1. Word Formation in Generative Grammar
   Mark Aronoff
2. X Syntax: A Study of Phrase Structure
   Ray Jackendoff
3. Recent Transformational Studies in European Languages
   Samuel Jay Keyser, editor
4. Studies in Abstract Phonology
   Edmund Gussmann
   Susan Steele
6. Some Concepts and Consequences of the Theory of Government and Binding
   Noam Chomsky
7. The Syntax of Words
   Elisabeth Selkirk
8. Syllable Structure and Stress in Spanish: A Nonlinear Analysis
   James W. Harris
9. CV Phonology: A Generative Theory of the Syllable
   George N. Clements and Samuel Jay Keyser
10. On the Nature of Grammatical Relations
    Alec Marantz
11. A Grammar of Anaphora
    Joseph Aoun
12. Logical Form: Its Structure and Derivation
    Robert May
13. Barriers
    Noam Chomsky
14. On the Definition of Word
    Anna-Maria Di Sciullo and Edwin Williams
15. Japanese Tone Structure
    Janet Pierrehumbert and Mary Beckman
16. Relativized Minimality
    Luigi Rizzi
17. Types of Á-Dependencies
    Guglielmo Cinque

The MIT Press
Cambridge, Massachusetts
London, England
Chapter 1
Long and Successive Cyclic
Wh-Movement

1.1 Introduction

It is generally agreed that two different types of Wh-Movement must be recognized: long and successive cyclic.¹ Certain elements, prototypically adjuncts, must undergo successive cyclic Wh-Movement. Others, prototypically complements of verbs, have an additional option: long Wh-Movement.

Pretheoretically, the two types of movement differ in that successive cyclic movement is subject to both strong and weak islands, whereas long movement is subject only to strong islands. See (1)–(3), which are cases of strong islands, and (4)–(7), which are cases of weak islands (the (a) cases represent long movement; the (b) cases successive cyclic movement):²

(1) Subject island
  a. *Which books did [talking about t] become difficult?
  b. *How would [to behave t] be inappropriate?

(2) Complex NP island
  a. *To whom have you found someone who would speak t?
  b. *How have you found someone who would fix it t?

(3) Adjunct island
  a. *To whom did you leave without speaking t?
  b. *How was he fired after behaving t?

(4) Wh-Island
  a. **To whom didn’t they know when to give their present t?
  b. *How did they ask you who behaved t?

(5) Inner (negative) island
  a. To whom didn’t you speak t?
  b. *How didn’t you behave t?
Chapter 1

(6) Factive island
   a. To whom do you regret that you could not speak it?
   b. *How do you regret that you behaved it?

(7) Extrapolation island
   a. To whom is it time to speak it?
   b. *How is it time to behave it?

Beyond the necessity of distinguishing between long and successive cyclic Wh-Movement, no general agreement exists in the literature on many fundamental questions concerning this domain, in particular on the questions in (8):

(8) a. What classes of elements undergo long and successive cyclic Wh-Movement?
   b. From what principles of the theory does the existence of long and successive cyclic Wh-Movement follow?
   c. What is the nature of the locality conditions on long and successive cyclic Wh-Movement?

I will begin with questions (8a) and (8b). After examining the answers given by Chomsky (1986b) and Rizzi (1990), who elaborate on much important work of the 1980s, I will discuss some new facts bearing on question (8a), which appear to be more easily integrated into Rizzi’s system than into the Barriers system, which appears to suggest a particular refinement of the system proposed by Rizzi whereby the referential status of the trace is a prerequisite for long Wh-Movement.

In the second part of the chapter I will reconsider the locality conditions on long and successive cyclic Wh-Movement (question (8c). Taking Chomsky’s (1986b) proposals as a starting point, I will reformulate them in such a way as to eliminate certain redundancies and capture the relevant generalizations in a maximally simple way.

1.2 Chomsky’s (1986b) Analysis of Long and Successive Cyclic Wh-Movement

Chomsky (1986b), building on work by Huang (1982) and by Lasnik and Saito (1984), suggests that the answer to (8a) and (8b) is provided by (a particular formulation of) the ECP. Consider (9)–(13):

(9) Proper government
   \( \alpha \) properly governs \( \beta \) iff \( \alpha \) \( \theta \)-governs or antecedent-governs \( \beta \). (p. 17)

(10) \( \gamma \)-marking
   If \( \beta \) is properly governed, it is assigned \( [+\gamma] \) indelibly. If \( \beta \) is not properly governed, it is assigned \( [-\gamma] \) indelibly. (pp. 17–18)

(11) Such \( \gamma \)-marking takes place at S-Structure for A-positions, and at LF for \( \bar{A} \)-positions. (p. 18)

(12) Empty categories (ECs) not required by the Extended Projection Principle can delete between S-Structure and LF (possibly after having properly governed and \( \gamma \)-marked other ECs). (p. 21)

(13) \( \gamma \)-checking (applying at LF)
   \( *[-\gamma] \)

This formulation of the ECP has, among other consequences, that of forcing a strict successive cyclic derivation for all categories that (like adjuncts) are in \( \bar{A} \)-positions, and of permitting long Wh-Movement of all categories that are in A-positions (pace the Subjacency Condition). Consider briefly how. (14a) is a case of extraction of a \( \theta \)-governed category; (14b) a case of a \( \bar{\theta} \)-governed category in A-position; and (14c–d) cases of \( \bar{\theta} \)-governed categories in A-position:

(14) a. ?Which particular problem were you wondering how to it [phrase it]?
   b. ?Which student did he wonder whether to it [consider it intelligent]?
   c. How have you it [decided it to it [phrase the problem it]]?[
   d. *How are you it [wondering which problem to it [phrase it]]?

Let us start with (14c). The trace there is not \( \bar{\theta} \)-governed. Hence, to be properly governed (to be assigned \( [+\gamma] \)), it must be antecedent-governed. It is if how moves successive cyclically, adjoining first to the embedded VP. Given Chomsky’s (1986b) definition of government, \( t \) antecedent-governs \( t \). \( t \) cannot delete between S-Structure and LF because, for ECs in \( \bar{A} \)-position like the trace of how, \( \gamma \)-marking applies only at LF. If it deleted, \( t \) would not be \( \bar{\gamma} \)-marked. This also means that the trace left adjointed to the embedded VP must itself satisfy the ECP (that is, be antecedent-governed), as must every higher EC needed to antecedent-govern a lower EC.

Clearly, none of the ECs represented in (14c) can be missing (whence the strict successive cyclic movement of how), nor can any of them delete prior to LF. The reason why (14d) is ill formed within this system is now transparent. It violates the ECP. If not \( t \), then \( t \) fails to be properly governed (antecedent-governed). As the Spec of the embedded CP is filled
Chapter 1

by another $wh$-phrase, the closest antecedent is $t'$. But this phrase does not antecedent-govern $t'$. A barrier (by inheritance), CP, intervenes between them.

Consider now (14a) and (14b). In (14b) the trace is not $\theta$-governed. Hence, it must be antecedent-governed. This forces the phrase to adjoin to the higher VP, from which position it can antecedent-govern the original trace and assign it $[+y]$ at S-Structure since the latter is in an A-position. But once it has $\gamma$-marked the original trace at S-Structure, the VP-adjointed trace can delete, so that no other intermediate EC will be needed to antecedent-govern it. As a result, the phrase will be free to undergo long $wh$-Movement from the VP-adjointed position (again, pace Subjacency).

Finally, in (14a) the trace is $\theta$-governed by the verb, which thus assigns it $[+y]$ at S-Structure. This marking, carried along to LF, by itself satisfies the ECP. No antecedent government is required and long $wh$-Movement is again permitted by the ECP.

1.3 Rizzi’s (1990) Binding and Government Approach

Despite its remarkable success in deriving many important distinctions, the Barriers system raises a number of questions, both conceptual and empirical. Concerning the former, Rizzi (1990) notes the existence of a redundancy between the generalized requirement of head government (see note 4 here) and the $\theta$-government requirement of “proper government.” Every phrase that is $\theta$-governed is, a fortiori, head-governed. Thus, both clauses of the conjunctive formulation of the ECP turn out to require some sort of head government. A second conceptual problem inherent in the formulation of proper government is, as often noted, the disjunction between $\theta$-government and antecedent government. To admit a disjunctive statement of this sort amounts to admitting that the nature of the relevant generalization is not understood.

The empirical problems are inherent in the general prediction that phrases in A-positions should be able to undergo long $wh$-Movement. This does not always seem to be the case.

Consider, for example, measure object NPs and objects of idiomatic VPs. As pointed out by Rizzi (1990) (also see Koopman and Sportiche 1988), these elements fail to undergo long $wh$-Movement even though they are $\theta$-marked by the verb that selects them. See, for example, (15)–(16), which contain measure phrases, and (17)–(18), which contain the VP idioms fare giustizia ‘do justice’ and prestare attenzione ‘pay attention’:

(15) *Quanti chili ti ha chiesto se pesavi? How many kilos has he asked you whether you weighed
(16) *Quanti chilometri non sai how many kilometers don’t you know se Venezia disti da Treviso? whether Venice is far from Treviso
(17) *GIUSTIZIA, mi domando quando faranno finalmente justice I wonder when they will finally do
the attention che non ho ancora deciso a chi prestare è poca. the attention that I haven’t decided yet to whom to pay is little
These NPs can of course be $wh$-Moved, apparently at an unbounded distance (via the successive cyclic option). See (19) and (20):

(19) a. Quanti chili credi che riuscirà a pesare how many kilos do you think that he will be able to weigh
dopo questa dieta? after this diet
b. Quanti chilometri credi che abbia detto how many kilometers do you think that he said
di distava, Venezia? che distava, Venice?
that Venice was far
(20) a. GIUSTIZIA, dice di voler fare! justice he says he wants to do
b. L’attenzione che ho deciso di prestare a Gianni è poca. the attention that I decided to pay Gianni is little

If we must conclude that long $wh$-Movement is not simply a prerogative of phrases in A-positions, of which class of elements is it a prerogative?

On the basis of the contrast between ordinary objects, which can be long $wh$-Moved, and measure or idiomatic NPs, which cannot, Rizzi (1990) suggests that it is the nature of the $\theta$-role involved that matters, over and above the requirement that the target of long $wh$-Movement be in an A-position. He expresses this condition unitarily by requiring that the target of long $wh$-Movement be a phrase receiving a $\theta$-role referring to the participants in the event described by the predicate: agent, theme, goal, and so on, but not measure, manner, or the role assigned to quasi arguments such as idiom chunks. He calls the former referential, and the latter nonreferential, $\theta$-roles.
Concerning the deeper question of why long *Wh*-Movement should be limited to phrases receiving a referential \( \theta \)-role, Rizzi proposes a solution in which the classical notion of referential index plays a crucial role.

The essence of Rizzi's proposal can be summarized as follows:

(21) a. The use of indices should be restricted, as in the classical theory of Chomsky (1965), to express referential dependencies between different arguments.

b. Movement does not create indices, but can only carry a (referential) index that is made legitimate by certain referential properties of the elements bearing it” (that is, is assigned to phrases receiving a referential \( \theta \)-role).

c. The binding relation \( X \text{ binds } Y \) iff (i) \( X \text{ c-commands } Y \) and (ii) \( X \) and \( Y \) have the same index) is defined in terms of the notion of referential index.

These assumptions have the effect of restricting binding relations to phrases bearing a referential \( \theta \)-role.

As Rizzi notes, this restriction subsumes the essential effect of the identification clause of the ECP, capturing the fundamental argument/adjunct asymmetries. The \( \Lambda \)-dependency between an operator phrase that receives a referential index at D-Structure and its trace can be expressed through binding. But the \( \Lambda \)-dependency between an operator phrase that does not receive a referential index at D-Structure and its trace cannot be so expressed. As the operator phrase must still be somehow connected to its trace, the system must resort to other available means.

In (14c), (19a–b) and (20a–b) are well-formed cases of successive cyclic *Wh*-Movement. The *wh*-phrases involved, not receiving a referential \( \theta \)-role (hence, index) at D-Structure, cannot connect to the original trace via binding. But they can connect to it via a chain of antecedent government relations since no government barrier intervenes between any of the pairs of positions involved.

