Linguistic Inquiry Monographs
Samuel Jay Keyser, general editor

1. *Word Formation in Generative Grammar*
   Mark Aronoff
2. *$\bar{X}$ 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
5. *An Encyclopedia of AUX: A Study in Cross-Linguistic Equivalence*
   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 $\bar{A}$-Dependencies*
    Guglielmo Cinque
18. *Argument Structure*
    Jane Grimshaw
19. *Locality: A Theory and Some of Its Empirical Consequences*
    Maria Rita Manzini

Locality: A Theory and Some of Its Empirical Consequences

Maria Rita Manzini

The MIT Press
Cambridge, Massachusetts
London, England

existence justified by the discovery of some new type of data and/or by the systematization of some previously unanalyzed language. Rather, the goal of the book is to organize the considerable mass of evidence now extant on local dependencies according to what I believe to be genuinely novel principles. Its main justification is that the proposals it advances, which in turn represent a simplification of other available analyses, place some of the most poorly understood evidence in the literature within a theoretical framework.

The chapters in the book differ somewhat in character. Chapter 1 is meant to provide accessible terms of reference for the ensuing debate. Though the perspective is admittedly biased, it is so in one respect only, namely, in that the crucial property of each theory examined is taken to be the extent to which it does or does not allow for unification. Thus, it could easily be used as (part of) a graduate introduction to locality theory. Chapter 2 presents the core statement of my theory, developing it step by step from previous models in the literature.

Chapter 3 represents the more open-ended side of the enterprise. It delves into various sets of data that either have been used to support alternative theories or on the contrary remain mysterious under such theories. Finally, in its central concern for a basic statement of the theory, chapter 4 is essentially a continuation of chapter 2, concerning locality conditions on anaphors and pronouns rather than on movement.

Chapter 1
The Locality Problem

1.1 The Theoretical Problem and Some Residual Empirical Ones

Two types of grammatical dependencies can be distinguished according to the locality constraints they are subject to. The first type is characterized by strict locality conditions. These "bounded" dependencies include (for instance) anaphoric relations with reflexives and reciprocals, and their mirror image, disjoint reference relations with pronouns. Thus, the well-formed anaphoric dependency in (1), which does not cross a sentence boundary, contrasts with the ill-formed dependency in (2), which does. Conversely, disjoint reference with a pronoun is enforced internally to a sentence, as in (3), but not across a sentence boundary, as in (4).

(1) John_i likes himself_i

(2) *John_i thinks that [Mary likes himself_i]

(3) *John_i likes him_i

(4) John_i thinks that [Mary likes him_i]

A second type of dependency that displays a strict locality behavior is A-movement. Thus, raising is allowed from an embedded subject to an immediately superordinate subject, as in (5) and (6); and successive applications of raising, as in (7), also produce well-formed results. But long-distance raising, crossing an intermediate subject, as in (8), is ill formed.

(5) John_i appears [t_i to win]

(6) John_i is likely [t_i to win]

(7) John_i appears [t_i to be likely [t_i to win]]

(8) *John_i appears that [it is likely [t_i to win]]

Examples like those in (1)–(8) and a discussion of their theoretical significance are found in Chomsky 1973. A third type of strictly local dependency, distinguished more recently, is head movement. Thus, we assume that a sentence like (9) includes a C(omplementizer) head, represented by *if*, an I(nflection) head, represented by the modal *will*, and a number of V heads, represented by the auxiliary *have* and by *finished*.

(9) ... if [Bill will [have [finished his work]]]

In the absence of a modal it can be argued, as does Pollock (1989a), that the auxiliary must move to I, in order to provide a realization for the inflectional material, as in (10).

(10) ... if [Bill has$_i$ [t$_i$ [finished his work]]]

Furthermore, in a main sentence, a modal or an auxiliary must move to C in order for a question to be formed (say, for the reason detailed by Rizzi (1991)), as in (11)–(12).

(11) Will$_i$ [Bill t$_i$ [have [finished his work]]]

(12) Has$_i$ [Bill t$_i$ [t$_i$ [finished his work]]]

But suppose an auxiliary is moved to C from its original V position, bypassing I as in (13), where I is already filled by a modal; the result is ungrammatical.

(13) *Have$_i$ [Bill will [t$_i$ [finished his work]]]

The descriptive generalization corresponding to the paradigm (10)–(13) is Travis's (1984) Head Movement Constraint, which states that movement is only possible from a head position to the immediately superordinate one. Thus, strict locality is again observed.

Contrasting with the dependencies in (1)–(13) is a class of grammatical dependencies that appear to be "unbounded," in that the two members of the dependency can occur an indefinite distance apart. A typical example is the sort of $\bar{A}$-movement shown in (14)–(16); and one can imagine examples in which the distance between the *wh*-phrase and its trace, measured for instance in terms of sentence boundaries crossed, is even greater.

(14) Who$_i$ do you like t$_i$

(15) Who$_i$ do you believe [Peter likes t$_i$]

(16) Who$_i$ do you believe [Peter said [that Mary likes t$_i$]]

As is well known, there is evidence, going back to Ross (1967), that this type of dependency is itself severely constrained. Thus, $\bar{A}$-movement out

of a subject, as in (17), out of an adjunct, as in (18), or out of a relative clause, as in (19), is ill formed, regardless of the total distance (again measured, for instance, in terms of sentence boundaries crossed) between the *wh*-phrase and its trace.

(17) *What$_i$ does [explaining t$_i$] bother you

(18) *What$_i$ was Mary bothered [because Peter explained t$_i$]

(19) *What$_i$ do you know the girl [that explained t$_i$]

Nevertheless, though *wh*-dependencies are subject to locality constraints in the form of islands, as in (17)–(19), contrasts like the ones between (2) and (15) and between (8) and (15) remain to be explained.

The situation is further complicated by the contrast discovered by Huang (1982) between the locality behavior of $\bar{A}$-dependencies created by arguments, such as those in (14)–(19), and that of $\bar{A}$-dependencies created by adjuncts, such as *how* or *why*. Thus, extracting an argument out of a sentence introduced by a *wh*-phrase produces an interpretable sentence, as in (20), whereas extracting *how* does not, as in (21). The Italian counterpart (22) of the English sentence (20) is perfectly good, as first noted by Rizzi (1980); but the Italian sentence (23), like its English counterpart (21), is uninterpretable.

(20) What$_i$ do you wonder [how$_j$ to fix t$_i$ t$_j$]

(21) *How$_i$ do you wonder [what$_j$ to fix t$_j$ t$_i$]

(22) Cosa$_i$ ti chiedi [come$_j$ aggiustare t$_i$ t$_j$]

(23) *Come$_i$ ti chiedi [cosa$_j$ aggiustare t$_i$ t$_j$]

Nevertheless, configurations like (14)–(16), where apparently unbounded *wh*-extraction of arguments is possible, also allow apparently unbounded extraction of adjuncts, as in (24)–(26).

(24) How$_i$ do you fix it t$_i$

(25) How$_i$ do you believe [Peter fixes it t$_i$]

(26) How$_i$ do you believe [Peter said that [Mary fixes it t$_i$]]

Suppose we assume with Aoun (1985) and Rizzi (1990), among others, that $\bar{A}$-dependencies involving arguments, as in (14)–(16), and only those, are in fact unbounded, in the sense that movement can take place in one step over an arbitrary number of sentence boundaries. By contrast, suppose that, as proposed by Chomsky (1973), the unboundedness of

Ā-dependencies involving adjuncts, as in (24)–(26), is only apparent. The adjunct moves in strictly local fashion, stopping in the initial position of each sentence, which is independently known to host *wh*-phrases; in other words, (25) and (26) have derivations like (27) and (28).

(27) How_i do [you believe [t_i [Peter fixed it t_i]]]

(28) How_i do [you believe [t_i [Peter said [t_i that [Mary fixes it t_i]]]]

Under these assumptions, the contrast between (20) and (21) becomes transparent. In (20) the argument *wh*-phrase can indeed move in one step, as indicated. But in (21) the adjunct *wh*-phrase ought to move first to the position that is already occupied by *what*. Since the presence of *what* makes this movement impossible, we correctly predict that (21) will be ill formed.

To sum up: If the preceding observations are correct, there are at least four types of grammatical dependencies that display strictly local behavior: anaphora, A-movement, head movement, and Ā-movement of adjuncts. There is also at least one type of dependency that does not display strictly local behavior, namely, Ā-movement of arguments (though this type of dependency too is subject to locality constraints in the form of islands, as in (17)–(19)). To complete the picture, strictly local dependencies are also subject to islands, as illustrated with *wh*-extraction of adjuncts in (29)–(31).

(29) *How_i did [fixing it t_i] bother Mary

(30) *How_i was Mary bothered [because John had fixed it t_i]

(31) *How_i do you know the girl [that fixed it t_i]

We are now in a position to formulate the question that this book addresses. Of course, each of the five above-mentioned dependencies can simply be listed along with its particular locality behavior, as in the previous paragraph. But can a unified account of these behaviors be given? All current theories achieve partial unification. What I propose to show here is that complete unification is possible.

Current theories identify two classes of configurations across which a local dependency cannot be formed. The first corresponds to islands of the type shown in (17)–(19) and (29)–(31), which appear to be entirely insensitive to the nature of the elements involved. In essence, according to Huang's (1982) classification, dependencies that cross object boundaries are allowed, as in (14)–(16) and (24)–(26), but not dependencies that cross subject boundaries, as in (17) and (29), or dependencies that cross

adjunct boundaries, as in (18)–(19) and (30)–(31), where the relative clause in (19) and (31) can be assimilated to an adjunct.

The second type of island that current theories identify appears to be crucially sensitive to the nature of the elements involved in dependencies. Thus, heads cannot form dependencies across other heads, as in (13) versus (10)–(12), and adjunct *wh*-phrases cannot form dependencies across other *wh*-phrases, as in (21) versus (24)–(26). More generally, following Rizzi (1990), an element cannot form a dependency across a potential antecedent, thus accounting for the contrasts between (1) and (2) and between (4)–(7) and (8) as well. Of course, in the contrast between (3) and (4) the ability of the pronoun to form a dependency only across a potential antecedent is involved. Argument *wh*-phrases constitute an important exception to this type of behavior, however, in that they are insensitive to antecedent islands, as in (20).

One of the immediate aims of this book is to provide a theory under which the difference with respect to antecedent islands between Ā-dependencies involving arguments and other dependencies does not need to be stipulated—in other words, to eliminate the disjunction between antecedent government and head government in Chomsky's (1981) ECP. The theoretical reasons for seeking nondisjunctive locality principles are evident. Such principles will yield a simpler grammar, where simpler can be understood as more highly modular and/or richer in deductive structure. If they are not simply a notational variant of other theories, then they will also have empirical motivation—in other words, they will be able to predict some new set of behaviors.

