complex DPs, etc.) while pronominal binding is not. I will continue to assume that all chains (including D-linked A'-chains) are submitted to an independent bounding principle, Subjacency, which is responsible for strong island effects; this principle is not operative on non-chain dependencies such as pronominal binding.
- 5 Argumental *wh*-PPs generally allow extraction from *wh*-islands; so either we extend (35) to PPs, or we assume that reconstruction can always put the preposition back in situ, thus transforming the PP dependency into a DP dependency at LF.
- 6 Notice that if this argument is correct, we have evidence for a representational approach to locality with RM applying at LF: a derivational variant of RM applying

on the attract operation would not naturally draw the distinction between (33a) and (33b), and, in general, between D-linked and non D-linked *wh*-DPs, because *wh*-attraction to the higher C involves the whole DP in both cases.

- 7 Consider also, in this respect, the argument given in connection with (17) and related facts: even if the particle is not attracted by the affix (or by some other functional head, like a *wh*-C) its presence suffices to block movement of V. In general, an approach based on the attraction of a specific feature seems to be too selective to account for the fact that a moving head can almost never skip an intervening head (i.e., the Head Movement Constraint holds very generally, with extremely limited exceptions): this approach predicts free movement of a head across a head not marked for the relevant attracting feature:
 - (i) $X_{+F} \dots X_{-F} \dots X_{+F} \dots$
a freedom which is not empirically supported.
- 8 The same effects are found in German, according to Bartsch (1976: 229), as reported in Cinque (1999):
 - (i) Wahrscheinlich kommt Peter oft
"Probably comes Peter often"
 - (ii) *Oft kommt Peter wahrscheinlich
"Often comes Peter probably"
- 9 Here I differ from Rizzi (1997) where adverb preposing was analyzed, in essence, as a subcase of topicalization. See Haegeman (1998) for arguments in favor of a structural distinction between argument topicalization and adverb preposing.
- 10 It has been observed that certain quantified DPs such as *everyone*,

- someone*, etc. differ from those in (48) in that they do not determine Minimality effects. Without going into the scopal properties of such DPs here, I will simply assume that their LF representations are different enough from those of (48) to allow a structural distinction to be drawn (without necessarily relying on the Specifier/adjunction distinction as in Rizzi 1990).
- 11 If theta-marking involves spec-head and head-complement relations, and cannot involve head/spec-of-the-complement relations (which is not obvious: Rizzi 1992b), this may be due to the fact that theta-marking involves some kind of binding of theta-slots in the grid associated to the theta-assigner (Stowell 1981), so that c-command from the assignee is demanded. If this view is correct, a head could not theta-mark the specifier of its complement because the latter would be unable to bind the appropriate slot in the theta-marker's grid.
- 12 In the spec-head relation and in the head/spec-of-the-complement relation asymmetric c-command holds (from the phrase and from the head, respectively), whereas heads and complements symmetrically c-command each other. Does this have adverse consequences? The issue may arise e.g. for the licensing of past participle agreement in French, not possible in situ but possible when the object moves:
- (i) a. J'ai mis(*e) la voiture dans le garage
 "I have put the car in the garage"
 b. Je l'ai mise dans le garage
 "I have put it in the garage"
- If this instance of Agr requires minimal c-command by the phrase, why is this requirement not satisfied by the object in its in situ position? But notice that under current assumptions this agreement specification is not licensed in the VP, but on a higher autonomous X-bar subtree headed by an Agr node. So, in (ia) the object is not in a position to license Agr, and the problem does not arise (see Belletti, this volume, and references cited there).
- Is there any genuine case of Agr licensing under minimal c-command by the complement? Our approach may lead us to expect it. One reasonable candidate may be C-agreement in Germanic (Haegeman 1992), which could then be analyzed as involving direct agreement from C to IP rather than LF (feature) movement of the subject to C. Perhaps also the pervasive agreement phenomena within the Romance DP (*I suoi molti bei libri* = The+MP his+MP many+MP beautiful+MP books+MP), not easily amenable to spec-head configurations, may be analyzed in this way.