(14d), (15)–(16), and (17)–(18) are ill-formed since the *wh*-phrase can connect to its original trace neither via binding (it does not receive a referential index at D-Structure) nor via (antecedent) government (a barrier by inheritance, the interrogative CP node, intervenes between the embedded VP-adjointed EC and the matrix VP-adjointed EC).9

(14a–b) are well-formed cases of long *Wh*-Movement since here the *wh*-phrase receives a referential \( \theta \)-role (hence a referential index) at D-Structure and can thus connect to its trace via binding, after movement. The fact that a government barrier (or a potential \( \Lambda \)-antecedent) intervenes between them is thus inconsequential.

The fact that *wh*-phrases not receiving a referential index at D-Structure cannot cross the other weak islands either also follows from the (antecedent) government requirement on each link of the successive cyclic derivation. See, for example, (22)–(24). The (a) cases exemplify the behavior of measure phrases, the (b) cases that of objects of VP idioms:

(22) **Negative island**

a. *Quanti chili non pesi?*

   how many kilos don’t you weigh

b. *Quanta attenzione non ti presti mai?*

   how much attention doesn’t she ever pay you

(23) **Factive island**

a. *Quanti chili ti rammarichi che lei pesi?*

   how many kilos do you regret that she weighs

b. *Quanta attenzione ti rammarichi di non avergli mai prestato?*

   how much attention do you regret you never paid him

(24) **Extraposition island**

a. *Quanti chili ha certamente contatto pesare per lei?*

   how many kilos did it certainly matter to weigh for her

b. *Quanta attenzione ha certamente contatto prestargli?*

   how much attention did it certainly matter to pay him

In (23)–(24) a government barrier, the non-L-marked embedded CP, intervenes between two positions of the successive cyclic derivation. In (22) a potential \( \Lambda \)-antecedent (the negation) intervenes (see Rizzi 1990, chap. 1).10

If so, the conceptual problems observed above dissolve. The irreducible disjunction between \( \theta \)-government and antecedent government in the “identification requirement” of the ECP has no primitive theoretical status.
any more, its effects having been subsumed under two independent principles, binding and government. The ECP can now reduce to the “formal licensing requirement” (a nonpronominal EC must be head-governed), thus eliminating the redundancy observed above and simplifying the overall system.

1.4 The Role of Referentiality

Having reviewed two particular partitionings of the classes of elements undergoing long and successive cyclic Wh-Movement, that of Chomsky (1986b) and that of Rizzi (1990), I now turn to some new facts bearing on this question, which appear to suggest the necessity of further restricting the class of elements that undergo long Wh-Movement. Since they point to the linguistic relevance of a particular notion of referentiality, they provide confirmation for the general approach of Rizzi (1990), while also suggesting an important refinement of that system.

To anticipate the main conclusions that will be reached, it appears that, of all the phrases that receive a referential θ-role, in Rizzi’s sense, only those can be long Wh-Moved that are used strictly referentially—in other words, that refer to specific members of a preestablished set. This characterization recalls Pesetsky’s (1987) important notion of D(sicourse)-linking, which I will later subsume under the relevant notion of referentiality used here. (For the significance of somewhat different conception of referentiality in the account of long Wh-Movement, see also Aoun 1986 and Aoun et al. 1987.)

Within Rizzi’s system, this result may be derived if, everything else being held constant, we limit the assignment of referential indices just to (wh)-phrases that are used referentially (that are D-linked). Only these will be able to enter binding relations via the referential index mechanism.

Whenever reference to members of a preestablished referential set is inherently impossible for some phrase, or hard to force, then, even if the phrase receives a referential θ-role, it will not be able to enter a binding relation with its trace, so that no long Wh-Movement will appear to be open to it.

To check the relevance of referentiality for binding relations, we need some independent way to discriminate between referential and non-referential phrases. One phenomenon that discriminates between referential and nonreferential phrases is coreference. Plainly, only referential phrases can enter coreference relations. So, for example, R-expressions, pronominals, and (lexical) anaphors can corefer, but, as is well known,
**Wh-Movement.** This is so because they are not referential, hence do not receive a referential index, and hence cannot enter a binding relation. Since only a successive cyclic derivation will be open to them, we expect that no such NPs will extract from weak islands, which block government relations.

### 1.4.1 Extraction from Weak Islands

As a matter of fact, neither the universal distributive quantifier *ogni* N nor the negative universal quantifiers *nessun* N, *niente* 'nothing', which we saw behave like nonreferential elements in (25b–c) and (27a–b), are extractable from weak islands. ((28) and (29) are examples of topicalization, which in Italian could more appropriately be termed *Focus Movement*.)

(28) a. *OGNI DICHIARAZIONE, mi chiedo perché abbia ritrattato. every statement I wonder why he retracted

b. *OGNI MUSEO, non vuole visitare. every museum he does not want to visit
c. *OGNI DICHIARAZIONE, mi rammarico every statement I regret

that he retracted
d. *OGNI MUSEO, è uno scandalo che chiedano. every museum it is a scandal that they shut

(29) a. *NESSUN LIBRO.\{NIENTE,\}

\{no book\}

\{nothing\}

I wonder why he bought

b. *NESSUN LIBRO.\{NIENTE,\}

\{no book\}

\{nothing\}

it is not true that he bought

c. *NESSUN LIBRO.\{NIENTE,\}

\{no book\}

\{nothing\}

I regret that he bought

d. *NESSUN LIBRO.\{NIENTE,\}

\{no book\}

\{nothing\}

it would be unacceptable

Within the present analysis, this implies that the quantifiers do not receive a referential index at D-Structure as a consequence of their nonreferential nature.

*Ogni* N and the similar *qualsunque* N, *qualsiasi* N 'whatever N' and *chiunque* 'whoever' contrast with the nondistributive universal quantifier *tutti* NP, which apparently can undergo long Wh-Movement (30) and consistently has an interpretation (the "collective" one) in which it is subject neither to the pronoun binding requirement (31) nor to weak crossover (32). (See also Reinhart 1983 for relevant discussion.)

(30) a. TUTTI I MUSEI, mi chiedo chi possa aver visitato.

all the museums I wonder who can have visited

b. TUTTI I MUSEI, non ha visitato. all the museums he has not visited

c. TUTTI I MUSEI, mi rammarico che abbiano fatto chiudere. all the museums I regret that they shut

d. TUTTI I MUSEI, sarebbe necessario che chiudessero. all the museums it would be necessary that they shut

(31) [Quelli di loro che hanno visitato tutti i musei] li hanno trovati those of them who visited all the museums they found

uno più interessante dell'altro one more interesting than the other

(32) [Le loro affermazioni incaute] hanno finito per rovinare their incautious statements ended up ruining

tutti i miei amici. all my friends

### 1.4.2 Longobardi's Scope Reconstruction Facts

Longobardi (1987b) makes the important observation that Reconstruction of the scope of an extracted quantifier is blocked not only by strong islands (Longobardi 1986) but also by weak islands. Normally, the scope properties that a quantifier has on the basis of its D-Structure position are preserved when it is moved to an A-position (see Van Riemsdijk and Williams 1981, Cinque 1982, Haik 1984, and references cited there), though new properties may arise as a consequence of its S-Structure position (see Chomsky 1980b). For example, in (33a), though not in (33b), the quantifier phrase *quantì pacienti* 'how many patients' can be in the scope of the universal distributive quantifier *ognuno dei medici* 'every one of the doc-
itors’. That is, (33a) can be satisfied by a family of answers (“I think that Dr. Rossi can visit 5 in one hour, Dr. Bianchi 7, and so on”), whereas (33b) can be answered by just one number (“Only 7”).

(33) a. Quanti pazienti pensi che ognuno dei medici riesca a visitare t_i in un’ora?
   can visit in one hour
   b. Quanti pazienti pensano che ognuno dei medici riesca a visitarli in un’ora?
   can visit them in one hour

As Longobardi noted, this scope Reconstruction is apparently blocked when the quantifier in question is extracted from a weak island, which otherwise normally allows extraction (and Reconstruction) of complements:

(34) a. ?Quanti pazienti ti chiedevi come ognuno dei medici riesca a visitare in un’ora?
   how many patients did you wonder how every one of the doctors could visit in one hour
   b. ?Quanti pazienti non pensi che ognuno dei medici riesca a visitare in un’ora?
   how many patients don’t you think that every one of the doctors can visit in one hour
   c. ?Quanti pazienti ti lamenti che ognuno dei medici sia riuscito a visitare in un’ora?
   how many patients do you regret that every one of the doctors managed to visit in one hour
   d. ?Quanti pazienti sarebbe uno scandalo che ognuno dei medici visitasse in un’ora?
   how many patients would it be a scandal that every one of the doctors visited in one hour

For each question in (34), the family of answers that was possible in (33a) is excluded. (34a–d) are (in fact, only marginally) possible with an interpretation roughly paraphrasable as ‘How many patients are such that you wondered how every one of the doctors could visit them in one hour?’ (see Longobardi 1987b), in which the quantified phrase acquires a referential reading. (Below we will see independent evidence to this effect.)

Longobardi (1987b) interprets this fact as suggesting that scope Reconstruction is only possible where the operator can be connected to its variable through a chain of antecedent government relations—hence, not across the boundary of even a weak island. This important insight still raises a question: why should scope Reconstruction require such a chain whereas Reconstruction of other properties does not? See (35a–c) and (36), which exemplify Reconstruction of Principles A, B, and C of the binding theory and pronominal binding, respectively, across a wh-island:

(35) a. E’ a se stessa che non so se lei abbia scritto.
   it’s to herself that I don’t know whether she has written
   b. *E’ a lei che non so se Maria abbia scritto.
   it’s to her that I don’t know whether Maria has written
   c. *E’ a Maria che non so se lei abbia scritto.
   it’s to Mary that I don’t know whether she has written

(36) A chi lo ha aiutato, non so to those who helped him I don’t know se ognuno di loro restituirà il favore.
   whether every one of them will reciprocate

Given the preceding discussion, it appears that this curious property of scope Reconstruction need not be stipulated. In other words, no special condition need be imposed on this type of Reconstruction. Its properties as exemplified in (34) can rather be seen as simple effects of the nonreferential nature of the extracted quantifier. When it interacts with another quantifier, quanti N must be interpreted nonreferentially. As a result, the wh-phrase is able to connect to its trace only via a chain of antecedent government relations, not via binding (whence the character of (34) noted by Longobardi). On the other hand, when quanti N does not interact with another quantifier, it (marginally) admits a referential reading—hence the marginal possibility of extracting it from a weak island. Compare (34) with (37):

(37) (?)Quanti pazienti non ricordi se lui avesse visitato?
   how many patients don’t you remember whether he had visited

It is apparently possible to check the correctness of this analysis by checking the twofold prediction that it makes.

Whenever quanti N interacts with another quantifier and must therefore be interpreted nonreferentially, no coindexation should be possible between it and a pronoun outside its c-command domain, since both coreference and pronominal binding are unavailable. Conversely, when it does not interact with another quantifier, such coindexation should be marginally possible, since reference (hence, coreference) is marginally available to it, as noted.
Chapter 1

This is precisely what we find. Compare (38) with (39):

(38) *[Quantipazienti ognuno de medi potesse visitare]
    how many patients every one of the doctors could visit
    non era chiaro neppure a loro.
    was not clear even to them

(39) *[Quantipazienti occorressero non era chiaro neppure a loro.
    how many patients were necessary was not clear even to them

Similar considerations apply to such quantifier phrases as tantii N...quantii N 'as many N...as', costi tantii N...che 'so many N...that', and the like.

1.4.3 Clitic Left Dislocated Bare Quantifiers

The evidence discussed in this section presupposes an analysis of the Romance construction of clitic left dislocation (CLLD) that will be justified extensively in chapter 2. Here I will simply state the relevant conclusions reached there.