Indeed, there are a number of known island behaviors that cannot (immediately) be classified as belonging to either type of island illustrated in (1)–(31). Thus, Ā-movement of an argument is well formed across a single nominal boundary, as in (32), but ill formed across adjacent nominal and sentential boundaries, as in (33), even if the sentence is an argument of N.

(32) Who_i did you see [many portraits of t_i]

(33) *Who_i did you see [many attempts [to portray t_i]]

Since no adjunct or subject configuration is involved in (33), and no other *wh*-element is present, neither currently distinguished type of island subsumes the island in (33), at least without further stipulation.

As before, the situation is complicated by the asymmetry in the behavior of arguments and adjuncts under extraction. Thus, an adjunct cannot even be extracted across a single nominal boundary, as in (34).

(34) *[With what kind of sleeves]$_i$ did you see [many sweaters t_i]

Again, this deviates from the characterization of both classes of islands illustrated in (1)–(31). (33)–(34) instead appear to require some stipulation about extractability across N heads, as opposed to V heads.

Similarly, if Chomsky (1986a) is correct, whether or not a *wh*-dependency is well formed across a *wh*-island depends on the presence or absence of Tense in the island. Thus, by contrast with (20), (35) is ill formed.

(35) *What$_i$ did you wonder [how I repaired t_i]

Again, object configurations are involved throughout; furthermore, *wh*-dependencies involving arguments are not sensitive to *wh*-antecedents. Hence, the island in (35) appears to be irreducible to either of the two well-understood types.

Finally, although $\bar{A}$-dependencies can be created across a single nominal boundary, as in (32), their acceptability depends on the nominal being indefinite. The presence of a definite determiner creates an island, as in (36).

(36) *Who$_i$ did you see [the many portraits of t_i]

Again, the contrast cannot be explained in terms of object versus subject or adjunct configurations, or in terms of the presence of other *wh*-elements.

As pointed out by T. Hoekstra (personal communication), all of the islands in (33)–(36) can be seen as belonging to a single type, which involves the presence of a head with denoting properties of some sort: an N head in (33)–(34), a T head in (35), and a definite D head in (36). The first major empirical argument in favor of the theory proposed here will be that it allows the islands in (33)–(36) to be distinguished as a coherent class, and eventually to be derived.

Overlapping *wh*-dependencies constitute another island behavior that current theories cannot explain. Thus, *wh*-movement of an argument across another *wh*-element is well formed, as in (20). But if a third *wh*-element, again an argument, is moved, creating a triple extraction, the resulting sentence is ungrammatical, as in (37).

(37) *[Which books]$_i$ did you wonder [to which student]$_j$ to ask [whether to give t_i t_j]

The second major empirical argument in favor of the theory proposed here will be that it correctly predicts that no more than two $\bar{A}$-dependencies can coexist in any given portion of a tree, thus automatically excluding (37).

To see that the two types of islands in (33)–(36) and (37) interact, it is sufficient to consider what happens when more than one argument of N, rather than of V, is *wh*-moved: In Italian, if not in English, both (38) and (39), where two different arguments of N are *wh*-moved one at a time, are fundamentally well formed.

(38) [Di che pittore]$_i$ hai visto [un ritratto di Aristotele t_i]
of which painter have you seen a portrait of Aristotle

(39) [Di che filosofo]$_i$ hai visto [un ritratto t_i di Rembrandt]
of which philosopher have you seen a portrait of Rembrandt

However, if both arguments are *wh*-moved at once, as in (40), the resulting sentence is as unacceptable as (37).

(40) *[Di che pittore]$_i$ ti chiedi [di che filosofo]$_j$ comprare
of which painter do you wonder of which philosopher to buy
[un ritratto t_j t_i]
a portrait

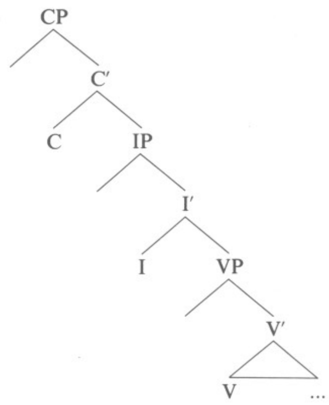
Thus, it appears that the presence of an N decreases the available *wh*-movement paths from two to one. This prediction also follows from the theory proposed here, as desired.

In brief, the general theoretical problem that we will be addressing is to unify the various locality conditions on grammatical dependencies. This is essentially a simplicity problem, or a problem in maximizing the modularity and/or deductive depth of the model. The empirical problems that we will address at the same time are the head islands in (33)–(36) and the multiple extraction islands in (37) and (40). I will claim that the theoretical problem and the empirical ones admit of one and the same solution.

1.2 The Barriers Solution

As a preliminary step, I will introduce some major current theories of locality, beginning with Chomsky's (1986a) barriers theory. According to Chomsky the canonical structure of a sentence includes two functional heads, I and C, in addition to the lexical head V, as in (41) (where the subject position is identified with the Spec of IP position).

(41)



The theory of Subjacency is based on the notion of L-marking, which in turn is based on the notion of θ-government. θ-government is θ-marking by a head, as in (42). L-marking is θ-government by a lexical category, as in (43).

(42) β θ-governs α iff β is a head, β θ-marks α, and β is a sister to α.

(43) β L-marks α iff β is lexical and β θ-governs α.

Subjacency is then defined in terms of the notion of barrier. A maximal projection that dominates a given category can be a barrier for it, inherently or by inheritance. It is inherently a barrier if it is not L-marked (i.e., in Chomsky's (1986a) terms, if it is a blocking category (BC)); it is a barrier by inheritance if it is the first maximal projection that dominates a BC. IP, however, is an exception. It can be a barrier by inheritance, and barrierhood can be inherited from it; but it cannot be an inherent barrier. The relevant definitions are as follows:

(44) β is a BC for α iff β is an XP, β dominates α, and β is not L-marked.

(45) β is a barrier for α iff β (other than IP) is a BC for α, or β is the first XP that dominates a BC for α.

Crucially, although the notion of barrier is defined in terms of dominance, Subjacency itself is stated in terms of exclusion, where dominance is defined as in (46), following May (1985), and exclusion is defined as in (47).

(46) β dominates α only if all segments of β dominate α.

(47) β excludes α iff no segments of β dominate α.

The interplay of dominance and exclusion has the effect of making any adjunction to a maximal projection into an escape hatch. However, adjunction to two maximal projections, CP and NP, is blocked by a constraint against adjunction to an argument. Furthermore, IP is exceptional in that nothing may be adjoined to it. (48) expresses these restrictions.

(48) Arguments and IPs cannot be adjoined to.

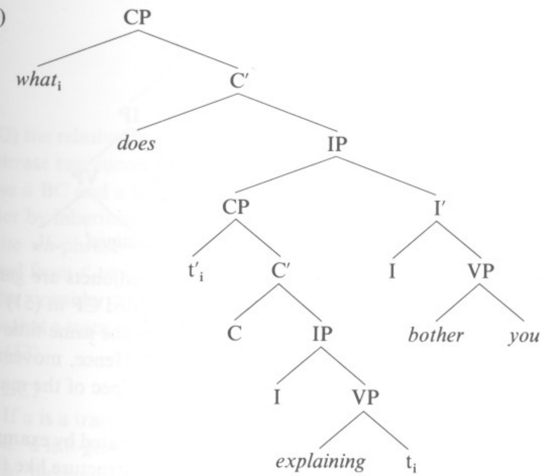
Under Subjacency no category can cross more than one barrier (i.e., no more than one barrier for the position moved from can exclude the position moved into); hence, Subjacency is violated if two or more barriers are crossed, as in (49).

(49) *Subjacency*
If α is a trace, there is an antecedent β for α such that at most one barrier for α excludes β.

The theory in (41)–(49) predicts three fundamental types of islands: subject islands, adjunct islands, and complex NP islands involving relative clauses. Consider first a typical subject island violation, as in (17), repeated here; the structure associated with it is (50).

(17) *What$_i$ does [explaining t$_i$] bother you

(50)

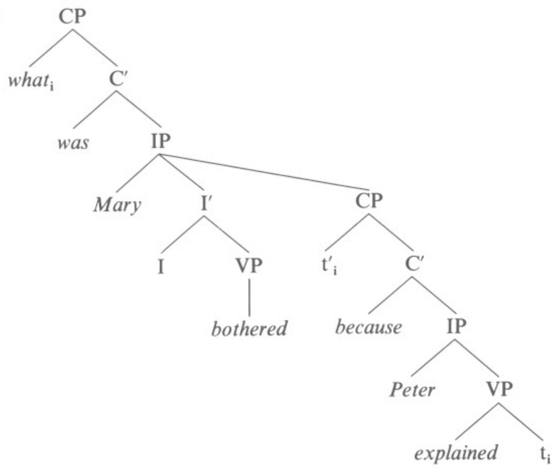


Here we assume that the *wh*-phrase being extracted can successfully move from *t* to *t′*. Its next possible landing site is the Spec of the matrix CP; but this movement crosses two barriers. The embedded CP is of course a BC and a barrier for *t′* since it is not θ-governed by a lexical head and hence is not L-marked; the matrix IP is a barrier for *t′* by inheritance, because it is the first maximal projection that dominates CP. Thus, the final movement violates Subjacency, and the ungrammaticality of (17) is correctly predicted.

Similar considerations apply in the case of adjunct islands. A relevant example is (18), repeated here; its corresponding structure is (51), where again we assume that the *wh*-phrase being extracted can successfully reach the position of *t′*.

(18) *What$_i$ was Mary bothered [because Peter explained t$_i$]

(51)

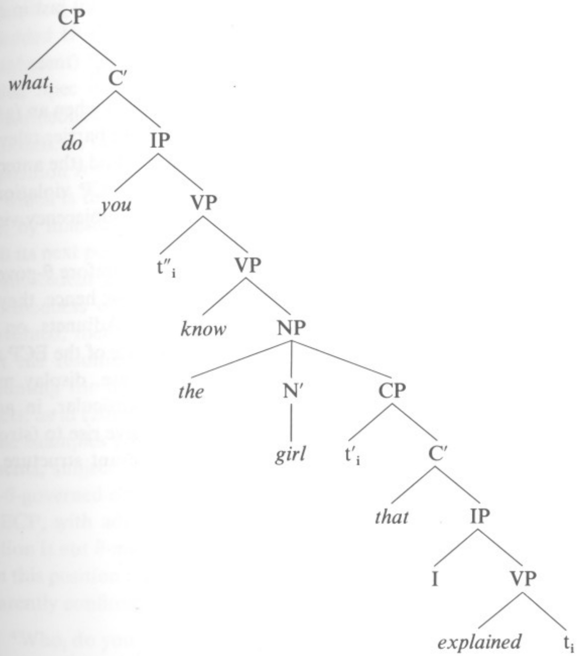


Following Chomsky (1986a), in (51) we assume that adjuncts are generated as sisters of I′. If so, it is obvious that the embedded CP in (51) is a BC and a barrier for *t′* because it is not L-marked; at the same time the matrix IP is a barrier for *t′* by inheritance from CP. Hence, movement from *t′* to the next possible landing site—namely, the Spec of the matrix CP—crosses two barriers and violates Subjacency.

