(13)

Mouton de Gruyter
A Division of Walter de Gruyter & Co. · Publishers

# Copyright Transfer Agreement

concerning the article entitled: Locality Theory: Towards a Unification

by: M. Rita Manzini

to be published in the work: **Levels, Principles, and Processes**

edited by: **W. Chao and G. Horrocks**

By their signature, the Author(s) warrant(s) that the article specified above in no way infringes upon any copyright or proprietary rights of others, and is in no way libelous or unlawful, and hereby transfers the copyright to this article to Walter de Gruyter,[1] effective immediately. The copyright transfer covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microform, electronic database, videodisks, or any other reproductions of a similar nature, and translations.

However, the Author(s) reserve(s):

1. All proprietary rights other than copyrights, such as patent rights.

2. The right to use all or part of this article in future works of their own published more than two years after the original publication in the above-named work and to grant or refuse permission to third parties to republish all or part of the article or translations thereof. To republish whole articles, such third parties must also obtain written permission from Walter de Gruyter. However, Walter de Gruyter alone may grant rights concerning the work as a whole.

The Author will receive one free copy of the work upon publication. In cases of multiple authorship, each Author will receive one copy of the work.

This Agreement must be signed by the Author. In cases of multiple authorship, this Agreement must be signed by at least one of the Authors, who agrees to inform the other(s):

Date: 19-6-95

Author's signature: M Rita M

[1] for US Government employees: to the extent transferable

Main office:
Genthiner Strasse 13
D-10785 Berlin
Federal Republic of Germany

Tel.: (+49-30) 260 05-235
Fax: (+49-30) 260 05-251

(letters only):
Postfach 30 34 21
D-10728 Berlin
Federal Republic of Germany

Telex: 184027 wdg d
E-mail: 100064.2307@compuserve.com

For North America
(sales and distribution):
Walter de Gruyter, Inc.
200 Saw Mill River Road
Hawthorne, NY 10532, USA
Tel.: (914) 747-0110
Telex: 64 66 77
Fax: (914) 747-1326

M. Rita Manzini

# LOCALITY THEORY: TOWARDS A UNIFICATION

Since Chomsky (1981), it has been assumed that under the ECP, traces must satisfy one of two requirements. Either they have a government relation to a head, or else they must have a government relation to an antecedent. This paper explores two questions, both related to the ECP. The first question is what determines the class of elements that must undergo antecedent-government; or in other words what is the government relation to a head which satisfies the ECP independently of antecedent- government. Recent proposals in the literature include Chomsky's (1986b) theta-government, and Rizzi's (1990) or Cinque's (1991) referential indices; our thesis will be that none of these is in fact correct. The second question concerns the formulation of antecedent-government itself, in particular the question whether antecedent-government makes use of an antecedent-based conception of locality, as under Rizzi's (1990) Relativized Minimality, or of a non-antecedent-based notion, as under Chomsky's (1986b) (rigid) Minimality. Our answer will favor the latter option.

In section 1, we argue that if the relation to a head relevant for the satisfaction of the ECP is theta-based, as in the references quoted, then traces of A-movement systematically satisfy it and cannot be predicted to undergo antecedent-government without stipulation. Instead we suggest that the relevant relation to a head is visibility-based. Thus A-traces are systematically predicted to be subject to antecedent-government in that they are not visible. The largest part of our discussion will be concerned with dealing with a number of arguments, crucially involving LF movement, that Rizzi (1990) and Cinque (1991) have put forward in favor of the notion of referential index. If we are correct, those arguments, contrary to our

argument from A-traces, do not choose between the available theories.

In section 2 on the other hand, we consider the arguments that Rizzi (1990) puts forward in favor of Relativized Minimality as opposed to (rigid) Minimality, namely two types of islands, pseudo-opacity and inner islands. We argue that there is an important theoretical reason why a Minimality-based formulation of antecedent-government is to be preferred to one which is based on Relativized Minimality, namely that the latter makes any unification of antecedent-government with Subjacency impossible. Crucially we show that both pseudo-opacity and inner islands can be accounted for under Minimality, if a restrictive theory of A'-positions is assumed under which at most one A'-position is associated with each maximal projection. Thus if we are correct the empirical argument in favor of Relativized Minimality from pseudo-opacity and inner islands disappears.

Though the issue is not pursued any further in this article, the disjunction between antecedent-government and Subjacency can be completely eliminated as in Manzini (1988); and the ECP itself can be given a non-disjunctive formulation as in Manzini (1990; to appear).