CLLD, as opposed to topicalization, does not involve (movement of) an empty operator. This entails, among other consequences, the following contrast:

(40a) a. GIANNI (*lo) ho visto.
    Gianni (locus) (him) I saw
    Gianni I saw him

A "resumptive" clitic is impossible with a topicalized object but is obligatory with a CLLD object. The contrast follows under Chomsky's (1977) analysis of topicalization and the above assumption concerning CLLD. (40a-b) receive the following analyses:

(41a) [tor GIANNI] [_*Q] [lo (*lo) ho visto e]]
    [tor Gianni] [cr*lo (*lo) ho visto e]

(41a), with a clitic locally binding the object EC, violates the principle barring vacuous quantification and is thus parallel to (42):

(42) Chi (*lo) hai visto?
    whom have you seen him

(41b), without the clitic, is not a well-formed structure, since the empty object qualifies as none of the various types of empty NPs: it can be neither an anaphor, nor pro, nor PRO, nor a variable, since no empty operator is permitted in CLLD.

Interestingly, if the object phrase in Top (an A-position) is a bare quantifier ([Q], qualcosa 'something', qualcuno 'someone', etc.), though not if it is a quantified NP (qualche N/alcuni N 'some N', molti N 'many N', etc.), the resumptive pronoun may be missing:

(43a) a. Qualcuno, (lo) troveremo.
    someone we (him) will find
    b. Qualcosa, di sicuro, lo (la) farò.
    something surely (it) I'll do

(44a) a. Qualche errore, Carlo *lo ha fatto.
    some error Carlo (it) has made
    some books (them) I bought
    c. Molte lettere, lui *le butta via.
    many letters he (them) throws away

The presence or absence of the resumptive clitic in (43) is not simply optional. It correlates with a difference in the referential properties of the quantifier. If the speaker has something or someone specific in mind (that is, if the bare quantifier is used referentially), the clitic is required. If the interpretation is 'something or other' or 'someone unspecified', the clitic is impossible.

This suggests that bare quantifiers used nonreferentially behave like intrinsic operators, which can identify an EC as a variable at S-Structure, whereas bare quantifiers used referentially and quantified NPs cannot, so that a resumptive clitic is required (see Cinque 1986, Dobrovie-Sorin 1987, 1990).\footnote{15}

Being nonreferential when they identify an EC as a variable, left-dislocated bare quantifiers should thus only be able to connect to the associated EC via a chain of antecedent government links and should accordingly be sensitive to weak islands. This is precisely the case.\footnote{16}

(45a) a. *Qualcuno mi chiedo come troverai.
    someone I wonder how you'll find
    b. *Qualcosa mi chiedo che farà per noi.
    something I wonder who will do for us

(46a) a. *Qualcuno, non credo che troverà.
    someone I don't think that he will find
    b. *Qualcosa, di sicuro, non farò.
    something surely I won't do

(47a) a. *Qualcuno, mi pento di aver aiutato.
    someone I regret I helped
b. *Qualcosa, mi pento di aver fatto per loro.
something I regret I did for them

(48) a. *Qualcuno, è un vero scandalo che abbia schiaffeggiato.
someone it’s a true scandal that he slapped in the face
b. *Qualcosa, è un vero scandalo che abbia ottenuto.
something it’s a true scandal that he obtained

1.4.4 Pesetsky’s Notion of D-Linking and Referentiality
As mentioned, the notion of “referentiality” (as the ability to refer to specific members of a set in the mind of the speaker or presupposed in discourse) recalls Pesetsky’s (1987) important notion of D-linking. I think, in fact, that the two are one and the same notion, or perhaps, more accurately, that the notion of referentiality subsumes that of D-linking.17

Pesetsky interprets certain asymmetries existing between two types of wh-phrases in situ as being due to their differing ability to relate to referential sets preestablished in discourse. Which N phrases can (in fact, must) refer to members of a set that both speaker and hearer have in mind (as do the relevant answers).

Bare wh-operators like who, what or how many N phrases can do so only quite marginally, and under very special contextual conditions forcing some linking to previous discourse. The simple addition of a phrase like the hell or on earth, which expresses surprise or ignorance of the possible answers and is thus incompatible with the choice among elements isolated in the previous discourse, suffices to exclude any such discourse linking.

Pesetsky calls the former D-linked and the latter non-D-linked wh-phrases. By assuming (1) that operators must occupy an A-position at LF and (2) that only non-D-linked wh-phrases are operators, he is able to account for a number of asymmetries between the two types of wh-in-situ. Only non-D-linked wh-phrases in situ (such as who, what, how many N phrases) are expected to move at LF and should thus show the usual diagnostics of movement (sensitivity to Subjacency, the ECP, and other conditions on movement like the Nested Dependency Constraint). D-linked wh-phrases in situ (such as which N) are instead expected not to move at LF, since they do not qualify as operators. They are rather interpreted in situ via a different mechanism, that of unselective binding (see Pesetsky 1987 for more careful discussion).

The relative contrasts between (49) and (50) and between (51) and (52) are thus expected:

(49) a. *Mary asked what who read.

Long and Successive Cyclic Wh-Movement

b. *I need to know whom how many people voted for.

(50) Mary asked what which man read.

(51) ??Tell me what proves that who is innocent.

(52) Tell me which piece of evidence proves that which person is innocent.

Pesetsky’s conclusions about the behavior of wh-phrases at LF are consistent with the conclusions reached in the previous sections about the (movement of) quantifier phrases at S-Structure. Only D-linked (in our terms, referential) phrases can indeed enter a binding relation, whether at S-Structure or at LF. Non-D-linked (nonreferential) phrases are instead forced to enter only chains of antecedent government relations, both at S-Structure and at LF.

It is thus not unexpected that non-D-linked wh-phrases do not take scope over a wh-island, even in languages (such as French) that allow for extractions out of wh-islands, since such LF movement is sensitive to antecedent government relations, just as overt syntactic movement of non-D-linked phrases is (see Huang 1982, Lasnik and Saito 1984):

(53) Qui sait quand Jean en a acheté combien?
who knows when Jean of them has bought how many

We will in fact see in the next section that there is complete parallelism between wh-extraction from a wh-island at LF and the same extraction in the syntax:

(54) *Combien te rappelles-tu quand Jean en a acheté?
how many do you remember when Jean of them has bought

That D-linking, in Pesetsky’s sense, implies referentiality is confirmed by the fact that a D-linked wh-in-situ (but crucially no non-D-linked wh-in-situ) can enter coreference relations. See the contrast between (55a) and (55b):18

(55) a. Which boy, started a fight with which girl, wasn’t clear even to them.

b. *Who, started a fight with whom, wasn’t clear even to them.

1.4.5 (Non-)D-Linked Wh-phrases and the Wh-Island Constraint
As noted independently by various authors, extraction of interrogative phrases out of indirect questions depends on the character of the extracted wh-phras. Rizzi (1982, chap. 2, fn. 5) and Engdahl (1980, a, b) observe that only relatively “heavy” wh-phrases can be extracted from wh-islands in Italian and Swedish, respectively. But, to judge from the examples, the
relevant feature seems once again to be the D-linked versus non-D-linked character of the extracted \textit{wh}-phrase. Only the former can apparently be extracted ((56a) is in fact marginally acceptable if a pause follows \textit{a chi}, and otherwise impossible):

(56) a. \(`A chi ti chiedi quanti soldi hai dato?\)  
   whom do you wonder how much money you gave  
   A quale dei tuoi figli ti chiedi  
   to which one of your children do you wonder  
   quanti soldi hai dato?  
   how much money you gave

b. \(`Vad visste ingen vem som skrev?\)  
   what knows no one who wrote  
   Sven undrar vilken bok alla studenter minns  
   Sven wonders which book all students remember  
   vilken författare som skrev.  
   who which author wrote

See Comorovski 1985, 1989, from which (57a–b) are drawn, and Bedzyk 1987 for the explicit claim that only D-linked \textit{wh}-phrases can be extracted from \textit{wh}-islands in Romanian and Bulgarian, respectively. As Comorovski (1989) indeed remarks, "The cross-linguistic generalization that emerges is that only D-linked \textit{wh}-phrases can be questioned out of indirect questions." She proposes a solution for this contrast based on certain presupposition properties of D- and non-D-linked \textit{wh}-phrases in interaction with the semantics of \textit{wh}-islands. Within the present analysis of long \textit{Wh}-Movement, however, nothing at all is needed to account for the asymmetry in question.

Only D-linked (hence, referential) \textit{wh}-phrases can enter a binding relation with their trace, whence their insensitivity to \textit{wh}-islands (weak islands, more generally). Non-D-linked (nonreferential) \textit{wh}-phrases instead can only move successive cyclically, whence their sensitivity to weak islands.

The other often noted asymmetry between questioning (impossible or marginal) and relativization (marginal or perfect) in extraction from \textit{wh}-islands\textsuperscript{19} appears to be amenable to the same account. The relative \textit{wh}-phrase can plausibly be referential more easily than the interrogative \textit{wh}-phrase, which behaves more typically like a nonreferential operator.\textsuperscript{20}

1.4.6 \textbf{A-Chains and Government and Binding Chains}  
The evidence discussed in the previous sections thus points to the existence of a further condition on binding. For a phrase to enter a binding relation with its trace, it must not only occupy an A-position and receive a referential \textit{\theta}-role. \textit{It must also have intrinsic referential properties}—properties that allow it, for example, to be D-linked, in Pesetsky’s (1987) sense, partially a matter of lexical variation (\textit{which N versus who (the hell)}, and so on).

We have seen that such a requirement can be built in Rizzi’s (1990) system by requiring that the principle licensing referential indices be made sensitive to the intrinsic referential nature of the phrase. This additional condition on binding allows us to reconsider Rizzi’s (1990) account for why NP-Movement can only enter a chain of government relations (Chomsky 1986b, sect. 11). Since NPs in NP-Movement processes (Passive, Raising, and so on), as well as in Clitic Movement, are clearly in A-positions receiving a referential \textit{\theta}-role at D-Structure, Rizzi (1990, chap. 3) suggests that their obligatorily entering a government-type rather than a binding-type relation is forced by an independent requirement: namely, that \textit{\theta}-role and Case transmission, crucially involved in such processes, is a property of chains, and that chains can be taken to be defined in terms of antecedent government. As no comparable \textit{\theta}-role and Case transmission are involved in A-relations, no requirement of antecedent government is then forced with them.

However, the suggested referentiality requirement for binding opens up a different way of interpreting the necessary antecedent government requirement on A-chains (and clitic chains). It seems plausible to say that traces of NP-Movement and Clitic Movement, unlike variables, are not referentially autonomous, in that they are only subparts of a discontinuous referential element: the A-chain (see Cinque 1989, Browning 1989a). If so, they will not be able to enter a binding relation, provided that we impose on binding the requirement that the \textit{trace} be intrinsically referential (this being, in part, a function of the nature of its antecedent, as noted).

One desirable consequence of this analysis is that we may then keep the notion of chain for both government and binding relations—a move that is in fact necessary if Reconstruction is a property of chains (see Cinque 1982), given that binding as well as government relations display Reconstruction effects. See, for example, (35b), repeated here as (58), in which the pronominal violates Principle B of the binding theory under Reconstruction (however Reconstruction is formally derived; see Cinque 1982, Bars 1986, 1988, and references cited there):

(58) `E’ \textit{a lei} che non so \textit{se Maria abbia scritto.}  
   It’s to \textit{her} that I don’t know whether \textit{Maria} has written
Chains (and the concomitant property of Reconstruction) seem to find their motivation on other grounds: whenever a phrase in an A- or A'-position is not licensed independently of another A- or A'-position, it must enter a chain with it (again see Cinque 1982, for discussion along these lines). Whether it enters a binding or a government chain (where, in the latter, each link must satisfy antecedent government) will depend on the referential or nonreferential nature of the trace, which depends in turn on the referential or nonreferential nature of the operator, as noted.

Related to this is the question whether referential traces (in the above sense) can enter both a binding chain and the stricter government chain (in other, more traditional terms, whether elements participating in long Wh-Movement can also participate in successive cyclic Wh-Movement).

The implicit assumption is that they can. But one may conceive of the possibility that the two modes of connecting a phrase in an A'-position and its trace could be mutually exclusive, so that nonreferential phrases will enter only a government chain and referential elements only a binding chain. This possibility may be desirable on other grounds (such as parsing), as pointed out to me by Maria Teresa Guasti. (The experimental results of De Vincenzi (1989) in fact appear to lend some credence to this idea.)

