c-commanded by the variable (you can check this in (A4, A5, A9, A13, and A15)). However, pronouns do not have to be c-commanded by their antecedent, as we saw in 3.1.2. So, in principle, we might get a configuration where a pronoun is not c-commanded by a variable that it is coindexed with. This arises in such configurations of 'weak crossover' (which is distinct from the 'strong crossover' discussed in 3.3.2.3), as:

(A16) *Who$_i$ does [his$_i$ mother] love t$_i$?

This example is ungrammatical on the interpretation indicated by the coindexing relations. It should mean 'For which *x*, *x* a person, does *x*'s mother love *x*?' Here, unlike all the other examples we've looked at, the pronoun isn't c-commanded by the trace. Accordingly, we can conclude that *for a pronoun to be interpreted as a bound variable, it must be c-commanded by an A'-bound trace.* In (A15), this requirement isn't fulfilled and the pronoun fails to have an interpretation with the indexing given. Different approaches to weak crossover have been proposed by Wasow (1972), Chomsky (1976) and Koopman and Sportiche (1982). A further argument for the existence of QR is the presence of weak-crossover effects with quantifiers, as in:

(A17) *[His$_i$ mother] loves everyone$_i$

After QR, this example has the LF representation in (A17), which is clearly parallel to (A15) and violates the condition on pronoun-interpretation in the same way:

(A18) Everyone$_i$ [[his$_i$ mother] loves t$_i$]

In conclusion, logical binding corresponds to A'-binding of variables, where the relation may be set up either by covert or overt movement to an A'-position. Anaphoric binding of the sort discussed in this chapter is A-binding. Binding in both cases is c-command combined with coindexation.

One final note: the picture presented above assumes a primitive distinction between quantificational and non-quantificational expressions, as do many kinds of logic. However, we can introduce an operator which makes anything into an expression capable of binding a variable; this is the λ-operator. It has been argued that all DPs should be interpreted this way. So *John left* might have the interpretation:

(A19) John (λ*x* (Left (*x*)))

This is read as 'John is *x* such that *x* left'. As we saw, this notation is exploited in Reinhart and Reuland's account of pronominal anaphora (see 3.5.5).

4

Locality

4.0 Introduction

This is the last chapter that deals with the technical parts of syntactic theory. As in the previous chapter, our theme is the kinds of relations that can hold among syntactic constituents. However, unlike binding relations, here we are dealing exclusively with movement dependencies, and so the theoretical construct that I'll focus on here is Move-α. What we're interested in here is 'how far' movement can go: when Move-α applies to a category α, how far away from α's original position can α be taken? It's clear that, in order to answer this question properly, we'll need a workable measure of 'distance' that can be defined over syntactic structures – figuring out what distance is in syntactic terms, and how much of it Move-α can cover, have been central currents of research in generative grammar. Here I'll give an overview of what's known.

We can set the scene for the discussion by looking at what we've already said about Move-α in earlier chapters. In Section 2.4 of Chapter 2 we looked at the properties of the different kinds of Move-α: head-movement, DP-movement, and *wh*-movement. Let's look at those again:

Head-movement:

(1a) moves a head

(1b) leaves a trace in the starting position

(1c) moves to a position c-commanding the trace

(1d) moves to the closest possible position to the trace

(1e) obeys the Structure Preservation Constraint

DP-movement:

(2a) moves a DP

(2b) leaves a trace in the starting position

(2c) moves to a position c-commanding the trace

(2d) moves to the closest possible position to the trace

(2e) obeys the Structure Preservation Constraint

(2f) only ever moves to Specifiers of functional heads

wh-*movement:*

(3a) moves a *wh*-XP

(3b) leaves a trace in the starting position

(3c) moves to a position c-commanding the trace

(3d) does NOT move to the closest possible position to its trace (or so it appears)

(3e) obeys the Structure Preservation Constraint

(3f) only ever moves to Specifiers of CP

We commented on the similarities between the different kinds of movement in 2.3. Let's see what the common factors of them are:

Move-α:

(4a) moves α = either X or XP

(4b) leaves a trace in the starting position

(4c) moves α to a position c-commanding the trace

(4d) moves α to the closest possible position to the trace (except *wh*-movement)