## 1. Against theta-government and referential indices

As discovered in Huang (1982), the locality behavior of wh-dependencies created by arguments differs from that of wh-dependencies created by adjuncts. Thus following Chomsky (1986b), while extracting an adjunct out of a wh-island gives ungrammatical results, as in (1), extracting an argument produces a fundamentally grammatical sentence, as in (2):

(1)*How do you wonder [what to repair t t]

(2)What do you wonder [how to repair t t]

According to Chomsky (1986b), the two classes of elements that do and do

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

(3) $\beta$ theta-governs $\alpha$ iff

$\beta$ is a head, $\beta$ theta-marks $\alpha$ and $\beta$ is a sister of $\alpha$

According to Rizzi (1990), however, the notion of theta-government is empirically inadequate in at least three respects. First, adverbs that obligatorily occur 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, undergoing wh-islands, as 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 is substituted by a notion of referential index, where a referential index is licenced under theta-role assignment, provided the theta-role is a referential one, as in (7). 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 though each of them is theta-marked, it is not referentially theta-marked:

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

In Cinque (1991), furthermore, the notion of referential index is

extended 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; or to the contrast between (10) and (11), where a D-linked and non-D-linked wh-phrase, in the sense of Pesetsky (1987), are extracted from wh-islands. The first member of the pair is systematically more acceptable than the second, which can be accounted for, according to Cinque (1991), if existential quantifiers such as qualche and D-linked wh-phrases have a referential index, while a distributive universal quantifier such as ogni and non-D-linked wh-phrases do not:

(8)Qualche museo, mi chiedo chi visitera'
Some museums, I wonder who will visit

(9)*Ogni museo, mi chiedo chi visitera'
Each museum, I wonder who will visit

(10) Che macchina ti chiedi come riparare
Which car do you wonder how to repair

(11) *Che diavolo ti chiedi come riparare
What the hell do you wonder how to repair

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

(12) A referential index must be licenced by referential content

What we want to suggest here is an alternative to all three conceptions

in (3), (7) and (12). Chomsky (1986b), Rizzi (1990) and Cinque (1991) agree that the constraint that determines wh-island violations, namely antecedent-government, is also the constraint that compels movement to an A-position to be strictly local, as exemplified in (13)-(14). If so, antecedent-government ranges argument wh-dependencies against not only adjunct wh-dependencies but also A-movement dependencies:

(13) John appears [t to be likely [t to win]]

(14) *John appears [that it is likely [t to win]]

The question then arises whether the theories in (3), (7) and (12) correctly predict this property of A-movement. The answer is obviously negative with respect to (3). Under (3), A-movement is predicted to pattern with wh-movement of arguments in that movement takes place from a theta-governed position in both cases. Given (3), then, is is clear that the class of elements subject to antecedent-government must be disjunctively characterized, including both non-theta-governed elements and A-traces.

Similarly given (7), since A-movement in general takes place from positions assigned a referential theta-role and hence a referential index, we expect the class of elements subject to antecedent-government to be disjunctively characterized as including both non-referentially indexed elements and A-traces. Instead, Rizzi (1990) includes antecedent-government in the definition of chain. Since A-traces are compelled to form chains to satisfy the Theta-Criterion, they are compelled to satisfy antecedent-government. Of course, wh-traces of adjuncts are compelled to form chains and hence to satisfy antecedent-government for a different reason. In particular, since chains do not require coindexing, and adjuncts do not have a referential index, chains represent the only means through which an adjunct wh-trace can establish a relation with an antecedent. Thus the disjunction between non-referentially indexed elements and A-traces is resolved only at the price of

an enrichment of the notion of chain.

Finally under (12), a disjunctive characterization of the elements that are subject to antecedent-government can be avoided by assuming, as in Cinque (1991) that while a trace of wh-movement is referentially indexed or not, depending on the content of the phrase moved, a trace of A-movement is never referentially indexed because of its own lack of referential content. But this represents just a stipulation, making the theory equivalent to one in which the elements subject to antecedent-government are disjunctively characterized.

What we want to suggest here is that though a non-disjunctive characterization of the class of elements that undergo antecedent-government is impossible given notions based on theta-marking and/or inherent content, such as those in (3) and (7)/(12), it is possible given a visibility-based notion. Indeed A-traces are standardly assumed not to be Case-marked and hence visible, and the same can be assumed to hold of adjuncts. By contrast argument wh-traces are typically Case-marked and visible. Hence the notion of Case-marking or visibility will generally range A-traces and wh-traces of adjuncts against wh-traces of arguments, exactly as desired for the purposes of antecedent-government.

As a first approximation, we can define a notion of K-government so that $\beta$ K-governs $\alpha$ just in case $\beta$ is a head and $\beta$ Case-marks $\alpha$, as in (15). Under (15), an argument wh-trace, as in (2) is K-governed, but an A-trace, as in (13)-(14), is not; nor, we can assume, is an adjunct wh-trace, as in (1). If so, the antecedent-government requirement can be correlated non-disjunctively with lack of K-government. Wh-traces of adjuncts and A-traces are immediately predicted to be subject to it, as opposed to wh-traces of arguments, which are not:

(15) $\beta$ K-governs $\alpha$ iff

$\beta$ is a head and $\beta$ Case-marks $\alpha$

Thus our contention is that a theory that includes a Case-based notion such as K-government can predict that A-traces behave identically to adjunct wh-traces from the point of view of locality. Any theta-based theory predicts that A-traces pattern with wh-traces of argument, and can only derive that A-traces are subject to antecedent-government only by stipulating it separately, as in Chomsky (1986b) and Cinque (1991), or by otherwise enriching the grammar, as in Rizzi (1990).