This points, then, to the possible existence of two operator/variable configurations, differing “semantically” in terms of referentiality, and syntactically in terms of the kind of chain that (as a consequence of their “semantics”) they are able to enter: binding chains or government chains.

Let us now consider the locality conditions on long and successive cyclic Wh-Movement, now identified with binding chains and (antecedent) government chains, respectively.

1.5 Chomsky’s (1986b) Locality Conditions on Long and Successive Cyclic Wh-Movement

As noted, long Wh-Movement is subject to strong islands, whereas successive cyclic Wh-Movement is subject to both strong and weak islands.

Chomsky (1986b) proposes that the sensitivity to strong islands (as well as weak islands) manifested by successive cyclic Wh-Movement follows from the requirement that each link of the successive cyclic chain satisfy (antecedent) government. So, for example, in the subject, complex NP, and adjunct island cases seen above, there is a maximal projection (the subject and relative CP in the first two, and the adjunct CP or PP in the latter) that fails to be L-marked, hence functioning as a blocking category and a barrier to (antecedent) government. The sensitivity to strong islands manifested by long Wh-Movement follows instead from a separate principle: Subjacency.

There are thus two distinct ways to derive strong island effects, one based on government and the other on Subjacency. This amounts to building a certain redundancy into the system, though perhaps an unavoidable one, if it were the case that Subjacency effects are overridden, for successive cyclic Wh-Movement, by the more restrictive condition that also derives weak islands.

Chomsky (1986b) attempts to reduce this redundancy by employing a common notion of barrier in the formulation of (antecedent) government and Subjacency.  

\[(59) \gamma \text{ is a barrier for } \beta \text{ iff } (a) \text{ or } (b):\]

\[a. \gamma \text{ is a maximal projection which immediately dominates } \delta, \delta \text{ a non-L-marked maximal projection dominating } \beta. \text{ (definition of barrier “by inheritance”)}\]

\[b. \gamma \text{ is a non-L-marked maximal projection dominating } \beta, \gamma \neq \text{IP.} \text{ (definition of “inherent” barrier)}\]

These definitions give some form to the intuitive idea of Cattell (1976) and others that every maximal projection that is not the complement of a lexical category is a barrier (except IP). And they add that even a maximal projection that is a complement of a lexical category becomes a barrier for some element if it immediately dominates another barrier (or IP) dominating that element. We will come back later to the apparent special character of IP in Chomsky’s (1986b) definition of barrier.

The definition of barrier (59) takes part in both the definition of (antecedent) government (60) and that of Subjacency (61).

\[(60) \alpha \text{ governs } \beta \text{ iff } \alpha \text{ m-commands } \beta \text{ and there is no } \gamma, \gamma \text{ a barrier for } \beta, \text{ such that } \gamma \text{ excludes } \alpha. \text{ (}= (18) \text{ of Chomsky 1986b, 9})\]

\[(61) \beta \text{ is subjacent to } \alpha \text{ iff there are fewer than two barriers for } \beta \text{ that exclude } \alpha. \text{ (see (59) of Chomsky 1986b, 30 and corresponding text)}\]

By utilizing the same notion of barrier for the theories of government and bounding (Subjacency), Chomsky thus aims at a unified approach to these theories and at the elimination of the redundancy noted above in the derivation of strong islands.

The unification (and the concomitant elimination of the redundancy) is only partial, however. First, because one barrier suffices to derive strong islands for government, whereas two are needed to derive them for Subjacency—clearly a residue of asymmetry. Second, because two additional
notions of barrier are introduced, one holding exclusively of government (the “minimality” barrier), the other exclusively of Subjacency (the weak barrier constituted by the most embedded IP, or CP, designed specially for wh-islands)—another instance of asymmetry. Third, because the notion of barrier holding of government and Subjacency can be given a unified definition only at a certain cost. As we will see, the notion of barrier by inheritance can be dispensed with for government but is crucial for Subjacency, so that the retention of a unified notion ultimately entails a redundancy (for government).

It would be desirable to eliminate these asymmetries and redundancies between the theories of government and bounding, or at least to minimize them, if we were to discover an irreducible residue of difference. I attempt this in the next section. After eliminating some such asymmetries and redundancies, I arrive in fact at the conclusion that a residue of difference exists and that it may be convenient to expose it in a perspicuous way. In section 1.7 I formulate this residue of difference in the form of two (minimally) different definitions of barrier: one for government chains and the other for binding chains, in the sense of section 1.4.

1.6 Simplifying Chomsky’s (1986b) Locality Conditions

1.6.1 Eliminating the Notion of Minimality Barrier

Let us begin with the notion of minimality barrier, which holds of government alone (Chomsky 1986b, 42). If we could eliminate it entirely, the theories of government and bounding would come one step closer together. This indeed appears feasible.

Chomsky (1986b) defines the notion of minimality barrier as follows (see (92) and corresponding text):

(62) In the configuration \ldots \ldots [y \ldots \delta \ldots \beta], y is a barrier for \beta if y is the immediate projection of \delta, a zero-level category distinct from \beta.24

In addition to its role in barring government by a head into the domain of another head, the minimality barrier has two further consequences in Chomsky's analysis. The first is to block the extraction of an adjunct from a complex NP of the noun complement type, independently of choice of inherent barriers. For example, in (63) N qualifies as a minimality barrier, thus blocking antecedent government of t" by t"', even if CP and NP were not to qualify as barriers:

(63) *How did John [vP t" announce [NP a [vP t" to [vP t' fix the car t]]]]?}

Long and Successive Cyclic Wh-Movement

The second consequence of the minimality barrier is that it seems to provide an analysis of that-trace effects:

(64) *Who did you believe [CP t' [C that [IP t would win]]]?

In (64) C qualifies as a minimality barrier (when C is lexical), thus blocking antecedent government of t by t'.

It appears, however, that the latter two results of the notion of minimality barrier follow from independent principles. Concerning the complex NP case (63), there is evidence that an inherent barrier is indeed present. As Chomsky notes elsewhere (1986b, 36), that the CP complement of N may be a barrier is suggested, if Stowell's (1981) theory is correct, by the fact that the complementizer of CP cannot be missing (John expressed the feeling *(that) the meeting should not be held). This conclusion is confirmed by the impossibility of genuine complement extraction from such a CP. Chomsky (1986b, 35) contrasts the ill-formed (63) with the almost unexceptionable (65):

(65) (?) Which book did John announce a plan (for you) to read? But (65) is not really representative, because it involves an NP gap, which, if we are right, can be related to an antecedent via a different strategy (see chapter 3). Indeed, if we select a PP complement, which has no access to such a strategy, the result is clearly ungrammatical:

(66) *With whom did John announce a plan to go out?

This points to the barrierhood of CP, and that of NP, by inheritance. As to why the complement CP of a noun should count as a barrier, Chomsky (1986b, 36 and sec. 11) suggests that nouns may not be L-markers, and Grimshaw (1990) gives ample evidence that they are defective -l- (hence, L-)markers.

The second alleged consequence of minimalism, the derivation of that-trace effects, raises certain conceptual problems: Why should an empty C not trigger a minimality violation? Why does that not trigger a minimality violation with simple adjunct extractions How did you say that he fixed the car?]

Besides, an independent way exists to derive that-trace effects that does not incur these problems: the condition of head government, imposed on the trace over and above antecedent government (see Rizzi 1990, chap. 2, Koopman and Sportiche 1988, Longobardi 1985b, forthcoming).25

Thus, adjunct extraction from noun complements and that-trace effects do not provide any independent grounds for a notion of minimality barrier.
This leaves us with the original motivation for minimality: the necessity of preventing a head from governing into the domain of another head. But this can be expressed without recourse to a notion of minimality barrier valid for government alone. For example, if Rizzi's (1990) notion of Relativized Minimality is adopted, this result follows independently.

(67) Relativized Minimality
X α-governs Y only if there is no Z such that
a. Z is a potential α-governor for Y, and
b. Z c-commands Y and does not c-command X
(where α ranges over heads, A-specifiers, or A-specifiers)

The intuitive idea expressed by (67) is that minimality effects are triggered only in the presence of intervening elements of the same type: heads for head government, and A- and A-specifiers for antecedent government in A- and A-chains, respectively. So, government by a head X into the c-domain of another head Z is barred even in the absence of a special notion of minimality barrier, which can thus be eliminated altogether (along with the problems outlined in note 25). See Rizzi 1990 for further discussion of these and related issues.

Now, if it were also possible to eliminate the special notion of barrier valid for Subjacency only (the most embedded tensed IP, or CP, barrier), then we would have a truly unified notion of barrier (of the inherent and inheritance varieties) for both government and bounding.

Let us suppose that this is possible (I will come back to this issue in section 1.9, suggesting that it is). An asymmetry persists, however, despite the higher degree of unification attained: the fact that one barrier is sufficient for government but two are needed for bounding.

Is this really so for all cases? A still higher degree of unification could perhaps be reached, if we could show that even for bounding one barrier suffices (at least, for a proper subset of its cases). We will explore the feasibility of this further simplification starting from a particular redundancy hidden in the Barriers system: namely, the fact that the notion of barrier by inheritance for government is almost always redundant, and, where prima facie it is not, it can apparently be eliminated.

1.6.2 A Redundancy: The Notion of Barrier by Inheritance for Government (and Bounding)
Let us begin with strong islands. The notion of inherent barrier appears to be sufficient to derive them:

Long and Successive Cyclic Wh-Movement

(68) a. Subject island
*Non so come, [IP poterla riparare τ] sarebbe utile.
I don't know how to be able to fix it would be useful

b. Complex NP island
i. *Non so come, giustifichino [NP quelli
I don't know how they justify those
[CP che l'hanno riparato τ] who fixed it

ii. *Non so come, giustifichino [NP i tentativi
I don't know how they justify the attempts
[CP di liberarlo τ]].
to free him

c. Adjunct island
*Non so come, [IP fosse entrato] [CP che piangeva τ].
I don't know how he had gone in while she cried

In all of (68a–c) (and the same is true of the other strong islands), there is an inherent barrier, the non-L-marked CP, which blocks antecedent government quite independently of the barriers by inheritance, IP, NP, and CP, respectively.

The same conclusion holds for some of the weak islands, factive and extraposition CPs. The reason is the same, if—as is generally assumed—the CP is not L-marked (and hence is a barrier). See Kayne 1984 and section 1.7.1.

(69) Factive island
*Non so come, si sia pentito [CP di essersi comportato τ].
I don't know how he repented of having behaved

(70) Extraposition island
*Non so come, vi potrebbe danneggiare
I don't know how it could harm you
[CP che si comportasse τ].
that he behaved

True enough, certain cases exist that would seem to require a notion of barrier by inheritance even for government. For each of them, however, there is an independent way to block government that does not involve such a notion.

The first has to do with government across sentential maximal projections. As shown by Kayne (1984), government must be permitted to cross a CP (as in the case of a V head governing a trace in Spec CP), just as it is permitted to cross IP (as in exceptional Case marking, Raising, and other
Chapter 1

Long and Successive Cyclic Wh-Movement

Three cases falling under Subjacency, where there is but one barrier, are complement extraction from (certain) adjuncts, complement extraction from relative clauses extraposed from an object, and complement extraction from degree clauses (the CP associated with such adjectival modifiers as too and enough). Let us begin with the first case, exemplified in (73) (recall that, to be sure we are dealing with real movement, we must select a non-NP target):

(73) *To whom, did [IP they leave [NP before speaking t]]?

Chomsky (1986b, 31) suggests that this case of Huang’s (1982) Condition on Extraction Domains (CED) follows from (1) Subjacency in that the adverbial PP qualifies as an (inherent) barrier (it is not L-marked) and transmits barrierhood to IP so that two barriers are crossed simultaneously. But this account crucially presupposes that the adjunct inside IP be excluded by VP, as shown in (74a). If it were adjoined to VP, as in (74b), then to whom could itself adjoin to VP, thus voiding the barrierhood of IP and consequently crossing only one barrier (PP).