Finally, complex NP island violations, which are illustrated by examples like (19), repeated here, are predicted on the basis of a structure like (52).

(19) *What$_i$ do you know the girl [that explained t$_i$]

(52)



In (52) the relative clause is a sister of N′, and we assume as usual that the *wh*-phrase can successfully reach the *t′* position. The embedded CP is of course a BC and a barrier for *t′* because it is not L-marked, and NP is a barrier by inheritance from CP. In this case the next possible landing site for the *wh*-phrase is a VP-adjoined position, *t″*; since two barriers are crossed from *t′* to *t″*, Subjacency is again violated.

Now consider then the ECP. Chomsky's (1986a) version of the ECP states that a trace must either be θ-governed or governed by an antecedent, as in (53).

(53) *ECP*

If α is a trace

a. α is θ-governed; or

b. there is an antecedent β for α such that β governs α.

θ-government has already been defined in (42); government on the other hand is defined to hold between a category and its antecedent just in case no barriers intervene between them, as in (54).

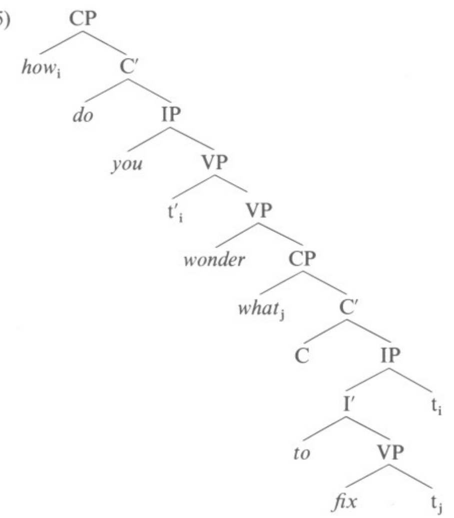
(54) β governs α iff there is no barrier for α that excludes β.

Hence, government does not hold and the ECP is violated when an (antecedent, trace) link crosses one barrier. Since the notion of barrier relevant for Subjacency is also relevant for the ECP, Subjacency and (the antecedent government clause of) the ECP differ in that an ECP violation is defined in terms of crossing just one barrier, whereas a Subjacency violation is defined in terms of crossing two or more.

Objects, which are clearly θ-marked by a head and therefore θ-governed, always satisfy the ECP under the θ-government clause; hence, they do not need to satisfy the antecedent government clause. Adjuncts, on the other hand, which do not satisfy the θ-government clause of the ECP and must therefore satisfy the antecedent government clause, display more restrictive locality patterns than direct objects. In particular, in addition to Subjacency violations like (50)–(52), adjuncts give rise to (strong) *wh*-island violations like (21), repeated here. The relevant structure for such violations is provided in (55).

(21) *How$_i$ do you wonder [what$_j$ to fix t$_j$ t$_i$]

(55)



t_i in (55) represents the base-generated position of the adjunct, on the assumption that an adjunct is a daughter of IP, much as in (51). The embedded IP is of course a BC since it is not L-marked; but it is not a(n inherent) barrier for t_i. Hence, if the adjunct can move to the embedded Spec of CP position, no violation ensues, accounting for the well-formedness of long-distance adjunct extractions in the absence of *wh*-islands, as in (25) (*How$_i$ do you believe Peter fixes it t$_i$*). If the Spec of CP position is filled, however, a barrier is crossed. Indeed, the embedded CP, which is the first maximal projection that dominates IP, is a barrier for t_i by inheritance. Thus, when the adjunct crosses the CP barrier to reach its next possible landing site, the VP-adjoined position t'_i, it violates the antecedent government clause of the ECP, correctly predicting the ill-formedness of (21). On the other hand, if the extracted element is θ-governed, then in the absence of any antecedent government requirement the resulting sentence is predicted to be well formed, correctly accounting for the (relative) acceptability of *wh*-island violations with objects, as in (20) (*What$_i$ do you wonder how$_j$ to fix t$_i$ t$_j$*).

The examples given so far have illustrated the behavior of θ-governed elements, subject only to Subjacency, with objects, and the behavior of non-θ-governed elements, subject to the antecedent government clause of the ECP, with adjuncts. Let us now turn to subjects. Since the subject position is not θ-marked by a head, it is not θ-governed; hence, extraction from this position is predicted to show *wh*-island effects. This prediction is apparently confirmed by ill-formed examples of the following type:

(56) *Who$_i$ do you wonder [what$_j$ t$_i$ painted t$_j$]

On the other hand, there is one major respect in which subject extraction differs from both object and adjunct extraction: the former exhibits *that-t* effects, but the latter does not. The relevant examples are (57)–(58), where (57), with no overt complementizer adjacent to the subject extraction site, is well formed, but (58), which contains a *that-t* configuration, is not. No comparable violation arises with objects, as in (59), or with adjuncts, as in (60).

(57) Who$_i$ do you believe [t$_i$ is a painter]

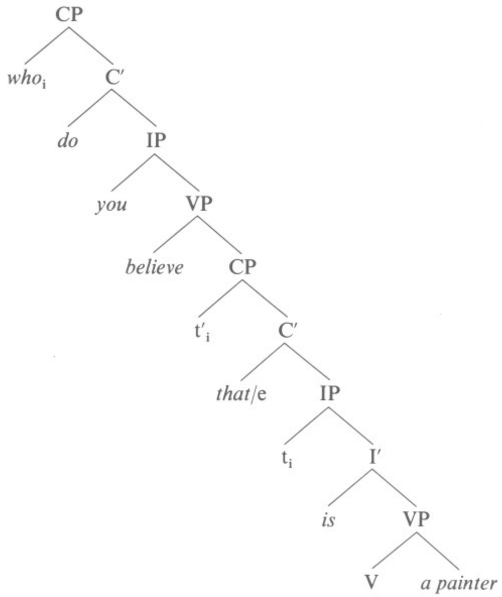
(58) *Who$_i$ do you believe [that t$_i$ is a painter]

(59) Who$_i$ do you believe [(that) Peter likes t$_i$]

(60) How$_i$ do you believe [(that) Peter fixes it t$_i$]

The relevant configuration for *that-t* violations is provided in (61).

(61)



Here, under the theory outlined in (41)–(49), movement from t_i to the next landing site (the Spec of CP position t'_i) creates no violation, whether the embedded C is overt or not. In order to predict *that-t* violations as violations of antecedent government, the notion of minimality barrier must be introduced, as in (62). As before, the exceptional status of I is confirmed, since I′ cannot be a minimality barrier. Furthermore, unlike other barriers, minimality barriers do not count for both Subjacency and government, but only for government, as in (63).

(62) β is a minimality barrier for α iff β is an X′ (other than I′), β dominates α, and the head of β, X, is lexical (X other than α).

(63) Minimality barriers are barriers only for government.

With these notions in place, consider (61). If the embedded C is lexical, then C′ is a minimality barrier for t_i. This means that movement to the Spec of CP position t'_i is blocked by antecedent government, because one barrier is crossed. Suppose, however, that the embedded C is empty. Then C′ is not a minimality barrier for t_i, and t_i can move to the Spec of CP

position t'_i without crossing any barriers, hence satisfying antecedent government. This explains the data in (57)–(58). As for (59), objects are predicted not to display *that-t* effects because they satisfy the ECP under the θ-government clause. The fact that I′ is never a barrier means that (62)–(63) have no consequences for movement (of objects) across I′ projections. However, a technical problem arises with respect to movement across V′, which should always be a minimality barrier. This is solved by assuming that in the absence of a Spec of VP, V′ is not projected and VP remains the only projection of V.

Finally, consider adjuncts. Essentially following Lasnik and Saito (1984), it is assumed that the relevant level for the satisfaction of antecedent government by traces in $\bar{A}$-position, and hence by adjuncts, is LF, as stated in (64), and that at LF *that* can systematically delete.

(64) Antecedent government applies to $\bar{A}$-traces at LF.

If so, then adjuncts are predicted not to display *that-t* effects, as the example in (60) indeed attests. Similarly, the theory predicts the contrast between (57)–(58) and (65).

(65) Who$_i$ do you believe [$_{CP}$ t‴$_i$ (that) [$_{IP}$ I [$_{VP}$ t″$_i$ said [$_{CP}$ t′$_i$ [$_{IP}$ t$_i$ is a painter]]]]]

In (57)–(58) the subject trace must be antecedent-governed at S-structure; as we have seen, it is so governed in the absence of *that*, as in (57), but not in the presence of *that*, as in (58). In (65), on the other hand, only intermediate traces are found on either side of *that*—namely, t''_i and t'''_i. Since intermediate traces are in $\bar{A}$-position, they only need to satisfy antecedent government at LF. Crucially, *that* can delete at LF, eliminating the minimality barrier it defines. Thus, (65) is well formed, whether *that* is present or absent at S-structure.

Head movement and A-movement remain to be accounted for. Head movement must satisfy antecedent government, since a head cannot be θ-governed. Antecedent government in turn correctly derives the descriptive generalization on head movement, the Head Movement Constraint. Consider for instance the contrast between (11) and (13), repeated here.

(11) Will$_i$ [Bill t$_i$ [have finished his work]]

(13) *Have$_i$ [Bill will [t$_i$ finished his work]]

Suppose that the base-generated position of the higher auxiliary in (11) and (13) corresponds to the I position in (41), and the position of the lower auxiliary to the V position in (41). Movement from the I to the C position satisfies antecedent government, since neither I′ not IP is ever a barrier except by inheritance. However, movement from the V to the C position crosses at least two barriers, VP and IP, the latter in this case a barrier by inheritance. Thus, (13) violates Subjacency as well as antecedent government.

The antecedent government clause of the ECP also accounts for A-movement. One problem that the theory faces, however, is that movement from the object position in (41) to the Spec of IP position is blocked by the VP barrier. Of course, such movement must be allowed in order for passive sentences to be generated, as in (66).

(66) John$_i$ was [hit t$_i$]

Since a movement dependency will in general relate the V and I positions in (41), Chomsky (1986a) proposes that this dependency and the object-to-subject dependency can be combined, allowing the latter to bypass the VP barrier. A second problem is that the theory cannot predict that A-movement is subject to antecedent government, since the trace of A-movement typically is θ-governed. This fact must therefore be stipulated, as in (67).