Still, leaving aside the core cases in (1)-(2) and (13)-(14), that are indeed accounted for under (15), 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), as well as any evidence in favor of Chomsky's (1986b) original notion of theta-government, can also be handled by (15). Consider Cinque's (1991) evidence. 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 wellformed results across a wh-island, as in (10). By contrast extraction of a non-D-linked wh-phrase is illformed across a wh-islands, as seen in (11). Such contrasts are of course 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-phases cannot.

Cinque (1991) furthermore shows that Italian bare quantifiers such as qualcuno ("someone") either behave as other quantified phrases associated with a resumptive clitic under Clitic Left Dislocation, as in (16); or else are not associated with a resumptive clitic and hence bind a variable, as in (17). When construed as in (16), qualcuno is not sensitive to wh-islands, though it is when construed as in (17), as shown by the contrast between (18) and (19). In order to capture this behavior, Cinque (1991) has recourse to the

assumption that bare quantifier can be used referentially or non-referentially; a bare quantifier used non-referentially will then behave as an operator, binding a variable:

(16) Qualcuno penso che lo troveremo

Someone I think we will find (him)

(17) Qualcuno penso che troveremo

Someone I think we will find

(18) Qualcuno mi chiedo chi lo trovera'

Someone I wonder who will find (him)

(19) *Qualcuno mi chiedo chi trovera'

Someone I wonder who will find

Finally, Cinque (1991) considers scope reconstruction phenomena of the type in (20)-(21), originally discussed in Longobardi (1991). In (20) the wh-phrase can be interpreted as having narrow scope with respect to the embedded quantifier; in (21), however, which differs from (20) 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 the wh-phrase can interact in scope with another quantifier only if it is non-referential, even though it can be referential when it does not interact with another quantifier:

(20) Quanti pazienti pensi che ogni medico visiti

How many patients do you think that every doctor visits

(21) Quanti pazienti ti chiedi come ogni medico visiti

How many patients do you wonder how every doctor visits

All of the data in (8)-(11) and (16)-(21) raise a potential problem for our theory. Indeed all of the phrases moved in (8)-(11) and (16)-(21) are apparently K-governed, so that K-government appears to be unable to predict any of the relevant contrasts. Consider however the further paradigm in (22)-

(25). Sentences of the type in (22), where no wh-island is crossed, are perfectly wellformed in Italian with extraction of quanti ("how many") and stranding of the clitic ne ("of them"); but sentences of the type in (23), where quanti is moved across a wh-island, again leaving behind ne, are illformed. Examples of the type in (24) where quanti does not strand the nominal part of its phrase are also wellformed; in this case extraction out of a wh-island appears to be acceptable, as in (25):

(22) Quanti pensi che ne visitera'
How many do you think he will visit of them

(23) *Quanti ti chiedi chi ne visitera'
How many do you wonder who will visit of them

(24) Quanti pazienti pensi che visitera'
How many patients do you believe he will visit

(25) Quanti pazienti ti chiedi chi visitera'
How many patients do you wonder who will visit

Under the theory in Cinque (1991) it can of course be assumed that quanti is not referential, thus predicting the illformedness of (23). In this case however a solution is also available within our theory. Under our usual assumptions the DP quanti-ne is Case-marked and hence K-governed by the embedded V. On the other hand, quanti presumably corresponds to the D head of DP. If so, we only need to say that Case-marking always is to a maximal projection, to predict that a D head is never K-governed. Hence the fact that quanti is subject to antecedent-government, as in (23), is correctly predicted under the notion of K-government.

Let us now consider the examples in (24)-(25) more closely. While they both are wellformed, they do not necessarily have the same interpretation. Suppose then that the two options available in the syntax, in (22) and (24)-(25), are available at LF as well. In other words the whole DP can be in

operator position, or only the D, as suggested in Dobrovie-Sorin (1990; to appear). This means that there are two LF's corresponding to examples of the form in (24)-(25). The first type of LF straightforwardly corresponds to the S-structure configuration, as in (26). The second type of LF corresponds to the S-structure configuration in (22) where the quantifier head has been raised out of its DP, as in (27):

(26) a. [How many patients: x] you think that he will visit x

b. [How many patients: x] you wonder who will visit x

(27) 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 (26) the DP is of course Case-marked and hence K-governed by the embedded verb; and antecedent-government is predicted not to apply. Consider on the other hand the properties of the LF's in (27). D is not K-governed by the embedded verb, on the assumption that Case-marking always is to a maximal projection. This means that (27b) is excluded as an antecedent-government violation. Our theory then predicts that an example like (24) can be associated with both the interpretations in (26a) and in (27a), while an example of the type in (25) can only be associated with the interpretation in (26b). This prediction corresponds to native speakers intuitions, in that (24) can have an interpretation, the obvious one for native speakers, which (25) cannot have. If so, our hypothesis that the two different LF's in (26) and (27) are possible is also confirmed; for, the different range of interpretations for (24) and (25) has now been derived from this hypothesis.