(74) a. ...[IP they [VP leave] [IP before ... t]]
   b. ...[IP they [VP [VP leave] [IP before ... t]]]

There is in fact evidence from VP-Preposing and do so substitution that the adverbial PP in (73) can be adjoined to VP.27

(75) a. ... and leave before speaking to John, they certainly did.
   b. They left before speaking to John, and I did so too.

Therefore, it must be concluded that to whom has crossed only one (inherent) barrier and that that suffices to trigger a Subjacency violation.28

A comparable conclusion can be reached on the basis of complement extraction from a relative clause extraposed from an object, if, as is standardly assumed, the extraposed CP is in this case adjoined to VP.29 See the ungrammaticality of (76b), derived from (76a) by crossing only one barrier (CP).30

(76) a. Avevo [VP presente[ata qualcuno t] a Gianni
   I had introduced someone to Gianni
   [CP che voleva parlare con sua figlia],
     who wanted to speak with his daughter
   b. *Con chi avevi [VP presente[ata qualcuno t] a Gianni]
      with whom had you introduced someone to Gianni
      [CP che voleva parlare t]?

Degree clauses also qualify as strong islands, as shown by (77a–b), in spite

cases). If so, neither CP nor IP must count as an inherent barrier (when the CP is L-marked). Nonetheless, one must ensure that in (71) the verb does not come to govern the Spec IP:

(71) John decided [CP e [IP PRO to see the movie]].

So it must be that CP counts as a barrier by inheritance, blocking government.26 But this conclusion is not necessary. The notion of Relativized Minimality suffices to block government of a head across CP and IP since another head, C (itself inert for government), intervenes.

A second case, discussed by Chomsky (1986b), that would seem to crucially involve the notion of barrier by inheritance for government is provided by adjunct extraction from a wh-island:

(72) *How did Bill [VP t'] [VP wonder [CP* who [IP VP t']
   [VP wanted [CP t' [IP to fix the car]]]]]

Here, no inherent barrier intervenes between any of the pairs of positions \{t, t', \{t', t'\}, \{t', t''\}, \{t'', how\}. But since extraction is forbidden, there must be a barrier. The barrier by inheritance (from IP), CP*, is apparently the only candidate—whence the conclusion that this notion is crucial for government too. But, once again, that conclusion is not necessary, at least if we adopt Relativized Minimality (which is violated in (72) by the intervention of who between t' and t'').

It thus seems that there is no real need for a notion of barrier by inheritance in the case of government. It is completely redundant, its effects following either from the notion of inherent barrier or from that of Relativized Minimality.

Suppose that, following the fruitful practice of eliminating redundancies from the system, we discard the notion of barrier by inheritance from the theory of government (adopting Relativized Minimality). By doing so, however, we introduce a new asymmetry between government and bounding (Subjacency). The notion of barrier by inheritance now holds of the latter theory only.

Related to this asymmetry appears to be the other noted by Chomsky (1986b) that one barrier suffices for government, whereas two are apparently needed for Subjacency. The barrier by inheritance provides the extra barrier needed for Subjacency. However, there is evidence, to be reviewed directly, that at least for the derivation of strong islands, one barrier (an inherent barrier) is sufficient even for the theory of bounding. This will force us to reconsider the very opportunity of retaining the notion of barrier by inheritance even for bounding.
of the fact that they contain only one (inherent) barrier, the extraposed CP complement of *too and enough (see Chomsky 1986b, 33ff.): 31

(77) a. *To whom were they [CP PRO to talk]?
   b. *To whom were they [CP PRO not to talk]?

The existence of cases such as these, where only one (inherent) barrier suffices to trigger a Subjacency violation, raises the question whether we really need two barriers in the remaining cases falling under Subjacency. In the other strong islands (CP and NP subjects, complex NPs, and so on) an inherent barrier also exists, the non-L-marked CP* and NP* of (78)–(79), which could well suffice to trigger a Subjacency violation, by analogy with the case just considered of adjunct, extraposed relative, and degree clauses:

(78) a. *A chi, crede che [CP parlare t_1] sarebbe vitale?
   b. *Whom, did [NP your interest in t_1] disturb John?

(79) a. *A chi, hai conosciuto [NP qualcuno
to whom have you met someone [CP che voleste parlare t_1]]?
   who wanted to speak
   b. *Con chi, avevete discusso [NP la possibilità
with whom have you discussed the possibility [CP di uscire t_1]]?
of going out

The logic of the matter is clear. Since there are at least three strong islands for which one inherent barrier must suffice to trigger a bounding violation, and since all other strong islands also contain an inherent barrier (in addition to a barrier by inheritance: IP and NP in (78) and (79), respectively), barriers by inheritance come to be entirely redundant for bounding as well, with respect to the class of strong islands.

The second conclusion is that, for strong islands, Subjacency must apparently qualify as 0- rather than 1-Subjacency, in Chomsky’s (1986b) terms.

Let us summarize the discussion so far.

In attempting to attain a higher degree of unification of the theories of government and bounding, we have seen that it is possible to dispose of the government notion of minimality barrier, adopting Relativized Minimality. We have also seen that the notion of barrier by inheritance (and the concomitant redundancy) can be eliminated from the theory of government, and even from the theory of bounding, at least for the class of strong islands.

It would thus appear possible to attain an extreme degree of unification and simplification of the notion of barrier valid for the two theories (any non-L-marked XP). But there is a residue, the weak islands, which constitute a barrier for government, though not for bounding. See the contrasts in (4)–(7), repeated here:

(4) Wh-island
   a. ??To whom didn’t they know when to give their present?
   b. *How did they ask you who behaved?

(5) Inner (negative) island
   a. To whom didn’t you speak?
   b. *How didn’t you behave?

(6) Factive island
   a. To whom do you regret that you could not speak?
   b. *How do you regret that you behaved?

(7) Extrapolation island
   a. To whom is it time to speak?
   b. *How is it time to behave?

If the contrasts in (4) and (5) can be attributed to the independent application of Relativized Minimality and are thus compatible with the unified notion of inherent barrier, those in (6) and (7) cannot. The (inherent) barrier that suffices to block government must not suffice to block binding. The latter class of weak islands thus appears to constitute the ineliminable residue of difference between government and binding/bounding.

The Barriers approach, though keeping a (largely) unified notion of (inherent and inheritance) barrier, expresses this residue of difference in the number of barriers relevant for government (one) and for bounding (two). We have seen, however, that quite apart from the cost inherent in this (in terms of redundancies), the choice of expressing the difference between the two theories in terms of the number of barriers crossed fails in at least some cases, where one barrier must suffice even for bounding.32

This suggests, then, the opportunity of exploring the other logical possibility for expressing the above residue of difference. Instead of taking the notion of barrier to be the same and only the number of barriers to be different for the two theories, one can envisage the possibility that the notion of barrier itself is (partially) different for the two theories, while the number of barriers involved remains constant (one). This is the line that I
will pursue in the next section (see Manzini 1988 for a proposal in the same spirit).

1.7 Locality Conditions on Binding and Government Chains

1.7.1 The Definition of Barrier for Binding (the Theory of Bounding)

What appears to distinguish strong from weak islands is the fact that (with the exception of preverbal sentential subjects in Romance, English, and the like) strong islands are all cases of maximal projections that are neither L-marked nor θ-marked (by a [+V] element), whereas weak islands are all θ-marked maximal projections. This is clearly the case for wh-islands, which are even L-marked by a [+V] element (when complements of a V or A), but it is also the case for such other weak islands as the sentential complements of factive and manner-of-speaking verbs, and the extraposed subject clauses of transitive, unergative, and psych-movement verbs. Let us briefly consider each case in turn.

The CP complements of factive and manner-of-speaking verbs are presumably θ-marked by these verbs as internal arguments, under VP. Some evidence for this is provided by the fact that their infinitival complements in Italian are (obligatorily) introduced by the complementizer di (lit.) of (Deploro|rimpiano|etc. [di non averlo potuto leggere] 'I deplore/regret/etc. not to have been able to read it'): a clear diagnostic, in this language, of internal argumenthood (see Cinque 1990a).

Their weak islandhood suggests that they are not L-marked by the V. If L-marking is defined (as in Chomsky 1986b, 14) as direct θ-marking by a lexical head (where α directly θ-marks β if α and β are sisters, that is, are dominated by the same lexical projections), then factive and manner-of-speaking complements must not be dominated by V, as ordinary direct objects are. In fact, Kayne (1981b, fn. 23) provides evidence that manner-of-speaking verbs must be higher than direct objects, in fact even higher than indirect objects.

The same is presumably true of factive complements. (See Cardinali’s (1989a, chap. 3) argument based on German that they are higher than V and not dominated by NP, contra Kiparsky and Kiparsky 1970). See also Koster 1989 for the conclusion that the preverbal position of factive CPs in Dutch is derived (hence different from that of NPs). The extraposed sentential subject of transitive and unergative verbs is an external argument in VP-adjoined position, an A-position, as argued in much recent work (see, among others, Koopman and Sportiche 1988, Belletti 1988). In other words, it is indirectly θ-marked by the verb, but not L-marked, as required.

Finally, the extraposed sentential subject of unergative psych-verbs of the worry, frighten class is also in VP-adjoined (hence, non-L-marked) position at S-Structure, though it is clearly θ-marked by the verb (in fact, as an internal argument, at D-Structure, if Belletti and Rizzi (1988) are right).35

One could thus conceive of retaining the notion of (inherent) barrier based on L-marking for government relations, and of introducing a new notion of (inherent) barrier for binding relations (long movement), based on the somewhat looser condition of θ-marking (direct or indirect).

This is the alternative way, mentioned above, of expressing the ineliminable difference between government and binding/bounding. Strong islands are barriers for both government and binding, because they are neither L-marked nor θ-marked maximal projections. Weak islands are barriers to government but not binding, because they are not L-marked, but are θ-marked maximal projections.

The apparently exceptional behavior of preverbal sentential subjects, which, though (indirectly) θ-marked, nonetheless qualify as barriers for binding relations (that is, as strong rather than weak islands), can be brought under the aegis of “normality,” if we introduce into the definition of barrier for binding the notion of canonical direction proposed by Kayne (1983) and others.36

(80) Definition of barrier for binding (first approximation)

Every maximal projection that is not θ-marked in the canonical direction is a barrier for binding.

Some evidence that the requirement of canonical direction is a weaker condition additional to, and distinct from, the fundamental requirement of θ-marking is apparently provided by the fact that complement extraction from preverbal sentential subjects gives rise to an ungrammaticality milder than that produced by extracting a complement from the other strong islands (which violate the stricter θ-marking requirement). See the contrast between (81) and (82).

(81) ?Una persona a cui, credo [CP che a person to whom I think that [IP [CP risucire un parlaire t, oggi] sara impossibile]]... to be able to speak today will be impossible

(82) a. *Una persona a cui, sono uscito a person to whom I went out [CP senza riscuire un parlaire t,]... without being able to talk
b. *Una persona a cui, ho incontrato a person to whom I met
\[ [\text{NP}_{\text{CP}} \chi è riuscito a parlare } t_i]\]
who managed to speak

c. *Una persona a cui, sono troppo arrabbiato a person to whom I am too angry
\[ [\text{PP per parlare } t_i]\]
to talk

d. *Una persona a cui, io ho tanti soldi a person to whom I have as much money
\[ [\text{NP quanti Carlo ne ha regalati } t_i \text{ in un giorno}]\]
as Carlo gave in one day

e. *Una persona a cui, erano così arrabbiati a person to whom they were so angry
\[ [\text{CP che non hanno voluto parlare } t_i]\]
that they didn’t want to speak

The contrast can be explained as a consequence of the fact that the stricter \(\theta\)-marking requirement is violated in (82a–e), but not in (81), which merely violates the canonical direction requirement.\(^{37}\) In the present analysis, the special status of preverbal sentential subjects is thus expected to be a feature of binding alone (being due to an at least partial satisfaction of binding requirements). It is not expected to extend to government. And, indeed, no contrast is found between sentential subjects and the other strong islands in adjunct extraction (that is, with government relations). For example:

(83) *Il modo in cui, [\text{CP formulare la richiesta } t_i]\nthe way in which to phrase the demand
sarebbe un errore]\nwould be a mistake

(84) (a) *Il modo in cui, sono uscito [\text{PP senza formulare}\nthe way in which I went out without phrasing
la richiesta } t_i]\nthe demand
(b) *Il modo in cui, conosco [\text{NP}_{\text{CP}} chi ha formulato}\nthe way in which I know who phrased
la richiesta } t_i]\nthe demand

(c) *Il modo in cui, sono troppo arrabbiato [\text{PP per formulare}\nthe way in which I am too angry to phrase
la richiesta } t_i]\nthe demand

This is not surprising, since sentential subjects fail to be \(\text{L}-\)marked, just like any other strong island.