(4e) obeys the Structure Preservation Constraint

(4f) XP-movement only ever moves to Specifiers of functional heads; *wh*-movement is restricted to movement to SpecCP

Properties (4a, b, c, and e) are straightforward. Property (4a) just states that movement only affects some X'-entities (in fact, you might wonder why X'-level projections are inert for Move-α; one possibility, as we saw in 2.6.4, is that this is because such projections don't exist – we'll leave this aside here, however). Properties (4b, c, and e) all impose substantive constraints on Move-α and, in fact, amount to a little theory of Move-α. In fact, they fall together under the idea that we put forward in Section 3.3.2.2 of Chapter 3 that *a moved element α must bind a trace.* In terms of this idea, we can also make sense of (4f) by saying that *wh*-movement moves α to a non-L-related position, and so involves A'-binding of the trace, while DP-movement moves α to an L-related position and so involves A-binding of the trace, as we saw in 3.3.2.3.

So we have the following theory of movement, to restate what we've already seen in a synoptic form:

Move-α:

(5a) leaves a trace in the starting position

(5b) creates a binding relation between the moved category, α, and its trace (A-binding if α is in an L-related position; A'-binding if α is in a non-L-related position)

In fact, property (5a) of Move-α is a consequence of the Projection Principle (see 2.1). So the really central property that is specific to Move-α is (5b): the creation of a binding relation. So far so good (except that you might wonder about the binding relation between a moved head and its trace – let's just suppose that there can be binding, in the sense of coindexation combined with c-command, between X° elements too).

However, the locality condition in (4d) seems to be radically different for the different kinds of movement. We've seen locality conditions on both head-movement (the Head Movement Constraint, (67) in Chapter 1) and DP-Movement (Principle A of the Binding Theory). Besides Principle A, we also suggested that DP-movement may be required to move to the nearest available position: these two conditions aren't quite the same, as we'll see later. But, as we mentioned in Chapter 2, *wh*-movement doesn't seem to be constrained in the same way. *wh*-dependencies appear to be **unbounded dependencies**, able to hold across an arbitrarily large amount of syntactic material. The examples that we used to illustrate this in Chapter 2 (which were taken from classic work by Ross (1967)) were these:

(6a) What$_i$ did Bill buy t_i ?

(6b) What$_i$ did you force Bill to buy t_i ?

(6c) What$_i$ did Harry say you had forced Bill to buy t_i ?

(6d) What$_i$ was it obvious that Harry said you had forced Bill to buy t_i ?

And it seems intuitively clear that one could, circumstances permitting, interpose an arbitrary amount of material between *what* and its trace in this kind of example. One of the main goals of this chapter is to show that the unbounded nature of *wh*-dependencies is only apparent. We'll see that there is in fact a locality constraint on *wh*-movement, one comparable to the constraints on other kinds of movement. In this way, we'll eventually arrive at a unified theory of Move-α that includes (5) and a general locality condition.

As I said, the principal goal of this chapter is to show how movement theory can be unified by showing how *wh*-movement is, despite initial appearances, subject to a constraint like those which apply to head-movement and DP-movement. In Section 4.1, I'll present the evidence that *wh*-movement is not entirely unbounded – evidence which takes the form of 'island constraints', originally discussed in Ross (1967). In Sections 4.2 and 4.3, I'll present the two central locality principles that have been proposed to account for the locality of *wh*-movement: the Subjacency Principle and the Empty

Category Principle (ECP). Section 4.4 is about the concept of **barriers** (which we introduced in Chapter 2, Section 2.2.2), a first attempt to unify these two principles and integrate them with the rest of the theory. Section 4.5 deals with another kind of unification, namely the development of a theory of locality that applies to DP-movement and head-movement as well as *wh*-movement: Relativized Minimality. Finally, in Section 4.6 I'll present the most recent version of the theory of movement, Manzini's (1992) theory of locality.