With this much background, we can now return to the original problems in (8)-(11) and (16)-(21). Consider first (8)-(9). The contrast between (8) and (9) can be explained in terms of our theory without any appeal to referentiality, if we assume that a ogni-type quantifier, as in (9), is compatible only with an LF at which D raising has taken place, and not with an

LF where the whole DP is in operator position. For, D is not K-governed, hence antecedent-government must be satisfied; whence the ungrammaticality of (9). 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-governed and hence not subject to antecedent-government.

Similarly, consider (10)-(11). The claim in our 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 for a wh-island to be crossed, as in (11), reduces to the fact that D is not K-governed, and hence antecedent-government must be satisfied. 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-governed and hence the antecedent-government requirement does not apply.

As for (16)-(19), no stipulation about the same quantifier being either referential or non-referential is needed under our theory. 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-governed and able to cross wh-islands, as in (18), since antecedent-government does not apply. If it moves as a head D, it will not be K-governed, hence it will be subject to antecedent-government and not able to cross wh-islands, as in (19). In the latter case it will also not admit of a resumptive pronoun. Finally, with respect to (20)-(21) we can assume that scope interaction between quantifiers requires an LF with D's in operator position. Because of the lack of K-government, this LF is possible only when antecedent-government is satisfied, hence in the absence of a wh-island. This means that scope interaction between the two quantifiers is predicted to be possible in (20),

but not in (21), correctly.

Crucially, our theory is not any more complex than Cinque's (1991). Our link between antecedent-government and lack of K-government is of course strictly comparable with Cinque's (1991) link between antecedent-government and lack of referential indices. As for the option of raising either D or DP, this must be available in all theories, given examples of the type in (22)-(23). What our theory says is that the two interpretations possible in examples of the type in (24)-(25) correspond to the two LF's in (26) and (27). This is strictly comparable to saying that the two interpretations in (24)-(25) correspond to the presence or absence of a referential index. Finally, via the notion of referential index, the relevant interpretations are identified in Cinque (1991) as referential and non-referential respectively. We leave the relevant interpretations undefined, but given the lack of independent semantic characterization for referentiality in Cinque (1991), the matter appears to be purely terminological.

The potential problem posed by Rizzi's (1990) selected adverbs, idiom chunks and measure phrases, as in (4)-(6), remains then 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 from antecedent-government. By contrast our theory apparently predicts examples of the type in (6) to be wellformed, on a par with examples of the type in (21). One hypothesis that we can pursue in this respect is that (6) is illformed to the extent that not only it is incompatible with a D raising LF for syntactic reasons like (21), but it is also incompatible with a DP raising LF for semantic reasons, unlike (21). Indeed a similar explanation is necessitated by the conception of referential indices in Cinque (1991), as

opposed to Rizzi's (1990), if the former is construed as substituting for the latter rather than adding to it. Under (12), as opposed to (7), it must be a semantic property of measure phrases that they are not be associated with referential indices, so that (6) is blocked by antecedent-government.

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 in 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 furthemore 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 need not be theta-marked under Chomsky's (1981; 1986a) Theta-Criterion; hence under Chomsky's (1981; 1986a) Visibility it need not be Case-marked. We also know that only arguments can be theta-marked under the Theta-Criterion. If we assume that under Visibility only arguments can be Case-marked, then NP must also be Caseless. This means that an idiom chunk is not K-governed and hence it must be antecedent-governed. Its sensitivity to wh-islands as in (5) then follows.

Similarly, consider selected adverbs, as in (4). The suggestion in 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 need not satisfy Visibility and hence need not be Case-marked. More strongly, we can assume that adverbs cannot be visible, and hence they cannot be Case-marked. If so, adverbs are never K-governed, and the need for them to safisfy antecedent-government is straightforwardly predicted under our theory as well. Indeed though Rizzi (1990) uses selected adverbials and idiom chunks to argue against the notion

of theta-government in Chomsky (1986b), the argument does not really go through. In order for the notion of theta-government to be saved, it is sufficient to adopt something like the theory proposed here under which idiom chunks and adverbs are not arguments, hence they are not theta-marked.

Having now reviewed the empirical evidence in favor of Rizzi's (1990) and Cinque's (1991) notion of referential indices, a general conceptual issue remains to be taken into consideration. Given the two notions of referential theta-role in Rizzi (1990) and of referential content in Cinque (1991), we have seen that it is only the latter that can subsume the former, and not viceversa. In our discussion we have then used the fact that referential content is not independently characterized in Cinque (1991) as part of our argument that referential indices theories have no explanatory advantage over ours.