(67) If α is an A-trace, α is antecedent-governed.

We are now in a position to draw some preliminary conclusions. Let us consider first the empirical predictions of the theory, limiting ourselves to two sets of data discussed by Chomsky (1986a). First, it is evident that the theory faces empirical difficulties in the area of extractions across Ns. Nothing in the theory predicts differences between objects of Ns and objects of Vs. Thus, if complex NP islands like those in (32)–(33) are characterized by a CP in the object position of N, the CP is predicted to be L-marked, exactly like an object of V, and hence not to be a barrier. As a result, such constructions are not expected to violate Subjacency. Second, the theory does not account for complex *wh*-island behaviors. In particular, the interaction of *wh*-islands with Tense as in (35) can be captured only by stipulation, essentially by assuming that a tensed IP is an inherent barrier. Similarly, the ill-formedness of multiple *wh*-island violations, as in (37), must be attributed to an enrichment of the mechanism that computes Subjacency violations, such that barriers crossed at different stages of the derivation can be cumulated. These and other problems will be considered in detail in chapter 2. For the time being, all the island behaviors predicted by the theory, either in a principled way or by stipulation, are summarized in table 1.1.

Table 1.1
Island predictions made by Chomsky's (1986a) theory

		θ-governed $\bar{A}$-trace	Non-θ-governed $\bar{A}$-trace/A-trace
Subjacency			
	Subject island	*	*
	Adjunct island	*	*
	Relative island	*	*
	Tense island	*	*
	Multiple *wh*-islands	*	*
ECP			
	Wh-island	OK	*
	That-t	OK	*

Considering next the conceptual framework that the theory offers for the locality problem, we find that it requires at least three stipulations. The first concerns the systematically exceptional behavior of the projections of I built into definitions (45) and (62) and condition (48). The second is that A-movement is always subject to antecedent government, though its traces are always θ-governed, as in (67); this must be added to the stipulation inherited from Lasnik and Saito (1984) that traces in $\bar{A}$-position are subject to antecedent government only at LF, as in (64). A third concerns the role of minimality barriers, which are stipulated to be relevant for government but not for Subjacency, as in (63).

However, notwithstanding these drawbacks, the theory has a conceptually important advantage. Within a barrier-based approach to locality it is at least in principle possible to think of all locality conditions on movement as being based on a single notion of locality domain, or barrier. Thus, although the disjunction between Subjacency and the ECP survives, as well as the disjunction between θ-government and antecedent government within the ECP, the theory appears to incorporate no principled obstacle to unification. This of course is crucial for our discussion.

1.3 Antecedent-Based Solutions

Antecedent-based solutions to the locality problem have also been proposed. This type of solution is represented by the Generalized Binding theory of Aoun (1985), which makes crucial use of a rigid notion of antecedent, or SUBJECT, and by Rizzi's (1990) Relativized Minimality, which makes crucial use of a relative notion of antecedent. Consider the latter first. Relativized Minimality is essentially a theory of the ECP component of locality. According to Rizzi (1990), the ECP reduces to a proper head government requirement, as in (68). Proper head government is government and c-command by a head, as in (69), where β c-commands α only if the first branching node that dominates β dominates α (Reinhart 1976). Adopting Chomsky's (1986a) terminology, β is then said to m-command α only if the first maximal projection that dominates β dominates α.

(68) *ECP*
A trace must be properly head-governed.

(69) X properly head-governs Y iff X head-governs and c-commands Y.

Both head government and antecedent government are crucially constrained by Relativized Minimality, as in (70).

(70) *Relativized Minimality*
X α-governs Y (α = head or antecedent) only if there is no Z such that
a. Z is a typical potential α-governor for Y, and
b. Z c-commands Y and does not c-command X.

The notion of typical potential α-governor that enters into (70), where α ranges over head and antecedent, is in turn defined as in (71) and (72) for head government and antecedent government, respectively.

(71) Z is a typical potential head governor for Y iff Z is a head m-commanding Y.

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

$$\text{Y in an} \begin{Bmatrix} \text{A-chain iff Z is an A-Spec} \\ \bar{\text{A}}\text{-chain iff Z is an } \bar{\text{A}}\text{-Spec} \\ \text{head chain iff Z is a head} \end{Bmatrix} \text{c-commanding Y.}$$

Antecedent government now becomes a requirement on chains. This is not to say, however, that antecedent government becomes a requirement on movement in general. The crucial assumption is that referential indices are licensed by the assignment of a referential θ-role, as in (73); all and only those elements that are assigned a referential index can then form a binding dependency, as in (74).

(73) A referential index must be licensed by a referential θ-role.

(74) X binds Y iff X c-commands Y, and X and Y have the same referential index.

Elements that are *not* assigned a referential index must indeed form a chain, where a chain is partially defined in terms of antecedent government, as in (75). Antecedent government itself is defined as in (76).

(75) $(\alpha_1, \ldots, \alpha_n)$ is a chain only if for all i, α_i antecedent-governs α_{i+1}.

(76) X antecedent-governs Y iff X and Y are nondistinct, X c-commands Y, and Relativized Minimality is satisfied.

Let us consider how this theory works. Consider first *wh*-island violations, which constitute the crucial test for the disjunction between θ-government and antecedent government. Relevant examples are again (20) and (21), repeated here.

(20) What$_i$ do you wonder [how$_j$ to fix t_i t_j]

(21) *How$_i$ do you wonder [what$_j$ to fix t_j t_i]

In Rizzi's (1990) terms, the object *wh*-phrase is assigned a referential θ-role in (20), which licenses a referential index under (73). Hence, the *wh*-phrase can form a dependency through binding, as in (74), which must satisfy head government, as in (68), but not antecedent government. Of course, head government is satisfied, the head governor being V, and (20) is predicted to be well formed. Now consider (21). The adjunct *wh*-phrase is not assigned a referential index under (73) because it is not assigned a referential θ-role. This means that it can only form a dependency if it forms a chain in the technical sense of the term, as in (75). This in turn means that it must satisfy antecedent government, as defined in (76). Now suppose that the adjunct in (21) is generated under the embedded IP, like t_i in (55). In the presence of a *wh*-phrase in the embedded Spec of CP, its next possible landing site is adjoined to the matrix VP, t'_i in (55). Movement from t_i to t'_i is blocked by Relativized Minimality precisely because a typical potential antecedent governor intervenes between t_i and t'_i in the form of the *wh*-phrase *what*$_j$ in the Spec of CP. Hence, the ungrammaticality of (21) is again correctly predicted.

This theory predicts that subjects will pattern with objects rather than with adjuncts with respect to extraction, since they will typically be assigned a referential index. Indeed, it is possible to argue that *wh*-island effects with subjects differ substantially from *wh*-island effects with adjuncts. In particular, adjunct extraction out of a *wh*-island produces a violation at whichever point in the extraction path the *wh*-island is found. The violation in (21) is therefore parallel to the violation in (77), where the adjunct extraction site and the *wh*-island belong to two different sentences.

(77) *How$_i$ do you wonder [whether to believe [he is fixing it t$_i$]]

However, if the *wh*-island is not immediately adjacent to the subject extraction site, no (strong) violation ensues, as in (56), repeated here, versus (78).

(56) *Who$_i$ do you wonder [what$_j$ t$_i$ painted t$_j$]

(78) Who$_i$ do you wonder [whether to believe [t$_i$ is painting it]]

All theories under consideration correctly predict the behavior of the adjunct in (77). Differences arise, however, with respect to (56) and (78). Chomsky's (1986a) account of (56) in terms of antecedent government predicts that (78) is as unacceptable as (77). Given that the subject being extracted has a referential index and is not subject to antecedent government, Rizzi's (1990) theory predicts the well-formedness of (78), leaving the ungrammaticality of (56) to be accounted for. If adjacency to the subject extraction site is in fact crucial to *wh*-island violations with subjects, then these are essentially similar to *that-t* violations, as in (57)–(58), repeated here.

(57) Who$_i$ do you believe [t$_i$ is a painter]

(58) *Who$_i$ do you believe [that t$_i$ is a painter]

Rizzi (1990) indeed proposes a unified solution to *that-t* and *wh-t* violations. Remember that traces must be properly head-governed under (68). But consider the configuration in (61). The embedded I, *is*, that head-governs the subject, t_i, does not c-command it in Reinhart's (1976) sense, and hence does not properly head-govern it under (68). However, in order for this trace to be properly head-governed, it is sufficient that the embedded I and C agree, so that the subject is properly head-governed from the C position. Indeed, Rizzi (1990) takes the apparent absence of an embedded C in the well-formed (57) to be just a reflex of I-C agreement; quite simply, he assumes the zero form of C to be the agreeing form in English. The ill-formedness of (58) is then also explained, on the assumption *that* that is a nonagreeing form of C. *Wh-t* violations can also be viewed as a reflex of I-C agreement, since if I and C agree, the Spec of CP and the Spec of IP must also agree, and the Spec of CP cannot be filled by a *wh*-phrase that disagrees with the subject in the Spec of IP.

These subject extraction phenomena constitute one of three empirical areas where Rizzi's (1990) theory makes new predictions. The second involves the Relativized Minimality constraint in (70) itself. We have seen

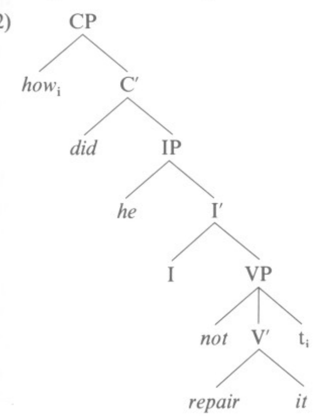
how this constraint accounts for *wh*-island violations with adjuncts. According to Rizzi, there are at least two other types of *wh*-islands: inner islands, created by a negation (Ross 1984), and pseudo-opacity islands, created by a floating quantifier (Obenauer 1976, 1984). Inner islands produce a violation with adjuncts, as in (79), but not with arguments, as in (80); for pseudo-opacity islands also produce a violation with adjuncts, as exemplified by the French construction in (81).