Let us return to the notion of barrier for binding chains, which needs to be further qualified. Though \(\theta\)-marked in the canonical direction, complements of nouns and prepositions (that is, \([-V]\) categories) still appear to qualify as strong islands for binding. This suggests the need to introduce in the definition of barrier for binding the specification that only \([+V]\) categories are capable of lifting the barrierhood of a maximal projection:

(85) Definition of barrier for binding (revised)
Every maximal projection that is not \(\theta\)-marked by a \([+V]\) category in the canonical direction is a barrier for binding.

Let us consider the relevant evidence, beginning with Ns. As (86a–b) show, extraction of (non-NPs, for the noted reasons) from a maximal projection \(\theta\)-marked in the canonical direction by a \([-V, +N]\) category is impossible (the asterisk in (86) suggests, incidentally, that antecedent government, to which we will return, is also impossible in the same context):\(^{39}\)

(86) a. *Gianni, da cui, disapprovo i tentativi
Gianni by whom I disapprove the attempts
\[ [\text{NP di andare a stare } t_i]\]
to stay

b. *Maria, alla quale, hanno respinto l’insinuazione
Maria to whom they rejected the insinuation
\[ [\text{CP di aver raccontato tutto } t_i]\]
of having told everything

Of course, the possibility of (limited) cases of reanalysis should be pointed out. With such expressions as avere l’impressione ‘have the impression’, avere speranza ‘have hopes’, make the claim, express a desire, and the like, even extraction of non-NPs is possible:\(^{40}\)

(87) a. Gianni, al quale, ho l’impressione che non vogliano
Gianni to whom I have the impression that they don’t want
parlare } t_i]\nto speak

b. John, to whom, I made the claim you would never talk } t_i]...
The peculiar behavior of such complex NPs was noted by Ross (1967, 139ff.), who observed one fact that can be taken as evidence that they indeed involve "reanalysis" of some sort. The complementizer that in English can be omitted in the complement of Vs but not in that of Ns.41

(88) a. Kleene proved that this set is recursive.
   b. The proof *(that) this set is recursive is difficult.

The that following the N of such "modal" expressions as have hopes and make the claim can, on the other hand, be omitted:42

(89) a. I'm making the claim the company squandered the money.
   b. I have a feeling the company will squander the money.

The paradigm can be regularized if the sentence is taken to be a complement of a reanalyzed (V + N) V rather than of an N.43 In most cases a single V exists that corresponds to the V + N complex: have hopes – hope, make the claim – claim, and so on.

Let us now turn to extractions from the sentential complement of a preposition, beginning with Italian.

At first sight, the well-formedness of the following examples would seem to provide evidence against (85):

(90) a. Un argomento di cui, non sono disposto a parlare t1, . . .
   a topic about which I am not willing to speak
   b. Una persona a cui, mi pento di non aver parlato t1, .
   a person to whom I regret not to have spoken

(90a–b) contain predicates that subcategorize for a PP headed by a and di, respectively. This can be seen from the ability of the Ps to be followed, with the same predicates, by a lexical NP instead of the infinitive:

(91) a. Sono disposto al ritiro.
   I am ready for the withdrawal
   b. Mi pento del mio gesto.
   I regret my gesture

It would thus seem plausible to analyze (90a–b) as instantiating the structure in (92),

(92) ···[PP[a/di]] [CP[PRO V\text{\textit{infinite}}]]···

parallel to the structure [PP[a/di]] [NP . . .] exemplified by (91).

These cases are generally taken to contrast with such cases as (93),

(93) a. Lui ha continuato a parlare.
   he went on talking

b. Lui cercava di parlare.
   he was trying to talk

for which the absence of corresponding a NP, di NP sequences (*Ho continuato alla lettura 'he continued to the reading'; *Lui cercava del colloquio 'he tried of the talk') renders it plausible to analyze a and di not as prepositions taking sentential objects but as infinitival complementizers followed by IP (for discussion along these general lines. See Kayne 1984, chap. 5, Rizzi 1982, chap. 3, and Manzini 1980).

If this standard analysis of (90)–(93) were correct, the conclusion that XP complements of [-V] categories are barriers for binding would indeed be falsified by the well-formedness of (90).

Before concluding from this that reference to [+V] categories in (85) is to be abandoned, we must be sure that the standard analysis is indeed correct. A deeper scrutiny, however, suggests that it is not.

One relevant observation is that, quite generally, the preposition of subcategorized PPs in Italian cannot take a sentential object.44 This is illustrated in (94)–(96), where the (a) cases show the P followed by a nominal object, the (b) cases a P followed by a tensed sentential object, and the (c) cases a P followed by an infinitival sentential object. The latter two cases are systematically excluded.45

(94) a. Contavo [sul \text{\textit{in}} la sua onestà].
   I was counting on his honesty
   b. *Contavo [su \text{\textit{in}} che Gianni fosse onesto].
      I was counting on that Gianni was honest
   c. *Contavo [su \text{\textit{in}} PRO essere onesto].
      I was counting on to be honest

(95) a. L'ho detto [da \text{\textit{in}} loro silenzi].
    I deduced it from their silence
   b. *L'ho detto [da \text{\textit{in}} che non sono stato accettato].
      I deduced it from that I was not accepted
   c. *L'ho detto [da \text{\textit{in}} PRO non essere stato accettato].
      I deduced it from not to have been accepted

(96) a. La sua fortuna consiste [in \text{\textit{in}} questo].
    his luck consists in this
   b. *La sua fortuna consiste [in \text{\textit{in}} che ha molti amici].
      his luck consists in that he has many friends
   c. *La sua fortuna consiste [in \text{\textit{in}} PRO avere molti amici].
      his luck consists in to have many friends
The generalization that subcategorized Ps in Italian cannot take a sentential object has apparently only two (partial) exceptions: the prepositions *a and *di. A ‘to/at’, when subcategorized, can be followed by a sentential object. For example:

(97) a. Sono contrario al [la tua partenza].
    I am against to your departure
b. *Sono contrario a [che tu parta subito].
    I am against to that you leave immediately
c. Sono contrario a [partire subito].
    I am against to leaving immediately

(“*” indicates that (97b) belongs to a more formal style than either (97a) or (97c).)

The preposition *di can apparently also be followed by a sentential object, but only when it is infinitival. For example:

(98) a. Sono contento del [la tua partenza].
    I am glad of your departure
b. *Sono contento di [che tu parta subito].
    I am glad of that you leave immediately
c. Sono contento di [partire subito].
    I am glad of leaving immediately

Note that if we abstract from the well-formedness of (97b), which belongs, in any event, to a different stylistic level, the behavior of the two prepositions becomes entirely parallel. The only normal type of sentential object they accept is the infinitival sentence.

This exception to the generalization that (“subcategorized”) Ps in Italian do not take sentential objects is very curious. For it is a striking coincidence that the only Ps partially contradicting that generalization (in that they may be followed by infinitival, though not tensed, sentences) are *a and *di, which we know independently have a usage as infinitival complementizers too.

The partial exception disappears if, instead, we take all instances of *a and *di followed by an infinitival sentence to be instances of their use as infinitival complementizers, even when they follow predicates that otherwise subcategorize for a PP headed by the preposition *a or *di (followed by an NP).

If we assume that, then the generalization turns out to hold virtually absolutely, and the strange limitation of *a and *di to infinitival complements is explained.

If so, the cases in (90) cease to be evidence that extraction from the complement of a P in Italian (perhaps, more generally, in Romance) is admitted. *A and *di are in the C of a subcategorized CP, which is quite compatible with the idea that CP complements of Ps are islands.

There is confirming evidence that such an extraction is indeed impossible. Recall that, in a more formal style of Italian, the P a must indeed be assumed to “subcategorize” for a CP, which can be either tensed or infinitival (the only marked real exception to the noted generalization; see (97)). If so, we do have a genuine case of the configuration (99), in Italian, in which to test the above prediction.

(99) ...[V_{pp} P_{cp} [sp ...]]]

The prediction is borne out, since extraction of a non-NP from a tensed complement of a appears to be impossible:

(100) a. *Carlo, [con il quale] Carlo with whom
    sono favorevole [sp a [che parlino tij]]... 50
    I am in favor that they speak
b. *Questo affare, [di cui], sono contrario [sp a [cp che trattiate tij
    this business of which I am against that you deal
    con loro]...]
    with them

These cases contrast with the following cases of “apparent NP extraction,” which are uniformly judged to be (relatively) acceptable:

(101) a. ?Questa disputa, che, sono favorevole
    this controversy which I am in favor
    [sp a [cp che voi dirimiate tij al più presto]]...
    of that you settle as soon as possible
b. Questo affare, che, sono contrario [sp a
    this business which I am against
    [cp che trattiate tij con loro]]...
    that you deal with them

This is not surprising given the possibility (to be justified in chapter 3) of analyzing these structures in terms of A-bound pros rather than in terms of genuine extraction with A-chain formation. On the basis of these facts, we thus conclude that the notion of barrier for binding (85) should indeed make reference to a [+V] category. If “preposition stranding” as four-A in English, though not in Romance, implies that the preposition is nondistinct from a verb (see below for discussion), then a specific prediction follows:
that extraction from a CP governed by a ("subcategorized") P should be possible in English but not in Romance, since the CP is a maximal projection \( \theta \)-marked in the canonical direction by a \([ + V ]\) category.

The prediction does not seem to be testable, however, given that CPs governed by a subcategorized P are disallowed in English as they are in Italian. One construction instantiating structure (99) would seem to be the NP-ing construction exemplified by such cases as I was counting on [him giving a present to Mary]. But Reuland (1983) provides evidence that the appropriate analysis construes the preposition in the C of the gerundive clause, so that the well-formedness of (102) is not informative:51

(102) a. The girl to whom, he was counting [CP on [them giving a present \( t_i \)]]
   b. The man from whom, we were looking forward [CP to [PRO receiving a letter \( t_i \)]]
   c. The man to whom, they insisted [CP on [PRO sending an invitation \( t_i \)]]

To summarize, we have seen that for binding to be possible, three requirements must be met simultaneously: that every maximal projection dominating the trace and not dominating the antecedent be (1) \( \theta \)-marked (either directly or indirectly) (2) by a \([ + V ]\) category, (3) in the canonical direction. If any one of these requirements is violated, the maximal projection counts as a barrier for binding and binding is impossible (even if the other requirements are satisfied).

The relevant examples showing a selective violation of each of these requirements are reproduced here for convenience:

(103) **Requirement of \( \theta \)-marking**

*Gianni, a cui, me ne sono andato [pp senza parlare \( t_i \)]]

Gianni to whom I went away without speaking

(104) **Requirement of \( \theta \)-marking by a \([ + V ]\) category**

*Gianni, da cui, disapprovo [cp i tentativi]

Gianni by whom I disapprove the attempts

*Gianni, a cui non so se [parlare \( t_i \) oggi]

Gianni to whom I don't know if to speak today

(105) **Requirement of canonical direction**

*How were you counting on him behaving in public?

b. *How strong do you look forward to drinking your coffee?
As with the notion of $\theta$-marking entering into the notion of barrier for binding, the Barriers notion of L-marking must, however, be restricted to $[+V]$ lexical heads. This is because neither nouns nor prepositions are apparently capable of L-marking their CP complements. For nouns, this was already pointed out in connection with such cases as (63), repeated here without structure as (109):

(109) *How did John announce a plan to fix the car?