4.1 Islands

Here I want to introduce the principal evidence that *wh*-dependencies are subject to constraints. The evidence, which was first systematically discussed and illustrated in Ross (1967) (although it was adumbrated in Chomsky (1964); in any case I'm going to take a number of historical liberties in my presentation), takes the form of a class of phenomena known as **islands** or **island constraints**. The terminology is metaphorical: an island is something that it's difficult to move from – at the very least you need some special means of transportation and, in fact, you can get completely stuck on one. Hence syntactic islands are those structures out of which *wh*-elements can be moved only with difficulty, if at all. It's important to see that, while island phenomena clearly show that *wh*-movement cannot apply just anywhere, they do not show us that *wh*-movement is *never* unbounded. What they show is that *wh*-movement is at least *sometimes* bounded.

As I said above, Ross discovered and illustrated many island phenomena. In the rest of this section I'll discuss the principal island phenomena that were originally identified by Ross, although a number of other constraints have been discovered in more recent work that I'll introduce as we go along. Ross instituted a terminological practice that has been largely followed ever since, that of calling each island 'The X Constraint' or 'The X Condition', where 'X' usually designates the particular structural configuration involved. Since it has at least mnemonic value, I'll follow this practice too.

The island constraints that Ross discussed are as follows:

The Complex NP Constraint (CNPC):
Ross (1986: 76) formulates this constraint as follows:

(7) No element contained in a sentence dominated by a noun phrase ... may be moved out of that noun phrase.

Putting (7) in terms of the assumptions about functional categories that we've made in earlier chapters, the CNPC prevents extraction (this term is often used for 'moving out', and I'll adopt this usage from now on) of α from a configuration like (8):

(8) $[_{DP} \ldots \quad [_{IP} \ldots \alpha \ldots]]$

The CNPC accounts for two main classes of fact: the impossibility of extraction from relative clauses, and the impossibility of extraction from the sentential complements of Nouns like *claim*, *fact*, and *story*. These facts are illustrated in (9). Here I've indicated the island in small capitals, with a trace in the position of α:

(9a) *Which band$_i$ did you write $[_{DP}$ A SONG WHICH $[_{IP}$ WAS ABOUT t_i]] ?

(9b) *Which band$_i$ did you believe $[_{DP}$ THE CLAIM THAT $[_{IP}$ WE HAD SEEN t_i]] ?

In (9a), the DP-island is a relative clause; in (9b) it's a DP containing a Noun and its sentential complement. In some respects, the two instances of the CNPC are distinct; for example, many speakers find sentences like (9a) worse than (9b).

Sentence (9a) should also be compared with (10):

(10) Which band$_i$ did you write $[_{DP}$ A SONG ABOUT t_i]] ?

Most people find (10) perfectly acceptable. If there is no 'reduced clause' in the complement of *song*, this fact is quite consistent with the CNPC as stated in (7) and (8). However, it is worth pointing out that the definiteness of the DP in (10) is important. The sentence deteriorates if the DP is definite, and is quite bad if there is a possessor DP in its Spec:

(11a) ??Which band$_i$ did you write $[_{DP}$ THAT SONG ABOUT t_i]] ?

(11b) *Which band$_i$ did you sing $[_{DP}$ MICK'S SONG ABOUT t_i]] ?

This fact was observed by Fiengo and Higginbotham (1981). Manzini (1992) calls this kind of case a **definiteness** island, since, as you can see, the blockage to extraction seems to be created by a definite D (remember that possessive DPs like *Mick's song* are always definite).

The Subject Condition:
Ross noticed that extraction from sentential subjects, as in (12a), was not allowed. Chomsky (1973) extended this to all cases of complex subjects. The generalization is that extraction from inside a subject is bad:

(12a) *Which rock star$_i$ was $[_{CP}$ THAT THE POLICE WOULD ARREST t_i] expected ?

(12b) ??Which rock star$_i$ were $[_{DP}$ ADMIRERS OF t_i] arrested ?

It's important to see that the Subject Condition bans extraction from within a subject, not extraction of a whole subject.

The Coordinate Structure Constraint (CSC):

(13) In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct. (Ross (1986: 99))

This means that no α can be extracted in structures like (14) (where '&' represents any kind of coordination):

(14) $[_{\alpha} \ldots \alpha \ldots] \ \& \ [_{\alpha} \ldots \alpha \ldots]$

This rules out examples like the following (again, the island configuration is given in small capitals):

(15a) *What$_i$ did Bill buy POTATOES AND t$_i$?