(79) What$_i$ didn't he repair t$_i$

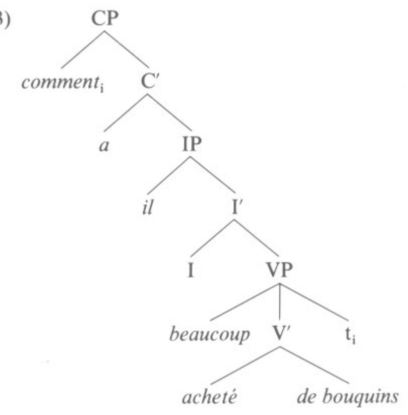
(80) *How$_i$ didn't he repair it t$_i$

(81) *Comment$_i$ a-t-il beaucoup$_j$ acheté [t$_j$ de bouquins] t$_i$
how did he a lot buy of books

Relativized Minimality easily accounts for both inner and pseudo-opacity islands, given the assumption that the elements that create them (the negative *not* and floating quantifiers) occupy an $\bar{A}$-Spec position, possibly Spec of VP. Now consider the structures in (82)–(83), corresponding respectively to (79) and (81). If we assume that the adjunct in (79) and (81) is generated under VP, its movement from a VP-internal position to the Spec of CP is blocked under (76) by a potential antecedent—namely, the negation or floating quantifier in the Spec of VP. Thus, Rizzi's Relativized Minimality theory is able to account for (79) and (81); Chomsky's barriers theory is not.

(82)


(83)



The third empirical domain on which Rizzi (1990) concentrates is related to the referential/nonreferential distinction that he draws, which replaces the θ-governed/non-θ-governed distinction proposed by Chomsky (1986a). Rizzi adduces three arguments to support his distinction. First, he argues that adverbs that obligatorily cooccur with certain Vs are θ-marked by them, and hence θ-governed; however, under extraction they behave exactly like adjuncts, causing *wh*-islands violations, as in (84). He suggests that the same is true of idiom chunk NPs, as in (85), and of measure phrases, as in (86).

(84) *[How well]$_i$ do you wonder [whether he worded his letter t$_i$]

(85) *[What headway]$_i$ do you wonder [whether he has made t$_i$]

(86) *[How many kilos]$_i$ do you wonder [whether he weighs t$_i$]

The shift from the notion of θ-government to the notion of referential index allows the ungrammaticality of (84)–(86) to be correctly predicted. Indeed, none of the *wh*-phrases involved has a referential index, since though each of them is θ-marked, it can be assumed not to be referentially θ-marked, as required by the licensing condition on referential indices in (73). This aspect of the theory is explored in more detail by Cinque (1991), who proposes that it should be extended to the contrast between (87) and (88), where an existential and a universal quantifier, respectively, are extracted from a *wh*-island, and to the contrast between (89) and (90), where a D(iscourse)-linked and a non-D-linked *wh*-phrase, respectively, in the sense of Pesetsky (1987), are extracted from a *wh*-island. Cinque proposes that these contrasts can be accounted for, if existential quantifiers and D-linked *wh*-phrases have a referential index, whereas universal quantifiers and non-D-linked *wh*-phrases do not.

(87) [Some museums]$_i$ I wonder [whether to visit t$_i$]

(88) *[Every museum]$_i$ I wonder [whether to visit t$_i$]

(89) [Which car]$_i$ did you wonder [how$_j$ to repair t$_i$ t$_j$]

(90) *[What the hell]$_i$ did you wonder [how$_j$ to repair t$_i$ t$_j$]

Rizzi's theory also predicts that head dependencies are subject to antecedent government, on the assumption that heads are not associated with referential indices. As for A-dependencies, it appears to be impossible to deny them referential indices, since A-movement typically involves referential phrases. However, Rizzi observes that the θ-Criterion of Chomsky (1981) applies to chains. If so, then because A-dependencies must satisfy the θ-Criterion, they must also conform to the chain format, independently of whether or not referential indices are involved, and they therefore must satisfy antecedent government.

Before we attempt to evaluate Rizzi's theory, let us consider the second antecedent-based solution to the locality problem, proposed by Aoun (1985). Aoun argues that the ECP can be eliminated in favor of binding theory. Thus, Chomsky's (1981) binding conditions, concerned entirely with A-binding, are extended to X-binding, as in (91), where X-binding ranges over both A- and $\bar{A}$-binding. The definition of governing category relevant for (91) is essentially preserved from Chomsky (1981), as in (92). However, the definition of accessible SUBJECT is generalized as in (93), so that accessibility fails when coindexing violates any grammatical principle, not just Chomsky's (1981) *i*-within-*i* Condition.

(91) *Generalized Binding Conditions*
A. An anaphor must be X-bound in its governing category.
B. A pronominal must be X-free in its governing category.
C. A name must be A-free.

(92) β is a governing category for α iff β is the minimal category containing α, a governor for α, and a SUBJECT accessible to α.

(93) β is accessible to α iff β c-commands α and coindexing of α and β does not violate any grammatical principles.

Consider then an object $\bar{A}$-trace, as in the *wh*-island configuration in (20). As a variable, t_i in (20) must be A-free under Generalized Binding

Condition C in (91). On the other hand, since all traces are assumed to be anaphoric, Generalized Binding Condition A in (91) also applies to t_i in (20), requiring it to be X-bound in its governing category. But according to (92), what is the governing category of t_i? No subject can be an accessible SUBJECT for t_i under (93), since coindexing between it and a subject would make it A-bound and hence lead to a violation of Generalized Binding Condition C. Furthermore, no Agr can be an accessible SUBJECT for t_i, since every Agr is coindexed with a subject, and coindexing between t_i and a subject is excluded as before. However, it is assumed that in the absence of any governing category, as defined in (92), the root sentence counts as a governing category for an anaphor, as in (94).

(94) A root sentence is a governing category for an anaphor.

If so, then Generalized Binding for an object Ā-trace, such as t_i in (20), reduces to the requirement that the trace must be bound within the root sentence. This correctly predicts that, as in (20), Ā-movement can take place in one step, bypassing any *wh*-islands on the way.

Next consider subject Ā-traces, such as t_i in (56). Crucially, Aoun (1985) assumes an S/S′ structure of the traditional type. For his results to be replicated within current theories of X-bar structure, it is necessary to assume (much the same way as Chomsky (1986a)) that IP is defective, so that β in the definition of governing category in (92) is other than IP. Again, under (91) t_i in (56) must be bound in its governing category. Furthermore, under (93) the Agr with which t_i is coindexed counts as an accessible SUBJECT for it. Hence, under (92) the governing category for t_i is the embedded CP, IP being defective by stipulation. Within CP t_i can be Ā-bound from the Spec of CP position. But if t_i is not bound from this position, a violation of Generalized Binding Condition A ensues. Thus, Generalized Binding Condition A predicts the antecedent government effects for subjects that are indeed found in (56).

Finally, consider adjunct Ā-traces, such as the one in the *wh*-island configuration in (21). The crucial aspect of Aoun's theory in this respect is that t_i in (21) does not count as a variable, because it is not referential. Thus, t_i is subject only to Generalized Binding Condition A. This in turn means that even if IP is defective, the embedded CP is a governing category for t_i under (92), since coindexing of the embedded subject or Agr with it does not produce any violation under (93). The antecedent government properties of adjunct Ā-traces, specifically their inability to bypass *wh*-islands, are then correctly predicted by Generalized Binding Condition A. Of course, this condition also derives the antecedent government properties of A-movement and of A-binding of lexical anaphors and pronouns. Similarly, the antecedent government properties of head movement can be derived if, following Koopman (1984), head-to-head movement is viewed as an instance of A- rather than Ā-movement.

One empirical domain with respect to which Aoun's (1985) theory differs from both Rizzi's (1990) and Chomsky's (1986a) is extraction from NPs. Remember that it is possible to extract an argument across a single nominal boundary, as in (32), but it is never possible to extract an adjunct across it, as in (34).

(32) Who$_i$ did you see [many portraits of t_i]

(34) *[With what kind of sleeves]$_i$ did you see [many sweaters t_i]

First consider (34). On the assumption that whatever fills the Spec of NP position, including Det as in (34), counts as the SUBJECT of NP, understood as the most prominent element in the phrase, the ungrammaticality of adjunct extraction from NP is immediately accounted for. NP indeed counts as a governing category for the adjunct in that it contains a SUBJECT accessible to it, and the crossing of NP violates Generalized Binding Condition A.

Next consider (32). The argument that undergoes extraction can be interpreted as either the subject or the object of N. If it is interpreted as the subject of N, then NP is not a governing category for it, on the assumption that there is no other accessible SUBJECT in NP. If on the other hand it is interpreted as the object of N, it can still move to the SUBJECT position first, so that its ability to be extracted follows as before. Such a theory predicts that if the SUBJECT position of an NP is independently filled, the object cannot be extracted. Crucially, Aoun (1985) claims that examples like (95) support this prediction:

(95) *Who$_i$ did you see [my many portraits of t_i]

We are now in a position to draw conclusions from the discussion so far. It seems clear that Rizzi's (1990) theory inherits two important conceptions from Aoun's (1985). The first is the distinction between referential and nonreferential elements, which in both theories distinguishes adjuncts from objects. The second, and the reason why the two theories are grouped together here, is that in both cases antecedent government is based on a notion of accessible, or potential, antecedent. The main difference between the two theories is that Aoun (1985) proposes a rigid notion of antecedent, coinciding with the notion of SUBJECT, whereas Rizzi (1990) relativizes the notion of antecedent to the various types of dependencies.

From an empirical point of view, Aoun's (1985) theory has an advantage with respect to extraction from NPs, but not with respect to any of our other residual problems, such as the interaction of *wh*-islands with Tense. Rizzi's (1990) theory appears to have empirical advantages with respect to inner and pseudo-opacity islands, but has essentially the same problems as Chomsky's (1986a) theory with respect to extraction from NPs, interaction of *wh*-islands with Tense, and so on. In addition, there are islands, which Rizzi (1990) does not consider, that display the typical pattern of *wh*-, inner, and pseudo-opacity islands, in that adjuncts are sensitive to them but arguments are not. In the terminology of Cinque (1991), they include factive islands, created by verbs such as *regret* in (96), and extraposition islands, as in (97).

(96) *How$_i$ do you regret [that I behaved t_i]

(97) *How$_i$ is it a shame [that they behaved t_i]

Not being antecedent-based, these islands are problematic for Relativized Minimality, as much as for any theory considered so far.

The empirical predictions of Rizzi's (1990) theory are summarized in table 1.2. What is crucial for present purposes is the general conceptual structure of Rizzi's (1990) and Aoun's (1985) theories. In both cases, at least one non-antecedent-based condition (Subjacency) is required in addition to the proposed antecedent-based conditions. Hence, antecedent-based theories appear to be fundamentally disjunctive and thus immediately excluded as candidates for a unified theory of locality.

Table 1.2
Island predictions made by Rizzi's (1990) theory

		Indexed traces	Nonindexed traces/chains
Subjacency			
	Subject island	*	*
	Adjunct island	*	*
	Relative island	*	*
Antecedent government			
	Wh-island	OK	*
	Inner island	OK	*
	Pseudo-opacity island	OK	*
Head government			
	That-t	OK (object)/ * (subject)	OK

1.4 Connectedness Solutions

Next we will take up connectedness solutions to the locality problem. These are solutions, such as the Connectedness Condition of Kayne (1983) and the Path Containment Condition of Pesetsky (1982), which are based entirely on the geometry of trees. In a sense, they are the opposite of antecedent-based solutions. First consider Kayne's (1983) proposal. Its crucial construct is the notion of g-projection, defined as in (98). In practice, a governor on the right, rather than on the left, always blocks the formation of g-projections in a language like English where canonical government, in the sense of (99), is left to right.