Cases such as (110) show that the same holds for prepositions:

(110) In che modo eravate contrari [$_{vp}$ a
in which way were you against to
$_{cp}$ che lo trattassimo $t_i$]? that we treated him

The need to refer to $[+V]$ categories in the definitions of barrier for both binding and government thus renders the two notions interestingly symmetrical (at least in part):

(111) a. Definition of barrier for binding

Every maximal projection that is not $\theta$-marked by a $[+V]$ category in the canonical direction is a barrier for binding.

b. Definition of barrier for government

Every maximal projection that is not L-marked by a $[+V]$ category is a barrier for government.

In the next section we will consider further amending both definitions by referring to the more general notion of selection.

1.7.3 A Refinement: The Role of Selection

Consider the status of VP and IP with respect to the definition of barrier for government. These maximal projections are not L-marked because I and C are not lexical categories (and presumably because they also fail to $\theta$-mark VP and IP, respectively; see, see further below). Hence, they qualify as barriers for government both in Chomsky’s (1986b) original definition and in the modified version suggested here. They must not count as barriers, however. Chomsky voids their barrierhood in two distinct ways: by admitting adjunction to VP, in one case, and by stipulating that IP is not an inherent barrier, in the other (since adjunction to IP is not allowed).

Chomsky adduces essentially two reasons for not admitting adjunction to IP. One is to prevent extraction of an adjunct from a wh-island (*How did you wonder [$_{cp}$ who $[\_p\ t_i\ [\_r\ fixed\ the\ car\ ]\ t_i]]$?). If how could adjoin to IP, CP would no longer inherit barrierhood, in Chomsky’s system. This reason loses its force, however, if one adopts Relativized Minimality. Extraction of how is blocked independently by the intervening who, even if adjunction to IP is permitted.

The second reason involves the reduction of strong islands to Subjacency. If adjunction to IP were allowed (in a system where Subjacency is triggered by the crossing of two barriers), no such reduction would be possible, as is apparent from (112a–b):

(112) a. *The book O$_i$ that $[\_p\ t_i\ [\_r\ reading\ t_i]]$ would be fun]...

b. *To whom$_i$, did $[\_p\ t_i\ [\_r\ they\ [\_v\ leave\ [\_v\ before\ speaking\ t_i]]]]$?

Only one barrier would be crossed in (112a–b): CP and PP, respectively. In other systems of assumptions, however, such as the present one or Manzini’s (1988), where a bounding violation is triggered by the crossing of only one barrier, adjunction to IP could well be admitted. This would still be beside the point, however. To admit free adjunction to VP and IP is tantamount to saying that these maximal projections are never (inherent) barriers: a statement nondistinct from the mere stipulation that they are not barriers.

A more interesting tack would be to find a principled reason why they behave like maximal projections L-marked by a $[+V]$ category, although they are not. Perhaps L-marking is but a special case of a more general property that includes, as a distinct case, the relation between the non-lexical categories C and I and their complements IP and VP.

XPs directly $\theta$-marked by some head are c-selected (ultimately, s-selected) by it, together with indirectly $\theta$-marked XPs (see Chomsky 1986a). VP and IP are also c-selected, it seems, by I and C, respectively, even if not s-selected by them: I cannot take any complement other than VP and C cannot take any complement other than IP. 54

This suggests, then, the possibility of utilizing the notion of c-selection in the definition of barrier for government or, to generalize, direct selection: a notion that, for lexical categories, will ultimately mean direct s-selection and, for non-lexical categories, direct c-selection (the notion of selection also plays a crucial role in Longobardi’s (1985c, 1988) theory of bounding).

Finally, concerning the $[+V]$ requirement, it can be noted that I and C, if not intrinsically $[+V]$ categories, are at least compatible with $[+V]$ elements (witness their ability to host verbs in some languages). Thus, they can be taken to be (at least) nondistinct from $[+V]$ categories.

The definition of barrier for government can therefore be reformulated as follows: 55
Long and Successive Cyclic Wh-Movement

Second, it permits the elimination of certain technical problems arising from the decision to admit (Wh-)Movement via adjunction to maximal projections. One was the stipulation that no intermediate adjunction must be allowed to maximal projections that are arguments, the rationale for which is not entirely clear (see Chomsky 1986b). Another is the fact that (Wh-)Movement via adjunction to maximal projections that are not arguments must also be severely limited. For example, one must prevent adjunction to adverbial PPs (see Chomsky’s 1986b, 66) stipulation that only NPs can adjoint to such PPs, and the discussion in chapter 3, to extraposed relatives or comparative CPs, and to predicative NPs (see Longobardi 1987b, fn. 43). Otherwise, no barrier would be there to block extraction from them.

Third, as the case of clitic left dislocation discussed in chapter 2 shows, the transparency of VP and IP for binding cannot be obtained by resorting to adjunction, since clitic left dislocation does not involve (Wh-)Movement, though it enters (“base-generated”) binding relations.

Finally, expressing the transparency of VP and IP for government independently of adjunction permits an optimally simple account of Head Movement and NP-Movement, which cannot resort to adjunction. We briefly discuss this in the next section.59

1.7.4 Head Movement and A-Chains

In Chomsky 1986b, sec. 11, the fact that Head Movement and NP-Movement only apply in a strict successive cyclic fashion (that is, only enter chains of antecedent government relations) follows essentially from abandoning the \(\theta\)-government half of proper government for Vs (or, rather, from subsuming it under the antecedent government half). That is, \(\theta\)-government by V is no longer sufficient to satisfy the ECP. This has practically no consequences for A-movement, for which adjunction to VP always yields proper government via antecedent government, and for which the \(\gamma\)-marking procedure at S-Structure for A-positions ultimately permits long Wh-Movement. It has the desired consequences for Head Movement and NP-Movement, since these can never adjoin to VP (without producing “improper movement”); see Chomsky 1986b, 73).

If, however, long Wh-Movement is nothing other than a binding relation, and if a binding relation requires a referential trace, then the strict successive cyclic derivation of Head Movement and NP-Movement may be taken to follow plainly from the nonreferential character of head traces and NP-traces, as noted earlier.
The point at issue is a different one, however. As noted, the definition of barrier for government assumed here, which considers VP and IP intrinsically transparent for government (that is, noninherent barriers) appears to yield a surprisingly simple account of Head Movement and NP-Movement (in interaction with Relativized Minimality).

Concerning V-Movement, for example, it is no longer necessary to hold the questionable (Chomsky 1986b, 71) assumption that I θ-marks VP in order to render L-marking of VP possible after V-Raising to I (in languages that allow it) to avoid the barrierhood of VP. Concerning NP-Movement, no extensions of the notion of chain (and antecedent government) of the type suggested by Chomsky (1986b, sec. 11) are needed anymore. If VP counts as an inherent barrier, a simple case of passive like (116)

(116) [IP John, [I [VP was [VP invited t1]]]].

requires an extended chain (Johne, t1, was, invited, t1) with chain coindexing via Spec-Head and Head-Head agreement, and the convention that there be no ‘accidental coindexing,’ to exclude superraising cases (see Chomsky 1986b, 75ff.). If VP is not a barrier, NP-Movement will apply undisturbed.

A question arises in connection with superraising and superpassive cases:

(117) a. *John, [VP seems ([IP it [VP appears [IP it to be intelligent]])]]

b. *John, [VP seems ([IP it [VP was [VP told t1 that he should leave]])]]

Here it would seem that NP-Movement is allowed to apply in too unrestricted a manner, for in (117a–b) none of the intervening maximal projections qualifies as a barrier for government under the definition suggested above, since each of them is directly selected by a head nondistinct from a [+V] category.

The derivations shown in (117a–b) are nonetheless blocked by Relativized Minimality. An A-position (the [NP, IP]-position filled by it) intervenes between John and its trace. (For the same conclusion, see Rizzi 1990, chap. 1).

1.8 The Conditions on the Trace: The (Residue of the) ECP

The notions of barrier for binding and for government can be viewed, in a certain sense, as conditions on extraction domains (for binding and government, respectively). They impose certain requirements on the domains (the XPs) that intervene between the trace and the c-commanding antecedent. What remains to be discussed is the nature of the conditions on the trace of a binding or government relation.

A condition on traces already exists: the ECP, now reduced, following Rizzi (1990) and Koopman and Sportiche (1988), to the “formal licensing” requirement of the conjunctive definition (the “identification” requirement, θ-government or antecedent government, having been subsumed under the two relations of binding and government). The formal licensing requirement, given earlier as (ia) in note 4, is repeated here as (118):

(118) Definition of the ECP

A nonpronominal EC must be properly head-governed.

The question is, Does this condition suffice, or are there further conditions, possibly distinct for traces of binding and government?

In this section I will argue that the conditions on the traces of binding and government are indeed the same and that they reduce to the ECP, provided that we slightly modify it by introducing into its definition reference to a head nondistinct from [+V], as turned out to be necessary for the definitions of binding and government barrier.

Consider (119a–c). As their ungrammaticality indicates, neither binding nor government is possible in such contexts:

(119) a. *Il presidente, al quale ho sentito
the president to whom I heard
[VP un appello t1] alla radio…
a petition on the radio

b. *Gianni, il quale, ho parlato
Gianni whom I spoke
[PP con t1]…
with

c. *Gianni, da cui ci comportiamo
Gianni from whom we behave
[ADv diversamente t1]…
differently

(Compare Ci comportiamo diversamente da Gianni.)

First, let us approach the question from the point of view of binding (which is potentially available since the trace appears to satisfy the conditions on binding discussed above: being in an A-position; having a referential θ-role; and being compatible with a referential use). The forms in (119) are not excluded by the theory of bounding developed above (since no barrier for binding intervenes between the trace and its antecedent); nor are they excluded, for that matter, by Subjacency (as formulated in Chomsky
Long and Successive Cyclic Wh-Movement

(125) a. He saw [a picture [of whom]]? (D-Structure)
   b. He saw [a picture [t] [of whom]]? (intermediate structure)
   c. Of whom did he see [a picture [t] [t]]? (S-Structure)

Chomsky (1986b, 45), however, reports examples like the Spanish (126a) and English (126b) (= his (97c–c’) as real cases of long extraction of the θ-marked object of a noun:

(126) a. [De cual de estas ediciones] no sabes si hay traducción francesa
   b. Of which of these editions don’t you know whether there is a
   French translation?

He contrasts them with extractions of the subject of a NP, which (for ECP reasons) have access only to a successive cyclic derivation (blocked by an intervening wh-island). See (127a–b) (= Chomsky’s (97b–b’)), which are judged deviant:

(127) a. [De qué pintor] me preguntaste si van a exponer varios dibujos?
   b. By which painter did you ask me whether they are going to
   exhibit several drawings?

That (126a–b) involve a genuine case of long extraction from NP is dubious, however. (Long or successive cyclic) extraction of genitive complements of Ns is normally excluded, as shown for example by (119a) and by (128):

(128) The country on which we remember [an attack] is Poland.

It would thus be surprising if only genitive complements of Ns were to be “long extractable.”

Suppose that they are not, just like any other complement. The well-formedness of (126a) must then be explained differently. Genitive PPs also appear to participate in a construction in which the PP is base-generated in initial position, yielding either a “whole/part” or a special “aboutness” interpretation. Base generation is, for example, inescapable in such cases as (129) (see Barbaud 1976, Koster 1984, 65):

(129) a. Dei tuoi parenti, mi chiedo se anche Gianni
   of your relatives I wonder if Gianni too
   sarebbe disposto a farlo.
   would be willing to do it

b. Di quale libro non ricordi
   of which book don’t you remember any longer
   più ora
   se trattasse di questi argomenti?
   if it dealt with these topics
Long and Successive Cyclic Wh-Movement

If we had reasons to postulate the same condition for the trace of government, then we would have an indication that one and the same condition does hold for both kinds of traces—a condition that looks too close to the ECP to be a different principle.\(^{64}\)

(132) Binding and government traces must be governed by a head non-different from \([+V]\).

Let us turn, then, to government relations for evidence to this effect. The ill-formedness of (119a–c), as instances of successive cyclic Wh-Movement, follows in the Barriers system from the presence of a minimality barrier (N, P and Adv, respectively).\(^{65}\) In the present system, which adopts Relativized Minimality, no such explanation is available. A head cannot count as a potential antecedent for a trace of \(\Lambda\)-movement. The successive cyclic derivation of (119) is, however, excluded if traces of government too must be governed by a head non-different from \([+V]\). As noted, it is plausible to assume that this condition on binding and government traces is nothing other than the ECP.