(15b) *What$_i$ did Bill buy t$_i$ AND POTATOES ?

(15c) *Which guitar$_i$ does KEITH [PLAY t$_i$] AND [SING MADRIGALS] ?

(15d) *Which madrigals$_i$ does KEITH [PLAY THE GUITAR] AND [SING t$_i$] ?

An important proviso to the CSC is that *wh*-movement can apply in coordinate structures as long as it applies **across the board**, in that identical elements are affected in each conjunct. We can see this if we compare the following examples of relative-clause formation (another instance of *wh*-movement):

(16a) Students$_i$ [WHO$_i$ t$_i$ FAIL THE FINAL EXAM] OR [WHO$_i$ t$_i$ DO NOT DO THE READING] will be executed (Ross (1986: 109))

(16b) *This is the student$_i$ [WHO$_i$ t$_i$ FAILED THE FINAL EXAM] AND [JOHN DID THE READING]

The CSC has to a large extent resisted satisfactory theoretical treatment in the principles-and-parameters framework.

The Left Branch Condition (LBC):

This condition prevents extraction of α in the following configuration, where X is any non-null material:

(17) $[_{DP} \ \alpha \ X \ ]$

Possessor DPs typically appear in the configuration in (18), and so the LBC tells us that *wh*-constituents that are possessors cannot be extracted from the DP that dominates them, as in:

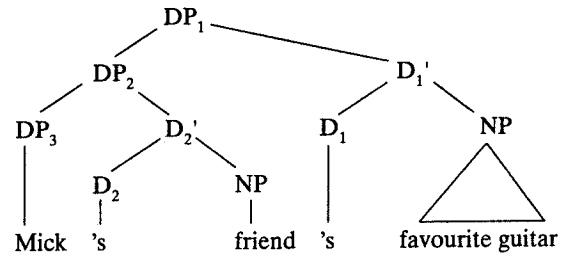
(18) *Whose$_i$ did you play [$_{DP}$ t$_i$ GUITAR] ?

Possessor DPs can be left-recursive, as in (19):

(19) Mick's friend's favourite guitar

In (19), *Mick's* is on the left branch of the DP *Mick's friend*, which is in turn on the left branch of the whole DP.

(20)



The LBC prevents extraction of the *wh*-element corresponding either to DP$_1$ or to DP$_2$:

(21a) *Whose$_i$ did you play [$_{DP}$ t$_i$ FRIEND'S FAVOURITE GUITAR] ?

(21b) *Whose friend's$_i$ did you play [$_{DP}$ t$_i$ FAVOURITE GUITAR] ?

(You might have noticed that we've treated the *'s* as a D in (20), in line with what we suggested in 1.3.2; if *whose* is really *who* + *'s*, then there might be a very simple explanation for the ungrammaticality of examples like (18) and (21): here a non-constituent made up of SpecDP and D is trying to move.) Instead of extraction from the left branch, when a possessor DP is a *wh*-element, it causes the entire DP containing the possessor to move. In other words, the grammatical version of (18) is (22):

(22) Whose guitar$_i$ did you play t$_i$?

This phenomenon, where a category moves which is not itself +*wh* but which contains a +*wh*-element, is known as **pied-piping** (the idea being that the other, non-*wh*, parts of the constituent containing the *wh*-element have to follow the *wh*-element wherever it goes, just as the rats of Hamelin followed the Pied Piper).

Up to now, I've been implicitly assuming that the island constraints are universal. This is in fact largely true; later in this chapter we'll see some examples of the kind of variation that exists across languages regarding islands, but it is mostly rather slight. However, the LBC appears to be violable in some languages. Ross gives the following examples from Russian and Latin:

(23a) Č'ju$_i$ ty čitaješ [$_{DP}$ t$_i$ KNIGU] ? (Russian)
whose you are-reading book
'Whose book are you reading?'

(23b) Cuius$_i$ legis [$_{DP}$ t$_i$ LIBRUM] ? (Latin)
whose you-are-reading book
'Whose book are you reading?'

We can add French *combien* ('how much/many') to this list:

(24) Combien$_i$ as-tu lu [$_{DP}$ t$_i$ DE LIVRES] ?
how-many have you read of books
'How many books have you read?'