Consider however the reconstruction of Rizzi's (1990) and Cinque's (1991) theory in Frampton (this volume; to appear) or Szabolcsi (1991), 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 by the presence of an individual variable, as in Heim (1987). It remains for us to show that the existence of a successful semantic characterization for those dependencies that are insensitive to wh-islands does not substantially change the terms of our discussion.

In Frampton (this volume), two representations of the type in (26)-(27) are motivated on semantic grounds for sentences involving extraction of <u>how many</u>, where (26) represents the individual variable reading, (27) the alternative reading. Once given (26)-(27), however, it is obvious that our discussion goes through as before. Consider then Longobardi's (1991) scope

interaction phenomena. What is involved apparently is the inability of an individual variable to be construed with anything but wide scope. If we assume as before that a DP variable is interpreted as an individual variable while a D variable is associated with the non-individual reading, once again our discussion goes through unchanged. Similarly, Szabolcsi (1991) comments that D-linked readings are always individual readings. But if individual readings correspond to DP variables, then the correlation of D-linking and DP variables holds, as desired. And so on.

In short, the logic of our discussion is as follows. 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. However under K-government, a DP is syntactically visible, while a D variable is not. This accounts for the fact that individual readings remain available across wh-islands, hence when antecedent-government is violated, but non-individual readings disappear.

We can now return to other potentially problematic predictions of our theory. If visibility is the relevant criterion for extractability from wh-islands, then VP's, which are not Case-marked, are predicted not to be extractable from wh-islands. Chomsky (1986b) however notices that extraction of a VP across a wh-island does not create a violation comparable to that produced by the extraction of an adjunct in the same context. On the basis of this, Chomsky (1986b) concludes that VP's are theta-governed by I, from which it follows that they need not be antecedent-governed. Under Rizzi's (1990) theory the data similarly require the assumption that VP's have a referential index, enabling them to move long-distance. But the conclusion that VP is argumental, as in Chomsky (1986b), or that it has referential properties, as

in Rizzi (1990), is as questionable as any conclusion that VP is visible. Thus at the conceptual level, if not at the realization level, the extractability of VP across a wh-island is a problem for all theories under consideration.

Notice on the other hand that none of the extractions from wh-islands of non-visible or non-referential constituents considered in Cinque (1991), as in (8)-(11) and (16)-(21), has the same uninterpretable results as extraction of an adjunct. On the contrary, it appears that the contrast between extractability of a VP in the absence of a wh-island and across a wh-island, as in (23)-(29), fully repeats contrasts of the type in (8)-(9), etc. If so, there is no need for any problematic assumption to the effect that VP's are arguments and/or they are visible so that they will be antecedent-government. Rather, no necessary correlation between strong ungrammaticality and antecedent-government violations holds, contrary to what suggested in Chomsky (1986b):

(28) Mangiato una mela, credo che abbia
Eaten an apple I believe that she has

(29) *Mangiato la mela, non so chi abbia
Eaten the apple, I wonder who has

Finally, consider whether our theory predicts that traces of V- or I-movement also fall under the antecedent-government requirement. Theta-based theories such as Chomsky's (1986b) or Rizzi's (1990) are unproblematic in this respect, since VP or IP are never theta-marked, hence V or I are never theta-governed or referentially indexed respectively. But our theory also yields the correct predictions, since neither VP nor IP are ever Case-marked; thus neither V nor I are ever K-governed. This of course means that they must satisfy antecedent-government.

On the other hand D or N movement appears to have exactly the same locality properties as V-movement. Consider for instance cliticization.

Following notably Kayne (1989), strictly local cliticization is possible, but not long-distance cliticization, as in French (30)-(31). Of course N/D heads differ from V/I heads in that their maximal projection, DP/NP is Case-marked. However as in our discussion of D-raising, we only need to assume that Case-marking is always to a maximal projection to predict that only a maximal projection, and never a head, can be K-governed. It is worth pointing out then that within our theory it is necessary to constrain the syntactically relevant notion of Case or visibility to maximal projections, no matter whether Case is in fact realized on heads. If so, head-movement is indeed predicted to be subject to antecedent-government in all cases, as desired :

(30) Jean le veut [t]
John it+wants

(31) *Jean le veut voir [t]
John it+wants to see

## 2. Against Relativized Minimality

So far we have been referring to antecedent-government as the condition on non-visible traces, without providing any definition for it. In the literature, essentially two definitions are available. One is provided in Chomsky (1986b) and its core is a notion of Minimality. The second corresponds to another major theoretical proposal put forward in Rizzi (1990), namely Relativized Minimality. In what follows we will consider them in turn.