(98) Y is a g-projection of X iff
  a. Y is a projection of X or of a g-projection of X; or
  b. Y immediately dominates W and Z, where Z is a g-projection of X, and W and Z are in a canonical government configuration.

(99) W and Z are in a canonical government configuration iff V governs NP to its right in the language and W precedes Z; or V governs NP to its left in the language and Z precedes W.

Given the notion of g-projection in (98), a g-projection set of a category α is defined as the set of all g-projections of β, where β is a (head) governor for α, together with α itself, and any node that dominates α and does not dominate β, as in (100). The Connectedness Condition is formulated in terms of g-projection sets as in (101). Roughly speaking, given a set of empty categories $\alpha_1, \ldots, \alpha_n$ with antecedent β in a tree T, under (101) β together with the g-projection sets of $\alpha_1, \ldots, \alpha_n$ must constitute a subtree of T.

(100) G_α is the g-projection set of a category α, where β governs α, iff
  a. all g-projections of β belong to G_α
  b. α belongs to G_α and
  b′. if δ dominates α and does not dominate β, δ belongs to G_α.

(101) *Connectedness Condition*
  Let $\alpha_1, \ldots, \alpha_n$ be a maximal set of empty categories in a tree T such that for some β, all α_i are bound by β. Then β together with the g-projection sets of all α_i must constitute a subtree of T.

First, consider a trace in the object position of a V, as in the *wh*-island configuration in (20), repeated here.

(20) What$_i$ do you wonder [how$_j$ to fix t_i t_j]

(101) requires that the antecedent for t_i (namely, the *wh*-phrase) be contained in a g-projection of the (head) governor for t_i (namely, V). In turn, g-projections as defined in (98)–(99) are sensitive only to the geometry of the tree and are not in any way sensitive to antecedent configurations. Thus, object extraction is indeed predicted not to be sensitive to the typical antecedent-based constraint (namely, *wh*-islands), as in (20).

Second, consider an object being extracted out of a subject island, as in (17), repeated here, whose configuration was illustrated earlier in (50).

(17) *What$_i$ does [explaining t_i] bother you

Since the object t_i is governed by V, the crucial question is whether there is a g-projection of V that contains the antecedent for it, namely, the *wh*-phrase. The embedded VP is a g-projection of V; the embedded IP is a g-projection of VP because I canonically governs VP; and the embedded CP is a g-projection of IP because C canonically governs IP. Thus, the embedded CP in (50) is a g-projection of V. However, the embedded CP in (50), being in subject position, is governed by the matrix I, but not canonically governed by it. This means that the matrix IP is not a g-projection of the embedded V; hence, the antecedent of t_i, the *wh*-phrase in the matrix Spec of CP, is not contained within a g-projection of its governor V. If so, Connectedness is violated; in other words, Connectedness correctly accounts for subject islands like the one in (17).

Third, consider an object being extracted out of an adjunct island, as in (18), repeated here, whose configuration was shown earlier in (51).

(18) *What$_i$ was Mary bothered [because Peter explained t_i]

Kayne's (1983) theory differs from all the ones considered so far in predicting such an extraction to be well formed. Indeed, the embedded CP in (51) is a g-projection of the V that governs t_i, very much like the embedded CP in (50). Unlike the subject CP in (50), however, the adjunct CP in (51) is in a canonical government configuration with its sister V′. Hence, the g-projection of the embedded V can be extended to the matrix VP in (51) and from there to the matrix CP, which includes the antecedent for t_i, namely, the *wh*-phrase. Examples like (18) are thus predicted to be well formed. Such a prediction is only apparently incorrect. Remember our discussion of the interaction of Tense with *wh*-islands, as in (35). If Tense indeed creates island effects, the presence of Tense in the adjunct in (18) is in itself sufficient to define a violation. We then expect extraction from an adjunct to be well formed if the adjunct is untensed. The prediction is arguably correct, given the relative well-formedness of examples like (102).

(102) What$_i$ did Peter leave [after explaining t_i]

Indeed, the contrast between (18) and (102) strengthens the Tense island problem for all theories under consideration, Kayne's (1983) included; moreover, (102) supports Connectedness over Chomsky's (1986a) Subjacency.

From the point of view of attempting to unify locality conditions, Connectedness obviously presents a fundamental problem. Let us again consider an object of V. When this object enters into an A-dependency, Connectedness may be a necessary condition on it, but it is certainly not a sufficient one, since the dependency must be strictly local even in the presence of an unbounded g-projection. A completely separate theory, not based uniquely on tree geometry, will then have to be invoked. Thus, Connectedness appears to be no more useful as a basis for a unified theory of locality than antecedent-based accounts, though for exactly opposite reasons.

Of course, saying that a principle based purely on tree geometry is not sufficient to encompass the whole of locality is not the same as showing that such a principle is unnecessary. In particular, apart from apparent exceptions to adjunct island violations like (102), independent evidence for Connectedness comes from parasitic gaps in the sense of Taraldsen 1981, Chomsky 1982, and much related work. In the parasitic gap phenomenon a single operator appears to bind two different variables. Kayne (1983) argues that Connectedness correctly predicts the contrast between (relatively) well-formed examples like (103)–(104) and ill-formed examples like (105).

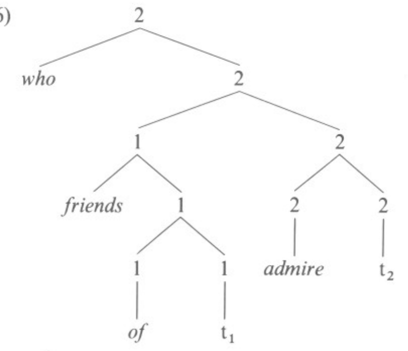
(103) A person who$_i$ [close friends of t_i] admire t_i

(104) A person who$_i$ you admire t_i [because you know [close friends of t_i]]

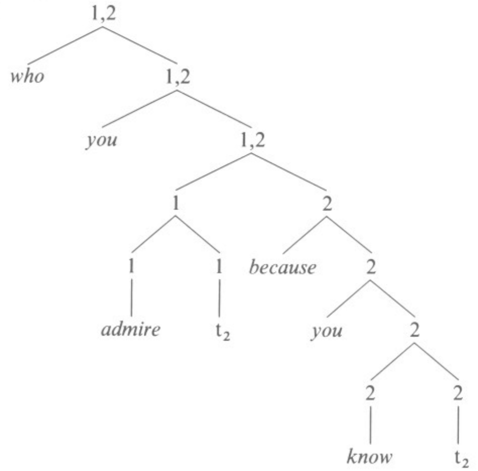
(105) *A person who$_i$ you admire t_i [because [close friends of t_i] become famous]

The relevant structures for (103)–(104) are (106)–(107), where the nodes marked 1 belong to the g-projection set of t_1 and the nodes marked 2 belong to the g-projection set of t_2.

(106)



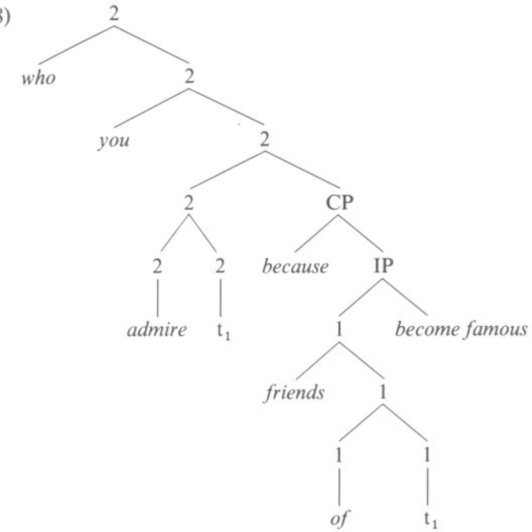
(107)



In (106) the t_1 g-projection set does not by itself form a subtree with the *wh*-phrase antecedent; but the t_2 g-projection set allows it to do so. Thus, descriptively, the potential subject island violation by t_1 is rescued by t_2. Notice that the g-projection set of t_2 is able to bypass the adjunct island in (107) without any assistance from the g-projection set of t_1. Indeed, we have seen that Connectedness, though predicting subject islands, does not predict adjunct islands.

By contrast with (106), consider (108), the relevant structure for (105).

(108)

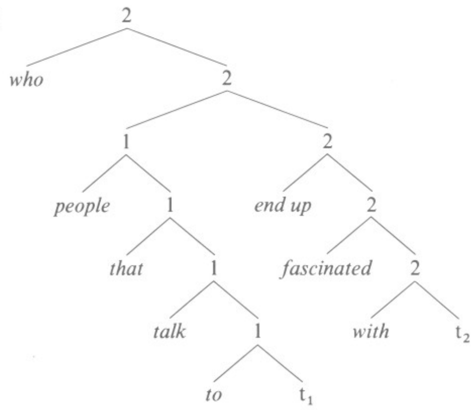


In (108) the g-projection set of t_1 stops at the subject node, as in (106), but contrary to (106), it forms a subtree distinct from the subtree formed by the g-projection set of t_2. Thus, in (108), contrary to (106), t_2 cannot rescue t_1 from its subject island violation.

Recall that under Chomsky's (1986a) Subjacency, relative clause islands are essentially a kind of adjunct islands, since relative clauses are attached in adjunct-like position within NP, as in (52). By contrast, Kayne's (1983) theory predicts that relative clauses (just like adjuncts) will not give rise to Connectedness violations, since they enter canonical government configurations (i.e., they are attached to the right in English). Thus, the parallelism with adjuncts is maintained, but the predictions are reversed. Examples like (109), with the structure in (110), indeed appear to support Connectedness. The g-projection set of t_1 bypasses the relative clause (island) in (110) without assistance from the g-projection set of t_2; the latter is then instrumental in rescuing the potential subject island violation.

(109) A person who$_i$ [people that talk to t_i] usually end up fascinated with t_i

(110)



Although relative clause island violations like (19), repeated here, are clearly ungrammatical, this can be attributed to the same Tense effect that causes the ungrammaticality of adjunct island violations like (18). If so, extraction from untensed relative clauses is expected to be grammatical, as in (111) and (112), where Italian differs from English in allowing overt *wh*-phrases in infinitival relatives.