It seems clear that the LBC is subject to parametric variation, unlike the other island constraints that we've looked at up to now. Beyond simply stating that some languages respect the LBC and others don't, at this point it's rather difficult to know what the parametric variation really consists of. It may be that it is connected to the nature of the D-position: Latin and Russian both lack overt D-elements, for example (although this doesn't explain the French fact in (24)). The precise status of the LBC is an open question at present.

The wh-*island Constraint:*

This was really the first kind of island phenomenon to be discussed, as it was first mentioned by Chomsky (1964).

The basic observation here is that a *wh*-element cannot be extracted out of clauses introduced by another *wh*-element. Now, there are two main types of clause that are introduced by *wh*-elements: questions and relatives. Relatives come under the CNPC, as we have seen. Interrogatives are thus the relevant case here:

(25a) ??Whose car were you wondering [HOW$_j$ TO FIX t_i t_j] ?

(25b) ?*Whose car$_i$ were you wondering [HOW$_j$ YOU SHOULD FIX t_i t_j] ?

(Since there are two instances of *wh*-movement in these examples, one to the lower SpecCP and one to the higher, there are two traces: the order of the traces is $t_{\text{whose car}}$ t_{how} in each case, as the order of the non-*wh*-elements shows: *I fixed Bill's car with a cocktail shaker.*)

Most speakers don't find either of the examples in (25) very good, but would probably agree that there's a contrast between (25a) and (25b). Tensed *wh*-complements seem to give rise to 'stronger' ungrammaticality than infinitival ones.

The above are the main islands discussed by Ross. So we see that *wh*-movement is subject to at least the CNPC, the Subject Condition, the CSC, the *wh*-island Constraint, and (in some languages) the LBC. However, as I said above, this only shows that *wh*-movement is bounded *sometimes*. It still appears to be fundamentally different from the other instantiations of Move-α in that a *wh*-category can move over an unlimited amount of material, provided that that material contains no islands.

Ross discussed another constraint indicating that *wh*-movement may be consistently bounded, one which later became known as the 'Right Roof Constraint'. Standard *wh*-movement in questions and relatives of the kind that we've been looking at always moves the *wh*-element to the left; in fact, movement is always to SpecCP. However, there are movement rules which appear to move constituents to adjoined positions on the right of the starting point. It's rather unclear what the landing-site of rules like extraposition might be, although it seems to be a position right-adjoined to a VP or to a clausal functional category (right-adjunction structures are ruled out by Kayne's LCA, discussed in Chapter 1, but we'll gloss over that matter here).

In any case, it seems pretty clear that the landing-site of extraposition and similar operations is not an L-related position, and so the operation is a kind of *wh*-movement, in that the moved constituent A'-binds its trace. One operation of this kind is extraposition, which relates pairs like (26):

(26a) The claim [$_{CP}$ that the world was round] was made by the Greeks

(26b) The claim t_i was made by the Greeks [$_{CPi}$ that the world was round]

Here the CP *that the world was round* is apparently moved to the right of the main clause. Ross observed that this operation cannot go further; that is, it is impossible to extrapose 'two clauses up'. The ungrammaticality of this kind of 'long' extraposition is shown in (27):

(27) *The proof that the claim t_i was made by the Greeks was given in 1492 [$_{CPi}$ that the world was round]

Compare (27) with (28), where extraposition goes just 'one clause up':

(28) The proof that the claim t_i was made [$_{CPi}$ that the world was round] by the Greeks was given in 1492

Ross showed that all operations that move material to the right are bounded in this way. So we see that only leftward *wh*-movement from non-islands is unbounded: the other cases of *wh*-movement are bounded either by islands or by the Right Roof Constraint. The obvious question to ask now is: what's special about leftward *wh*-movement from non-islands? In Section 4.2, we'll see how Chomsky (1973) dealt with this.