According to Chomsky (1986b) antecedent-government is defined in terms of the notion of barrier in (32), as in (33). To be precise, the definition of barrier in (32), or minimality barrier, is only one of two definitions of barrier that enter into antecedent-government; (32) however is sufficient for the purposes of our discussion. The definition in (32) itself is recast with

respect to Chomsky's (1986b); in particular no exceptional status is accorded to IP, and the phrasal projection XP, rather than X', is conceived of as the 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:

(32) $\beta$ is a barrier for $\alpha$ iff

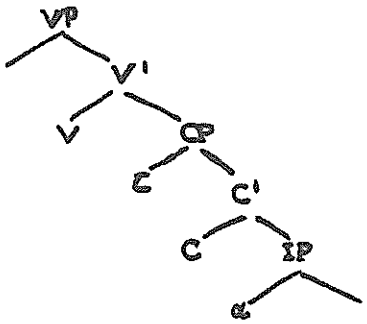
$\beta$ is an XP, $\beta$ dominates $\alpha$ and $\alpha$ is not the head or Spec of $\beta$

(33) $\beta$ governs $\alpha$ iff

there is no barrier for $\alpha$ that excludes $\beta$

Objects, which are typically theta-governed, or in our terms K-governed, typically 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 respectively, and must therefore satisfy antecedent-government, do give rise to wh-island violations. Thus antecedent-government accounts for the contrast in (1)-(2). The relevant structure for the violation in (1) is provided in (34). Since we are assuming, contrary to Chomsky (1986b), that IP does not have exceptional status, the adjunct is shown in the IP-adjoined position $\alpha$. We assume that such a position can be reached by the adjunct starting either from an IP-internal position or from a VP-internal one, in the latter case via adjunction to VP. Of course IP in (34) is not a barrier for $\alpha$, since it does not dominate it in Chomsky's (1986b) sense. Hence if $\alpha$ can move to the Spec of CP position, no barrier is crossed. If the Spec of CP position is filled, however, CP must be crossed, where CP is a barrier for $\alpha$ in (34). Indeed CP dominates $\alpha$, which is not in a head or Spec escape hatch position. Thus the crossing of the CP barrier violates antecedent-government, correctly predicting the illformedness of (1):

(34)




Consider now Rizzi's (1990) theory. This is based on the idea that antecedent-government is crucially constrained by Relativized Minimality, as in (35); informally, antecedent-government is blocked whenever a potential antecedent-governor is crossed. The notion of typical potential antecedent-governor, which enters into (35), is defined as in (36); 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 so on:

(35) Relativized Minimality

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-command Y and does not c-command X

(36) Z is a typical potential antecedent-governor for Y,

Y in an { A-chain iff Z is an A Spec / A'-chain iff Z is an A' Spec / head-chain iff Z is a head } 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 (34). 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 again correctly predictted.

Crucially, though the theories of antecedent-government in Chomsky (1986b) and Rizzi (1990) are empirically equivalent with respect to (1)-(2), two further types of islands are introduced in support of Relativized Minimality in 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 under the simplest hypothesis the component of the theory involved in predicting them must be antecedent-government.

Consider first pseudo-opacity. The relevant data are of the type in (37)-(40). Both of the extractions in (37) and (38) are wellformed in French, where in (38) the quantifier combien moves stranding the rest of its DP, while in (37) the whole DP moves. Consider however the contrast in (39)-(40). (40), where only combien moves across the quantificational adverb beaucoup, is illformed; (39), where the whole DP moves across the adverb, is wellformed. In our terms, movement of the D head combien out of its DP is predicted to give rise to antecedent-government effects, because it is not K-governed. Thus movement of combien across a wh-island in (42) is correctly predicted to be illformed, and movement of the whole DP to be wellformed, as in (41). The question however is whether the contrast between (39) and (40), and in particular the illformedness of (40), can also be accounted for:

(37) Combien de livres a-t-il consulte'
How much of books has he consulted

(38) Combien a-t-il consulte' de livres
How much has he consulted of books

(39) Combien de livres a-t-il beaucoup consulte'

How much of books has he much consulted

(40) *Combien a-t-il beaucoup de livres

How much has he much consulted of books

(41) Combien de problemes sais-tu comment resoudre

How much of problems do you know how to solve

(42) *Combien sais-tu comment resoudre de problemes

How much do you know how to solve of problems

Under Rizzi's (1990) theory a Relativized Minimality island is straightforwardly defined in (40) 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 (42); hence the illformedness of (42) follows, since movement takes place across a potential antecedent-governor. What is more, Relativized Minimality correctly predicts the contrast between (43) and (44), since in (44) an adjunct, comment, is extracted across the floated quantifier beaucoup, a potential antecedent-governor, while in (43) no potential antecedent-governor is crossed. Unfortunately, Chomsky's (1986b) Minimality does not yield any prediction with respect to either (40) or (44). In both cases the lower VP is a barrier for the non-theta-governed, or in our terms 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:

(43) Comment a-t-il resolu beaucoup de problemes

How did he solve much of problems

(44) *Comment a-t-il beaucoup resolu de problemes

How did he much solve of problems

Consider then inner islands. (46) differs from (45) only in the presence vs. absence of a sentential negation. (46) is illformed with the interpretation under which the adjunct is extracted from the embedded sentence, as indicated, though it is wellformed with the interpretation, irrelevant here, under which the adjunct is associated with the matrix verb. (45) is wellformed under both interpretations. Inner islands are the islands created by the negation in sentences of the type of (46). Crucially, no inner island effects arise with objects, as in (47):

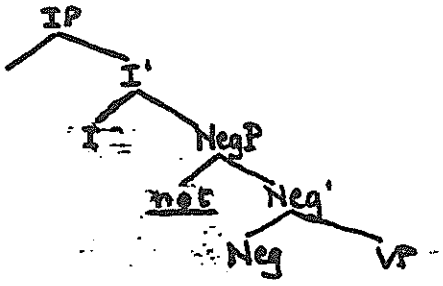
(45) Why do you think John left

(46) *Why don't you think [Peter left t]

(47) 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, as in Pollock (1989) or Chomsky (1989), the negation can be taken to fill the Spec position of NegP. If so, the resulting phrase structure configurations are then of the type in (48), on the assumption that NegP is generated above VP and under IP:

(48)




Given structures of the type in (48), the extraction of the adjunct in (46) must at some point cross not, a potential antecedent-governor for the wh-phrase in virtue of its A' Spec position. Hence antecedent-government does not hold under Rizzi's (1990) Relativized Minimality, and illformedness is correctly predicted to arise. Unfortunately, as before, no island is defined

in (46) under Chomsky's (1986b) Minimality; as the VP barrier can be escaped, the NegP barrier can be escaped by adjoining to it. Hence no predictions as to ungrammaticality of (46) follow under Chomsky's (1986b) 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 islands which remain unexplained under Minimality. Though the empirical evidence appears to favor Relativized Minimality, Minimality has at least one potential theoretical advantage over it. Wh-islands of the type in (1)-(2) or indeed pseudo-opacity and inner islands constitute only a subset of the islands to which (wh-)extraction is sensitive. Others include subject islands as in (49)/(52), adjunct islands, as in (50)/(53) and complex NP islands, as in (51)/(54), to which argument and non-argument wh-phrases are subject alike:

(49) *What does [explaining t] bother you

(50) *What was Mary bothered [because Peter explained t]

(51) *What do you know the girl [that explained t]

(52) *How did [fixing it t] bother Mary

(53) *How was Mary bothered [because John had fixed it t]

(54) *How do you know the girl [that fixed it]

Crucially, none of the islands in (49)-(54) can be antecedent-based, precisely because argument wh-traces are not sensitive to antecedent-based islands, but are sensitive to these islands. It appears then that any theory that embeds an antecedent-based definition of locality for (1)-(2) and similar islands must of necessity be disjunctive, since it must introduce a non-antecedent-based notion of locality for islands of the type in (49)-(54). Since Minimality however is a rigid notion, there is no principled reason why an appropriate definition of it could not cover non-antecedent-based islands. Hence there is no inbuilt reason why a theory that includes Minimality should

be disjunctive. This appears a good reason to investigate Relativized Minimality's claim to empirical superiority further.

Let us 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, without recourse to Relativized Minimality, by having recourse instead to the fact that the adjunct wh-phrase can move neither through the Spec of CP, which is already filled by hypothesis, nor through a CP-adjoined position, since following Chomsky (1986b) 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 radical 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 (44). 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 a VP-adjoined position.

Now, let us assume that in the absence of antecedent-government requirements, argument wh-phrases can move in a single step to their final landing site. If so, not only they never need use the Spec of CP, thus accounting for the wellformedness of (2), but they never need use the Spec of

VP either, thus accounting for the wellformedness of (39). 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 empirical consequences.

Much the same can of course be repeated for NegP. As for VP, the availability of multiple adjunction to NegP is irrelevant for argument wh-phrases, on the assumption that in the absence of antecedent-government 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 is reintroduced into the theory, essentially the reverse of Chomsky's (1986b) 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 impossibile 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. Unfortunately, this hypothesis raises a 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 (55), disregarding the question whether this is to be conceived as a Spec position or as an adjoined one. The question as to whether IP obeys (55) 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 (55) that there is at most one A'-position per maximal projection, can be maintained. Under (55) 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, exactly as a wh-phrase in the A'-Spec of CP does. This forces the crossing of a barrier and hence an antecedent-government violation:

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

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 (56), taken from May (1985). May (1985), in fact, basing

himself on Gueron (1981), explicitly suggests that a syntactic principle of the type in (55) holds. According to May (1985), multiple adjunction to IP, as in (56), is unnecessary. Instead (56) is argued to be equivalent to (57), where one quantified DP is adjoined to IP, and the second one is adjoined to the first:

(56) $[_{IP}$ Every spy $_i$ $[_{IP}$ some Russian$_j$ $[_{IP}$ $t_i$ suspects $t_j]]]$

(57) $[_{IP}$ $[_{DP}$ 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 (56) and (57) the two operators have the same absolute scope, since they have the same c-commands domain, or equivalently they commute in relative scope, since they c-command each other. Technically, May (1985) defines a sigma-sequence as in (58), where O stands for an operator. Under (58), it is obvious that the two quantified DP's form a sigma sequence in both (56) and (57). 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 (59); under (59), (56) and (57) give rise to exactly the same range of intepretations, as desired:

(58) 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 $O_i$

(59) 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 (57), we on the other hand expect that two wh-phrases that commute in scope can be adjoined to one another and can therefore cooccurr in the same Spec of CP. Since clusters of wh-phrases are in fact overtly excluded in English, as in (60), 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,

since a head is crossed. The revision of Relativized Minimality proposed in Roberts (1991) on the other hand predicts grammaticality, since movement takes place to an A'-head, C, across an A-head, Agr. Evaluating all of the potentially relevant evidence is beyond the scope of this article. It is striking however that in examples of the type in (62), Pollock's (1989) basic test for the conclusion that two different head positions are involved fails. In other words, no phrasal material, such as a subject, can ever intervene between the non-finite verb and the clitic+auxiliary complex.

REFERENCES


Baker, M. and K. Hale (1990), "Relativized Minimality and Pronoun Incorporation", LI 21.2

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

Chomsky, N. (1986a), Knowledge of Language, Praeger, New York

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

Chomsky, N. (1989), "Some Notes on Economy of Derivation and Representation", MIT Working Papers in Linguistics 10

Cinque, G. (1991), Types of A'-Dependencies, MIT Press, Cambridge, Mass.

Dobrovie-Sorin, C. (1990), "Clitic Doubling, Wh-Movement, and Quantification in Romanian", LI 21.3

Dobrovie-Sorin, C. (to appear), The Syntax of Romanian, Foris, Dordrecht

Frampton, J. (this volume), "The Fine Structure of Wh-Movement and the Proper Formulation of the ECP"

Frampton, J. (to appear), "Relativized Minimality, A Review", The Linguistic Review

Gueron, J. (1981), "Logical Operators, Complete Constituents, and Extraction Transformations", in R. May and J. Koster (eds.), Levels of Syntactic Representation, Foris, Dordrecht

Heim. I. (1987), "Where Does the Definiteness Restriction Apply?", in E. Reuland and A. ter Meulen (eds.), The Representation of (In)definiteness, MIT Press, Cambridge, Mass.

Huang, J. (1982), Logical Relations in Chinese and the Theory of Grammar, Ph.D. Dissertation, MIT, Cambridge, Mass.

Kayne, R. (1989), "Null Subjects and Clitic Climbing", in O. Jaeggli and K. Safir (eds.), The Null Subject Parameter, Kluwer, Dordrecht

Lema, J. and M.L. Rivero (1990), "Long Head Movement: ECP vs. HMC", in Proceedings of NELS 20, GLSA, Amherst, Mass.

Longobardi, G. (1991), "Extraction from NP and the Proper Notion of Head Government", in A. Giorgi and G. Longobardi, The Syntax of Noun Phrases, Cambridge University Press, Cambridge

Manzini, M.R. (1988), "Phrase Structure and Locality", in A. Cardinaletti, G. Cinque and G. Giusti (eds.), Constituent Structure, Foris, Dordrecht

Manzini, M.R. (1990), "A New Formalization of Locality Theory for Movement", in J. Mascaro and M. Nespor (eds.), Grammar in Progress, Foris, Dordrecht

Manzini, M.R. (to appear), Locality, MIT Press, Cambridge, Mass.

May, R. (1985), Logical Form, MIT Press, Cambridge, Mass.

Obenauer, H. (1976), Etudes de Syntaxe Interrogative du Francais, Niemeyer, Tubingen

Obenauer, H. (1984), "On the Identification of Empty Categories", The Linguistic Review 4.2

Pesetsky, D. (1987), in E. Reuland and A. ter Meulen (eds.), The Representation of (In)definiteness, MIT Press, Cambridge, Mass.

Pollock J.-Y. (1989), "Verb Movement, UG and the Structure of IP", LI 20.3

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

Roberts, I. (1991), "Head-Government and the Local Nature of Head-Movement", GLOW Newsletter 26

Roberts, I. (to appear), Verbs and Diachronic Syntax, Kluwer, Dordrecht

Ross, J.R. (1984), "Inner Islands", Berkeley Linguistic Society 10

Szabolcsi, A. (1991), "Islands, Monotonicity, Composition and Heads", GLOW Newsletter 26

Szabolcsi, A. and F. Zwarts (1990), "Semantic Properties of Composed Functions and the Distribution of Wh-Phrases", in Proceedings of the 7th Amsterdam Colloquium, ITLI