(18) *What$_i$ was Mary bothered [because Peter explained t$_i$]

(19) *What$_i$ do you know [the girl [that explained t$_i$]

(111) *The student that $_i$ I know [a book [O$_j$ to give t$_j$ to t$_i$]]

(112) *Il libro che$_i$ conosco [uno studente [a cui$_j$ dare t$_i$ t$_j$]]
the book that I know a student to whom to give

The fact that (111)–(112) are *not* grammatical can be attributed to a complex NP island effect. Indeed, (111)–(112) appear to have the same status as examples like (33), with extraction from a CP object of N. Connectedness then brings to the fore the problem of complex NP islands, as well as the problem of Tense islands in (18) versus (102).

The predictions that Connectedness makes for simple (nonparasitic) extractions are summarized in table 1.3. The table mentions argument traces only, because Kayne (1983) does not discuss the extraction of adjuncts. Thus, this theory leaves open the question of any differences between the behavior of adjuncts and the behavior of arguments.

A theory conceptually close to Kayne's (1983) is Pesetsky's (1982), which is based on a notion related to g-projection, namely, the notion

Table 1.3
Island predictions made by Kayne's (1983) theory

		Argument traces
Connectedness		
	Subject island	*
	Adjunct island	OK

"path." A path is technically defined in (113) as the set of all nodes that dominate a trace up to the first maximal projection that also dominates its antecedent.

(113) If t is locally $\bar{A}$-bound by α, β is the first nonlexical node dominating t, and β' is the first maximal projection dominating α, then the path between t and α is the set of nodes P including β, β' and all and only γ such that γ dominates β and γ does not dominate β'.

The major condition on paths is the Path Containment Condition (115), stated in terms of overlapping paths (114). In effect, (115) allows nested dependencies and disallows crossing ones.

(114) Two paths overlap iff their intersection is nonnull and nonsingleton.

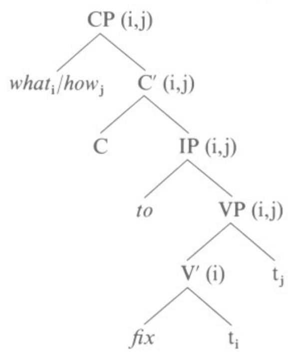
(115) *Path Containment Condition (PCC)*
If two paths overlap, one must contain the other.

Consider the standard contrast between adjunct extraction and object extraction from *wh*-islands, as in the familiar examples (20) and (21), repeated here along with the relevant structure (116).

(20) What$_i$ do you wonder [how$_j$ to fix t$_i$ t$_j$]

(21) *How$_i$ do you wonder [what$_j$ to fix t$_j$ t$_i$]

(116)



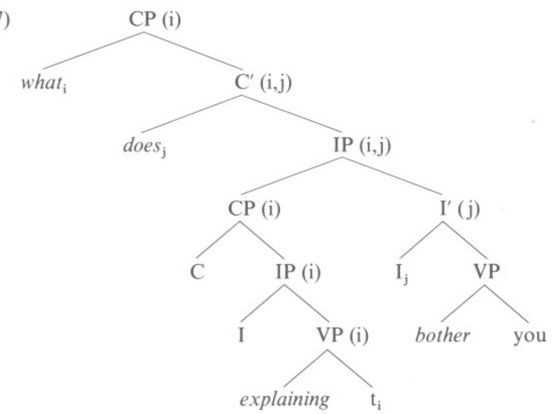
Pesetsky's (1982) theory treats this contrast as a PCC effect. In (116) the path created by object extraction includes the embedded V′, VP, and so on, as indicated; even assuming that the adjunct is attached as a sister to V, the path created by adjunct extraction includes the embedded VP and so on, but not V′. The two extraction paths will of course overlap, since they have common members such as VP. Suppose then that the object moves to a higher position than the adjunct, as in (20). If so, the path of the object completely contains the path of the adjunct, since it extends farther downward with respect to V′, as well as farther upward by hypothesis. Thus, the PCC is not violated. But suppose that the adjunct moves higher than the object, as in (21). Then by hypothesis the path of the adjunct extends farther upward than the path of the object; but the former, which does not include V′, cannot completely contain the latter, which does. Thus, the PCC is violated.

Next consider subject island violations, as in (17).

(17) *What$_i$ does [explaining t$_i$] bother you

Prima facie, the PCC does not make any predictions concerning such violations since a single dependency is involved, rather than two interacting dependencies. However, Pesetsky (1982) postulates that a dependency is systematically established between the I and C positions of any given sentence. This means that the relevant structure for a subject island violation is as in (117).

(117)



The path between the matrix I and C, which are involved in overt head-to-head movement in (117), contains I′, IP, and C′. The path for the extraction from the subject island never includes I′, but it does include a number of nodes, such as the embedded CP, which are not included in the I-to-C path. Thus, extraction from a subject inevitably violates the PCC.

It is evident that the PCC has the same basic problem as Connectedness with respect to unifying locality theory. Both of these accounts attempt to make use only of the pure geometry of trees. From the unification point of view, this approach has the fundamental fault that it must be supplemented by some nongeometrical approach, presumably an antecedent-based one, to account for typical antecedent government dependencies such as A-movement. On the other hand, to the extent that the PCC, unlike Connectedness, takes the interaction of different paths into account, it appears to be redundant with any antecedent-based constraint. If so, the PCC can be argued to have an additional conceptual disadvantage with respect to Connectedness.

Of course, showing that the PCC is insufficient is entirely different from showing that it is unnecessary. There is a particular series of contrasts, which look prima facie like crossing versus nesting contrasts, that the PCC accounts for, but other locality conditions do not. Thus, if a head has both an object and a prepositional object, it appears that the prepositional object can be extracted across the object, but not vice versa, as in (118)–(119).

(118) [Which violins]$_i$ are these sonatas easy [O$_j$ [to play t$_j$ on t$_i$]]

(119) *[Which sonatas]$_i$ are these violins easy [O$_j$ [to play t$_i$ on t$_j$]]

In terms of the PCC, the prepositional object path contains at least one node, PP, which the direct object path does not contain. Thus, if the prepositional object is extracted farther than the direct object, as in (118), the path of the former contains that of the latter, as required by the PCC. On the other hand, if the direct object is extracted farther, as in (119), its path cannot contain that of the prepositional object (and vice versa), and the PCC is violated. Contrasts such as these thus constitute crucial empirical arguments in favor of the PCC.

1.5 The Locality Solution

With the previous sections as background, I would now like to present my own solution to the locality problem, as an informal introduction to the discussion to follow. The fundamental ideas behind this solution are extremely simple and require no special technical apparatus. I assume that every lexical item and every other element individuated by lexical features of any sort is assigned what I call a *categorial index*. Roughly, categorial indices are indices of content. In the case of NPs these indices can be identified with the referential indices commonly used to express both anaphoric and movement dependencies. Thus, I am not so much introducing a new primitive, as extending an already existing one. Notice that to the extent that so-called referential indices are standardly used to express movement dependencies involving Vs rather than NPs, they are no longer referential in any identifiable sense of the term. Rizzi (1990) makes the same argument but reaches the opposite conclusion: namely, that referential indices should be maintained as such, and restricted to elements that indeed have referential content.

There is also a second type of index in the grammar, which I call an *address*, borrowing the term and in part the conception from Vergnaud (1985). Roughly speaking, addresses are indices of position and differ from categorial indices in being relational. Thus, in general a position Case-marked by a head is addressed by the head; in particular, we can assume that the address assigned to the Case-marked position corresponds to the categorial index of the Case-marking head. Again, then, addresses do not so much introduce a new primitive in the grammar, as encode an already existing relation, Case assignment by a head.

Each of the five dependencies considered earlier (anaphora as in (1)–(4), A-movement as in (5)–(8), head movement as in (9)–(13), Ā-movement of arguments as in (14)–(19), and Ā-movement of adjuncts as in (24)–(26)) can be established on the basis of a categorial index. An antecedent for a trace, anaphor, or pronoun can be defined simply as an element that shares a categorial index with it. Furthermore, each of the five dependencies is ordered by c-command. Thus, the antecedent, with which the categorial index resides, c-commands the dependent element, with which it shares the index.

Crucially, the five dependencies differ with respect to addresses. Three of them in fact involve a trace without an address. From θ-theory and Case theory, it follows that neither an A-trace nor the trace of a V or I head (or its maximal projection) is ever Case-marked; hence, such traces are never addressed. Whether an adjunct Ā-trace is Case-marked or not, we can assume that it is never Case-marked by a head; hence, it is never addressed either.

By contrast, Ā-traces of arguments typically are addressed, by V if they are assigned accusative Case, and by I if they are assigned nominative

Case. The same is true of lexical anaphors and pronouns. However, the notion of address serves to distinguish argument Ā-traces from lexical anaphors and pronouns. The antecedent of an anaphor or pronoun will be independently Case-marked and hence addressed differently from the referentially dependent element. The antecedent of an argument Ā-trace, typically a *wh*-phrase, will not be independently Case-marked and hence will not be independently addressed either.

Since dependencies can be defined on the basis of categorial indices and since an address is just another type of index, dependencies ought to be definable on the basis of a shared address as well. Let us consider the five dependencies in turn, each of which is well defined as a categorial index dependency. An (antecedent, anaphor) or (antecedent, pronoun) pair cannot in general form an address-based dependency for the simple reason that the two members of the dependency have different addresses; it would be like trying to form a categorial index dependency between two names. In the case of head movement and Ā-movement of adjuncts, movement is from a nonaddressed position to another nonaddressed position. No address-based dependency can be formed because these positions have no address to share. Finally, consider the dependencies corresponding to A-movement and to Ā-movement of arguments. In the former the trace is not addressed, but the antecedent generally is. In the latter the trace is generally addressed, but the antecedent is not. Thus, in the case of an A-chain the address of the antecedent can in principle be transmitted downward to the trace, whereas in the case of an argument Ā-chain the address can be transmitted upward from the trace to the antecedent. Only in Ā-chains is the direction of transmission congruent with the general direction of transmission of features and/or indices under movement; thus, we can assume that argument Ā-chains are well formed both as address-based dependencies and as categorial index dependencies, and that A-chains are not.

To sum up: Ā-dependencies involving arguments are distinguished from all other dependencies, in that they are well formed not only as categorial index dependencies but also as address-based ones. This means that one component of the unification problem for locality theory is solved. Dependencies with strict locality effects differ from other dependencies only in that the latter but not the former are well formed as address-based dependencies.

The formal notions involved are extremely simple. Categorial indices can be associated simply with lexicality, as in (119). Addressing can be defined as Case marking by a head, as in (120), and, adopting the cover term *sequence* for both categorial index and address-based dependencies, a sequence can be defined simply as in (121).

(119) α has a categorial index if α is lexical.

(120) α has an address if there is a head β that Case-marks α.

(121) $(\alpha_1, \ldots, \alpha_n)$ is a sequence iff for all i, α_i is coindexed with and c-commands α_{i+1}.