In this section, I've introduced a whole range of facts, all of which clearly indicate that *wh*-movement is subject to constraints of various kinds. Exactly what the theoretical interpretation of these facts should be remains to be seen; this is largely what we'll be looking at for the rest of this chapter. What we've seen is the following:

(29a) Rightward movement to A'-positions is always bounded

(29b) Leftward movement of *wh*-XPs is subject to island constraints

To recapitulate, here's a list of the island constraints we've seen:

CNPC:

(30a) *Which band$_i$ did you write [$_{DP}$ A SONG WHICH [$_{IP}$ WAS ABOUT t_i]] ?
*Which band$_i$ did you believe [$_{DP}$ THE CLAIM THAT [$_{IP}$ WE HAD SEEN t_i]] ? (see (9))

Subject Condition:

(30b) *Which rock star$_i$ was [$_{CP}$ THAT THE POLICE WOULD ARREST t_i] expected ?
??Which rock star$_i$ were [$_{DP}$ ADMIRERS OF t_i] arrested ? (see (12))

CSC:

(30c) *What$_i$ did Bill buy POTATOES AND t$_i$?
*What$_i$ did Bill buy t$_i$ AND POTATOES ?
*Which guitar$_i$ does KEITH [PLAY t$_i$] AND [SING MADRIGALS] ?
*Which madrigals$_i$ does KEITH [PLAY THE GUITAR] AND [SING t$_i$] ?
(see (16))

Left branch condition:

(30d) *Whose$_i$ did you play [$_{DP}$ t$_i$ guitar] ?

wh-*islands:*

(30e) ??Whose car were you wondering [HOW$_j$ TO FIX t$_i$ t$_j$] ?
?*Whose car$_i$ were you wondering [HOW$_j$ YOU SHOULD FIX t$_i$ t$_j$] ?

Every now and then in this book I've mentioned what a bore lists are. The reason for this is that lists don't really explain things, and what we're after when we're in the business of theory-construction is explanations. What we've seen in this section, I hope you'll agree, is a range of fascinating and complex facts (incidentally, facts that were totally unknown before the development of generative grammar, and mostly before Ross's 1967 Ph.D. dissertation). But we've just got a long list of island constraints. What we need now is some kind of unifying principle, which can tell us *why* the islands are the way they are. And, ultimately, of course, we want to link all these facts about *wh*-movement to what we've seen in connection with DP-movement and head-movement. The next sections all develop these points, beginning with the first real principle of locality: subjacency.

4.2 Subjacency

4.2.1 Successive Cyclicity

The basic idea behind subjacency was a fundamental conceptual move: we drop the assumption that *wh*-movement, at least in leftward-movement from non-islands, is unbounded. So: wh-*movement, despite appearances, is bounded.* This very important idea can only work if we assume that *wh*-movement can, like DP-movement, operate in successive cycles, or **successive-cyclically**. This idea makes it possible to, as it were, 'measure the distance' between the landing-site of a *wh*-category and its starting position.

In Section 2.3 of Chapter 2 we saw that DP-movement can give the impression of moving a DP a long way, but that each instance of apparently non-local movement can be broken down into a series of successive local hops, each of which actually moves the DP to the nearest available position. Here are the examples we saw then:

(31a) The train$_i$ seems [t'$_i$ to be likely [t$_i$ to be late]] (2.3 (48b))

(31b) The money$_i$ seems [t'$_i$ to have been stolen t$_i$]

(31c) The train$_i$ is expected [t'$_i$ to arrive t$_i$ an hour late]

(31d) The weeds$_i$ appear [t'$_i$ to have grown t$_i$ while we were on holiday]

(31e) The patient$_i$ seems [t''$_i$ to be expected [t'$_i$ to die t$_i$]] (2.3 (49a–d))

These examples show us that raising can take place 'cyclically'. All these movements are caused by the Case Filter (or feature-checking requirements, in terms of checking theory as introduced in Section 2.6). The cyclic movement shows that a DP which is looking for Case moves to the nearest available potential Case position, looks for Case, and, if it can't find it, moves to the *next* nearest position, and so on. In Chapter 1, we also saw some examples of successive-cyclic head-movement in our discussion of verb second in 1.4.2.4 (V-to-I-to-C movement).