(133) (Revised) definition of the ECP

A nonpronominal EC must be properly head-governed by a head non-different from \([+V]\).\(^{66}\)

A number of problems for this reformulation of the ECP must now be considered and explained away.

A first problem concerns the alleged capacity of nouns to head-govern the trace of internal NP-Movement in passive cases like the city's destruction \(t_i\).

A second, related problem is the N's apparent capacity to head-govern the trace of a postnominal subject extracted from NP via Spec, in Romance (see Longobardi 1987b). Given the arguments discussed by Giorgi and Longobardi (1990, chap. 3) that subjects of NPs (perhaps, more generally, XPs) are generated to the left of the head in Romance (and in Spec, in Germanic), the trace of the extracted subject must find an appropriate head governor.\(^{67}\)

Since Kayne's important work of the early 1980s (see, in particular, Kayne 1981b, 1983), it is generally admitted that nouns, unlike verbs, are not "structural governors" (that is, they cannot govern positions that they do not subcategorize)—whence the ill-formedness of (134a–c) and the like:

(134) a. *Mary's likelihood \(t_i\) to leave

(Compare Mary is likely to leave.)

b. *Mary's appearance \(t_i\) nice

(Compare Mary appears nice.)
Long and Successive Cyclic Wh-Movement

compatible with the fact that Italian disallows simple preposition stranding (recall (119b)), if we think that a P may be underspecified for [+V] only if it is nondistinct from V. Prepositions followed by an NP are distinct from V in Italian in terms of Case assignment, as originally proposed by Kayne (1981b), but they are nondistinct from V, even in Italian, whenever they do not assign Case, as when they take a PP complement.

There is supporting evidence for this conjecture. Extraction of the preposition’s PP complement in Italian obeys the same conditions that hold for simple preposition stranding in English: the PP from which extraction takes place must be subcategorized (governed) by a predicate (see the contrast between (138a) and (138b), noted in Rizzi 1988, 526):

(138) a. La ragazza alla quale Gianni si era messo accanto ti..., the girl to whom Gianni sat near
b. *La ragazza alla quale Gianni era felice accanto ti..., the girl to whom Gianni was happy near

The fact that when the relevant requirements are satisfied long extraction (binding) is possible (see (139a–b), which contain extractions from two weak islands) suggests that movement does not necessarily proceed via Spec PP (unlike, perhaps, what happens in Germanic, for which see Van Riemsdijk 1978 and Longobardi 1987b, 43):71

(139) a. Gianni, al quale, mi chiedo quando siano andati
Gianni to whom I wonder when they have gone
[pp incontro ti],...
toward
b. Gianni, al quale, non era caduta [pp addosso ti],...
Gianni (to) whom she had not fallen on

I thus tentatively conclude that the condition on traces (ECP) should refer to nondistinctness from [+V] heads in the general case, as do the conditions “on extraction domains” (the definitions of barrier for binding and government). This parallelism thus confirms the insight that underlies various analyses of bounding theory, from Kayne’s (1981b) extended notion of the ECP, to Huang’s (1982) CED, to Longobardi’s (1985b, 1988) symmetry principles, that the conditions on extraction domains are largely similar to the conditions on the trace.

1.9 Some Remarks on Wh-Islands and Subjacency

As we have seen, the islandhood of wh-islands for antecedent government relations (as in adjunct extraction) follows from Relativized Minimality.
Traditionally, wh-islands are also taken to be "mild" islands for long Wh-Movement (binding), falling under Subjacency, perhaps differently parametrized from language to language. See Rizzi 1982, chap. 2 and Chomsky 1981 for the original suggestion that the nodes counting for Subjacency are S (now IP), in English, and Σ (now CP), in Italian, giving different effects for the two languages.

Chomsky (1986b, 36–39) reinterprets the standard account along the following lines. Consider (140a–c):

(140) a. [To whom], did you wonder [CP what, in give t1/t1]?
   b. * [To whom], did you wonder [CP what, in they gave t1/t1]?
   c. [A chi, ti chiedevi [CP che cosa, in avessero dado t1/t1]]?

In such cases as (140a) movement of to whom to the matrix clause VP, and ultimately Spec CP, crosses one barrier (by inheritance), CP. If the most embedded tensed IP, in English, and CP, in Italian, is taken as an extra barrier, then two barriers are crossed in (140b), IP and CP, though still only one is crossed in (140c) (CP, which is a barrier both inherently and by inheritance). This roughly produces the reported judgments on (140). The lenient acceptability of (141)–(142) is taken by Chomsky (1986b, 38) to "suggest [...] that the violations are cumulative":

(141) a. What did you wonder [CP who [VP knew [CP who [VP saw t]]]]?
   = (Chomsky's (80))
   b. What did you wonder [CP who [VP said [CP who [VP saw t]]]]?
   = (Chomsky's (82a))

(142) a. Che cosa ti domandavi [CP chi [VP sapessi [CP chi [VP aveva visto t]]]]
   b. Che cosa ti domandavi [CP chi [VP avesse detto [CP che Bill [VP aveva visto t]]]]

In the system developed so far, essentially the same results could be obtained if we were to consider a tensed interrogative wh-clause as a special binding barrier (stronger in English-type languages than in Italian-type languages), with cumulative effects. However, this general approach raises some conceptual and empirical questions, the latter due in part to the often unclear status of the data. To begin with, English and Italian might be (and indeed have been claimed to be) much closer to each other than originally assumed, even though it is always delicate to compare judgments cross-linguistically. For one thing, many English speakers appear to accord fairly closely with the Italian judgments (see Grimshaw 1986, Chomsky 1986b, 37); moreover, extraction of interrogative wh-phrases from wh-islands in Italian gives rise to relatively degraded sentences, perhaps similar to the corresponding English cases considered originally (see Rizzi 1982, chap. 2, fn. 5, and related text). I suggested earlier that in selecting a wh-phrase to be extracted from a wh-island, care should be taken to choose one that allows a referential reading (and can thus enter into binding). Otherwise, binding will be unavailable and only the successive cyclic option will be left, which is filtered out by Relativized Minimality. Bare interrogative wh-phrases (who, what, chi, che cosa, and so on) appear to allow such a referential reading only marginally (as opposed to which-phrases, relatives, topicalized phrases, and the like)—a property that is perhaps at the root both of the intermediate status of many of the examples discussed in the literature and of the "mild" islandhood itself for long Wh-Movement attributed to wh-islands.

When a suitable phrase and context are chosen, even "double wh-island violations" in fact appear to be possible, both in Italian and in English:

(143) a. Carlo ? Maria, a cui, mi hanno chiesto se sapessi a cui
   Carlo è una persona a cui, non so
   Carlo is a person whom I don't know
   chi potrebbe chiedersi se affidare i propri figli ti help them
   chi potrebbe chiedersi se affidare i propri figli ti help them
   who could wonder whether to entrust his own children
   Maria, a cui, mi hanno chiesto se sapessi Maria, a cui, mi hanno chiesto se sapessi
   Maria to whom they asked me if I knew
   who had written
   Maria to whom they asked me if I knew
   Maria, a cui, mi hanno chiesto se sapessi Maria, a cui, mi hanno chiesto se sapessi
   Maria to whom they asked me if I knew
   Maria to whom they asked me if I knew
   Maria to whom they asked me if I knew
   Maria to whom they asked me if I knew
   Maria to whom they asked me if I knew
   Maria to whom they asked me if I knew

(144) a. A car that, I wouldn't know who to ask how to fix t1/... 74
   b. ? These are the only vegetables which, I don't know where to
   a. A car that, I wouldn't know who to ask how to fix t1/... 74
   b. ? These are the only vegetables which, I don't know where to
      find out how to plant t1.

The optimal case both for the Barriers system of assumptions and for the one developed here would be for long Wh-Movement to be essentially free out of wh-islands. 75 The facts appear to be almost so, but not quite. Certain residues remain. One is the tensed/untensed contrast, which is also found in other constructions (parasitic gap, easy-to-please, and related null object constructions, which will be discussed in chapter 3). A second is related to the availability of preposition stranding in English (but not in Italian), which gives rise to certain language-internal contrasts that are not entirely understood. Chomsky (1986b, 39) discusses the marginality of (145b), more severe than that of (145a):

(145) a. To whom did you wonder what John gave?
   b. Who did you wonder what John gave to?
A third residual problem is the fact that long Wh-Movement (binding) across one (or more) wh-island(s) degrades more rapidly, as structure is added, than that of ordinary complement CPs, although it never appears to give rise to discrete and dramatic changes, as does successive cyclic movement (government) out of a wh-island.

Pending better insights into these residual issues, I will conclude my discussion here and turn to the more general question of Subjacency as a separate condition on long Wh-Movement (binding).

On the basis of the empirical phenomena considered thus far, it seems that the notion of (1-)Subjacency is both too strong and too weak. It is too strong in that it rules out long Wh-Movement from two wh-islands (under the cumulative interpretation), whereas some such extractions must be admitted as in (143)–(144). It is too weak in that it fails to rule out long Wh-Movement from (certain) adjuncts, from relative clauses extraposed from an object or from degree clauses, all of which involve but one barrier (see section 1.6.2).

A question that arises is whether (abstracting from the important methodological role that the notion has played) (1-)Subjacency is but an artifact of two particular generalizations that may turn out not to be genuine. One is the decision to group together structures, such as the wh-island, the complex NP island, the adjunct island, and the subject island, that do not form a homogeneous class (one weak island and three strong islands). We have seen that, leaving aside wh-islands as the odd member, Subjacency could be interpreted as 0-Subjacency, with the notion of barrier for Subjacency (binding/bounding) defined as in (114), “every maximal projection that fails to be selected in the canonical direction by a category nondistinct from [+ V].” The other is the decision to employ the same notion of barrier for binding/bounding that holds for government. This decision also imposes 1-Subjacency, rather than 0-Subjacency, since the barrier that suffices for government (say, a non-L-marked factive complement CP) must not suffice for Subjacency. This in turn forces the introduction of a notion of barrier “by inheritance,” which is entirely dispensable for government and can be dispensed with for Subjacency too, if we give two (partially) different definitions of barrier for the two theories, as suggested.

1.10 Summary

We began the chapter by posing three questions, (8a–c). Concerning the first (What classes of elements undergo long and successive cyclic Wh-

Movement?), we found evidence that long Wh-Movement is limited to phrases that not only are in A-positions (Chomsky 1986b) and receive a referential θ-role (Rizzi 1990), but also are intrinsically referential (D-linked, in Pesetsky's (1987) terms). All phrases not satisfying these conditions have access only to successive cyclic Wh-Movement. This motivates recognizing two different types of operator/variable configurations, depending on the referential/nonreferential nature of the trace (and the relative operator).

Concerning the second question (From what principles of the theory does the existence of long and successive cyclic Wh-Movement follow?), we have essentially adopted Rizzi's (1990) theory, which recognizes only two ways in which the content of a trace can be identified: binding and (antecedent) government, which subsume long and successive cyclic movement, respectively. We have only slightly departed from Rizzi's system in admitting binding chains in addition to government chains.

Concerning the third question (What is the nature of the locality conditions on long and successive cyclic Wh-Movement?), we have tried to eliminate certain asymmetries and redundancies from the Barriers system of assumptions, eventually arriving at an interestingly simple set of locality conditions,

1. A (single) definition of barrier for binding/bounding

(114) Every maximal projection that fails to be (directly or indirectly) selected in the canonical direction by a category nondistinct from [+ V] is a barrier for binding.

2. A (single) definition of barrier for government

(113) Every maximal projection that fails to be directly selected by a category nondistinct from [+ V] is a barrier for government.

3. A condition of head government on traces (the ECP)

(133) A nonpronominal EC must be properly head-governed by a head nondistinct from [+ V],

in combination with Rizzi's (1990) notion of Relativized Minimality. All three conditions make crucial reference to a head nondistinct from [+ V], and the first two crucially refer to the notion of selection, whose introduction has also permitted other simplifications, from disposing of movement via adjunction to maximal projections, to eliminating the Barriers special notion of extended chain for NP-Movement.