It should be stressed again that (119)–(121) represent more a reorganization of the existing theory of grammar than an enrichment of it. In particular, (121) needs to be stated anyway in order to define a chain. In the proposed theory, of course, a chain will be defined simply as a sequence, which also satisfies certain other requirements.

We are now in a position to approach the second part of the unification problem. Given that the crucial difference between strictly local dependencies and all others is that only the latter can be construed as address-based dependencies, do their different locality behaviors follow from this single difference, without any disjunction being stipulated between them in the locality theory itself?

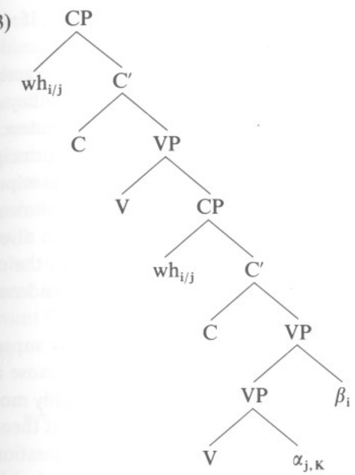
I assume that the locality theory for categorial index sequences reduces to a condition like (122)—that is, essentially the antecedent government clause of the ECP and/or binding theory formulated as a biconditional. In (122) *dependent element* is a cover term for traces and anaphors, the only nondependent elements relevant in this context being pronouns. A sequence is said to satisfy government iff all links of the sequence satisfy it. Government in turn is defined, as in Chomsky 1986a, as holding between two positions just in case no barrier is crossed.

(122) *Locality*
α is a dependent element iff there is an antecedent β for α and a sequence $(\beta, \ldots, \alpha)$ that satisfies government.

In order to prove that (122) represents the locality theory for categorial index sequences, we need to prove that (122) derives Subjacency as well. This is indeed the case, as I will argue in chapter 2, given the appropriate definition of barrier. Right now, however, the crucial issue is to determine the locality theory for address-based sequences. If it is also found to be (122), then the unification problem is essentially solved.

Consider for instance the structure in (123), which is a shortened bisentential structure with an embedded *wh*-island, where each structure contains a lexical V head and a functional C head.

(123)



IP projections are omitted here only in order to avoid technical discussions that are properly left for chapter 2. The embedded clause contains a direct object, α, and an adjunct, where again for simplicity the adjunct is adjoined to VP. Crucially, the adjunct is associated only with a categorial index, but the object is associated with both a categorial index and an address (notated with a capital subscript). The question is whether the single locality principle in (122) correctly predicts that the adjunct cannot move across the *wh*-island, as in (21), but the argument can, as in (20).

First consider the adjunct. This can move only if a well-formed categorial index dependency can be created. Since β_i in (123) is adjoined to VP, the first barrier for it is the embedded CP. But suppose that its next possible landing site, the Spec of CP, is already filled by wh_j. If so, β_i can only move by crossing the CP barrier, and Locality is correctly predicted to be violated, as in (21). This prediction is of course more or less shared by all current theories.

Next consider the direct object. α_j can move if it can enter into either a well-formed categorial index dependency or a well-formed address-based dependency. α_j cannot ultimately enter into a well-formed categorial index dependency for the same reason that β_i cannot, given the presence of another *wh*-phrase in the Spec of the embedded CP. However, consider the address-based alternative. None of the positions in (123) is addressed, except for α_j itself. Under the simplest conception of coaddressing, α_j can then form an address-based sequence with any of the positions in (123), yielding for instance the sequence (wh, C, V, C, V, α) of address K. If so, α_j is correctly predicted to satisfy government, as in (20).

To sum up once again: $\bar{A}$-traces of arguments differ from $\bar{A}$-traces of adjuncts, anaphors, pronouns, A-traces, and head traces, in that they can form address-based dependencies as well as categorial index dependencies. The two types of dependencies are subject to the same locality principle, (122). The different locality effects they give rise to follow without stipulation from their own different properties. This result has been shown to hold for the configuration in (123). But if (122) is correct, it can also be shown to hold for any other configuration. If so, then it is not the case that there are different locality theories for different types of dependencies; rather, there is only one locality theory, (122).

In sketching the proposed theory, I have claimed that it is superior because it eliminates a set of disjunctions—in other words, because it is simpler and/or because the grammar that includes it is more highly modular in nature. But can it also be distinguished from other current theories in terms of its empirical predictions? The answer to such a question is always fraught with difficulties, since most predictions require decisions quite independent of the theories being evaluated. However, three points can be made in this regard. First, most current theories of extraction face difficulties concerning complex NP island violations with sentences in the object position of N, as in (33), repeated here. This is a problem common to all theories of Subjacency based on the CED, including Chomsky's (1986a).

(33) *Who$_i$ did you see [many attempts [to portray t$_i$]]

Second, with the exception of Aoun's (1985) theory, they do not make clear predictions with respect to extractions from simple NPs, specifically with respect to the impossibility of extracting adjuncts, as in (34); nor do they make clear predictions with respect to the impossibility of extracting from definites, as in (36).

(34) *[With what kind of sleeves]$_i$ did you see [many sweaters t$_i$]

(36) *Who$_i$ did you see [the many portraits of t$_i$]

Third, there is no nonstipulative approach to Tense effects and to their interaction with *wh*-islands, as in (35), or to the contrast between double and triple *wh*-extractions, as in (37).

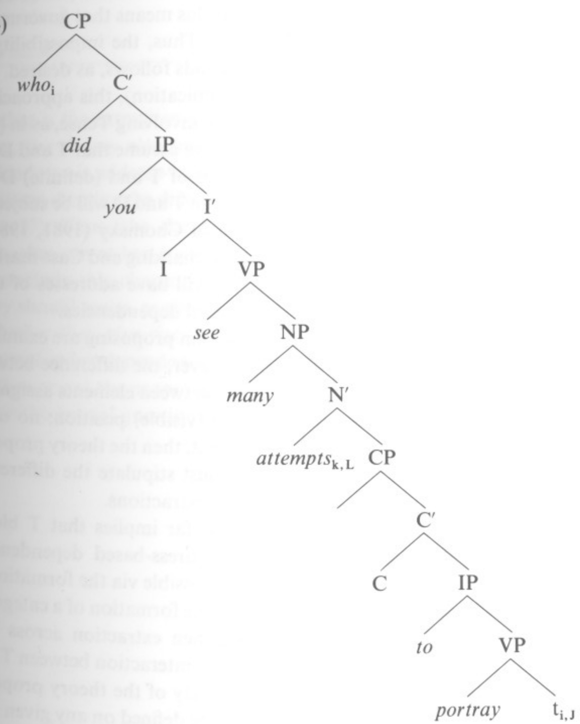
(35) *What$_i$ did you wonder [how$_j$ I repaired t$_i$ t$_j$]

(37) *[Which books]$_i$ did you wonder [to which students]$_j$ to ask [whether to give t$_i$ t$_j$]

And certainly, there is no proposal that draws together all of these residual problems.

Now consider the theory proposed here. There is one difference between Ns and Vs that follows from the basic statement of the theory and does not need to be stipulated: namely, that Vs are never addressed, because they are never visible in the sense of Chomsky (1981, 1986b), whereas we can assume that if an NP is visible and hence addressed, its address percolates to its head N. If so, then the basic difference between Ns and Vs with respect to Locality is that Vs (which have no address of their own) can be included in address-based sequences, whereas Ns (which do in general have an address of their own) cannot. Consider the structure relevant for complex NP island violations like (33), as in (124).

(124)



Suppose first that the *wh*-phrase in (33) moves by forming a categorial index dependency. If so, it can move as high as the embedded Spec of CP in (124), but not higher. Indeed, we can adopt the standard assumption that NP is not associated with any $\bar{A}$-Spec or adjoined position; if so, a categorial index dependency must bypass NP, creating a government violation. Since adjuncts can only form categorial index dependencies, this accounts for the impossibility of extracting adjuncts not only from complex NPs, but also from NPs in general, as in (34). Suppose on the other hand that the argument in (33) moves by forming an address-based dependency. Suppose further that this dependency includes the embedded V, I, and C heads in (124); crucially, it cannot include the N head, *attempts*, since this already has an address of its own. If so, the next head that can enter the dependency is the matrix V; but this means that government is violated, since NP is once more bypassed. Thus, the impossibility of extracting an argument from complex NP islands follows, as desired.

As noted by T. Hoekstra (personal communication), this approach to islands involving N can be generalized to those involving Tense, as in (35), and those involving Determiners, as in (36), if we assume that T and D are heads. In particular, if the denoting properties of T and (definite) D are sufficient to individuate them as arguments, then T and D will be subject to the θ-Criterion and the Visibility Condition of Chomsky (1981, 1986b); hence, some (appropriately abstract form of) θ-marking and Case marking will apply to them. If so, T and (definite) D will have addresses of their own and will therefore also block address-based dependencies.

These and other predictions of the theory I am proposing are examined in detail in chapters 2 and 3. In general, however, the difference between Ns (or Ds or Ts) and Vs is just the difference between elements assigned a (visible) position and elements that assign a (visible) position; no other difference is stated in the theory. If this is correct, then the theory proposed here is superior to its competitors, which must stipulate the differences between V- and N-type heads with respect to extractions.

Furthermore, nothing in the discussion so far implies that T blocks categorial index dependencies as well as address-based dependencies. Hence, extraction across T is predicted to be possible via the formation of a categorial index sequence. But suppose that the formation of a categorial index sequence is blocked by a *wh*-island; then extraction across T is predicted to be impossible. Thus, we derive the interaction between T and *wh*-islands, as in (35). In general, it is a property of the theory proposed here that at most two $\bar{A}$-extraction paths can be defined on any given subtree. One is the extraction path defined by a categorial index dependency; the other is the extraction path defined by an address-based dependency. Hence, the theory also predicts that two extraction paths can overlap, provided one is address-based; but a third extraction path cannot overlap with them, deriving the difference between double and triple *wh*-extractions, as in (37). Thus, the theory once more predicts behaviors that other approaches can at best handle by stipulation.

I conclude this chapter by summarizing the predictions an adequate theory should make, as shown in table 1.4. Islands are grouped according to whether they give rise to violations with all dependencies, only with categorial index dependencies, or only with address-based dependencies. Whether or not the proposed theory does indeed yield the desired predictions is the crucial question that will occupy the rest of this book.

Table 1.4
Island predictions an adequate theory must make

	Categorial index dependency	Address-based dependency
Subject island	*	*
Adjunct island	*	*
Relative clause island	*	*
(Complex) NP island	*	*
Definiteness island	*	*
Overlapping *wh*-islands	*	*
Wh-island	*	OK
Inner island	*	OK
Pseudo-opacity island	*	OK
Factive island	*	OK
Tense island	OK	*