If the other two types of movement (which seem to be phenomenologically simpler than *wh*-movement: of course, that doesn't mean that they really are simpler) can be successive-cyclic, it's entirely reasonable to think that *wh*-movement can be too. Specifically, the proposal is that wh-*elements move successively through SpecCP*. So, the structure for a case of apparently unbounded *wh*-movement like (6c) looks like (6c'):

(6c') [$_{CP1}$ What$_i$ [$_{C'1}$ did [$_{IP1}$ Harry say [$_{CP2}$ t''$_i$ [$_{C'2}$ [$_{IP2}$ you had forced Bill [$_{CP3}$ t'$_i$ [$_{C'3}$ [$_{IP3}$ to buy t$_i$]]]] ?

Here *what* moves from its base position, occupied by t_i, first to the most embedded SpecCP (SpecCP$_3$), then to the intermediate SpecCP$_2$, and then to the matrix SpecCP$_1$. These positions are marked by the appropriately-named 'intermediate traces' t'_i and t''_i. The seemingly unbounded movement thus in fact involves a series of relatively local hops from SpecCP to SpecCP.

If we want to make sure that *wh*-movement is genuinely local in the way indicated in (6c'), then we need to make sure that it moves from SpecCP to the next SpecCP that c-commands it. The c-command condition is built into the definition of Move-α (see (5b)), so what we need to do here is make sure that *wh* always moves to the closest SpecCP. There are two parts to this: (i) we have to make sure that *wh*-movement doesn't put a *wh* category in some position other than SpecCP, and (ii) we have to make sure that *wh*-movement always goes to the closest SpecCP position.

Let's just deal with point (i) for now by saying that *wh*-movement only takes place in order to check a *wh*-feature (as proposed in 2.6.3), and SpecCP is the only position where these features can be checked. Therefore *+wh* categories will never move to any other position.

Point (ii) is where subjacency comes in. To start with, I'll give a simplified version of the original formulation in Chomsky (1973:81)):

Subjacency:

(32) In the following structure, α and β cannot be related by movement:
... α ... [$_{BC}$ ... [$_{BC}$ ... β ....] ..]
where α and β are separated by more than one blocking category BC

Let's follow Chomsky's (1973) proposal in taking the blocking categories

(BCs) to be DP and IP. Now we can see how subjacency will have to go to the next SpecCP up. Consider what happens in (6c') if *wh*-movement 'skips' a SpecCP:

(6c") $[_{CP1}$ What$_i$ $[_{C'1}$ did $[_{IP1}$ Harry say $[_{CP2}$ $[_{C'2}$ $[_{IP2}$ you had forced Bill $[_{CP3}$ t'$_i$ $[_{C'3}$ $[_{IP3}$ to buy t$_i$]]]] ?

Here, *what* has moved from SpecCP3 (the position of t'_i) directly to SpecCP$_1$, skipping SpecCP$_2$. In so doing the movement crosses both IP$_2$ and IP$_1$. We've defined IP as a BC, and so this derivation, in crossing two BCs, violates subjacency. Movement through SpecCP$_2$ is necessary in order to avoid this violation. Since SpecCP$_2$ is available for *what* to transit through, (6c) is grammatical – but only with the derivation given in (6c').

4.2.2 Explaining Island Constraints

Subjacency gives us a kind of measure of syntactic distance in terms of BCs. As such, it can explain many of the island effects we saw in the previous section. Let's look at how this is done. Here are the typical CNPC violations that we saw in the previous section:

(33a) *Which band$_i$ did you write $[_{DP}$ A SONG WHICH $[_{IP}$ WAS ABOUT t$_i$]] ?

(33b) *Which band$_i$ did you believe $[_{DP}$ THE CLAIM THAT $[_{IP}$ WE HAD SEEN t$_i$]] ? (see (9))

Assuming successive-cyclic movement through SpecCP, (33b) has the following representation:

(33b') *$[_{CP1}$ Which band$_i$ did $[_{\mathbf{IP1}}$ you believe $[_{\mathbf{DP}}$ the claim $[_{CP2}$ t'$_i$ that $[_{\mathbf{IP2}}$ we had seen t$_i$]] ?

I've marked the BCs here in bold. The first step of movement, from the base position to SpecCP$_2$, crosses just one BC, IP$_2$. The second, on the other hand, crosses DP and IP$_1$. As such, it violates subjacency. So subjacency accounts for the complement case of the CNPC.

Now look at the relative-clause case:

(33a') *$[_{CP1}$ Which band$_i$ did $[_{\mathbf{IP1}}$ you write $[_{\mathbf{DP}}$ a song $[_{CP2}$ which$_j$ $[_{\mathbf{IP2}}$ t$_j$ was about t$_i$]] ?

Here there are two movements: one of *which* from the subject position of IP$_2$ to SpecCP$_2$, and of *which band* to SpecCP$_1$. If *which* moves first, then *which band* has to move in a single step all the way to SpecCP$_1$. This movement crosses three BCs, and so violates subjacency. If *which band* moves first, then it can presumably move cyclically through SpecCP$_2$. Movement from this position will nevertheless violate subjacency, as it crosses the two BCs DP and IP$_1$. I mentioned in 4.1 that relative-clause CNPC violations are worse than complement cases; a natural way to capture this would be by forcing the relative-clause to involve the crossing of three BCs as opposed to two in the complement case. To do this, we need to ensure that *which* moves first in (33a'). The following principle does this:

The Strict Cycle Condition:

(34) Nothing can move from a position c-commanded by an intermediate trace.

If we move *which band* to SpecCP$_2$ and on to SpecCP$_1$ in (33a') before moving *which* to SpecCP$_2$, then we'd have the following representation after the second movement of *which band* but before any movement of *which*:

(33a") *$[_{CP1}$ Which band$_i$ did $[_{IP1}$ you write $[_{DP}$ a song $[_{CP2}$ t'$_i$ $[_{IP2}$ which was about t$_i$]] ?

Here *which* is c-commanded by the intermediate trace t'_i, and so the Strict Cycle Condition is violated. So we force the derivation where *which* moves first to SpecCP$_2$ and so *which band* crosses three BCs en route to SpecCP$_1$. This gives rise to a stronger violation of subjacency than the complement CNPC example.

Next, let's look at the Subject Condition. Here are the examples from the last section with the successive-cyclic *wh*-movements indicated:

(35a) *$[_{CP1}$ Which rock star$_i$ was $[_{\mathbf{IP1}}$ $[_{CP2}$ t'$_i$ that $[_{\mathbf{IP2}}$ the police would arrest t$_i$]] expected]]?

(35b) ??$[_{CP1}$ Which rock star$_i$ were $[_{\mathbf{IP}}$ $[_{\mathbf{DP}}$ admirers of t$_i$] arrested] ? (see (12))

You can see straight away that the movement of *which rock star* in (35b) crosses two BCs, DP and IP. On the other hand, (35a) is allowed, since the movement can pass cyclically through SpecCP$_2$. However, it has often been suggested that subjects, meaning categories that occupy SpecIP, are always DPs. If that's true, then the correct representation of (35a) would be:

(35a') *$[_{CP1}$ Which rock star$_i$ was $[_{\mathbf{IP1}}$ $[_{\mathbf{DP}}$ $[_{CP2}$ t'$_i$ that $[_{\mathbf{IP2}}$ the police would arrest t$_i$]]] expected ?

Here we can see that *which rock star* crosses two BCs on its way to SpecCP$_1$. So subjacency can account for the Subject Condition.

Subjacency can also account for *wh*-islands. Let's look at the examples we gave in the previous section (see (25)), highlighting the BCs:

(36a) ?$[_{CP1}$ Whose car were $[_{\mathbf{IP1}}$ you wondering $[_{CP2}$ how$_j$ $[_{\mathbf{IP2}}$ to fix t$_i$ t$_j$] ?

(36b) ?*$[_{CP1}$ Whose car$_i$ were $[_{\mathbf{IP1}}$ you wondering $[_{CP2}$ how$_j$ $[_{\mathbf{IP2}}$ you should fix t$_i$ t$_j$] ?

Subjacency treats both of these examples alike. Remember that the Strict Cycle Condition (see (34)) makes the movement to SpecCP$_2$ take place first every time, so the category which undergoes 'long' movement to SpecCP$_1$ crosses both IP$_2$ and IP$_1$ in both of these examples, and subjacency rules them

both out. Obviously, this is basically a good result, but we'd like to know why tensed *wh*-islands are worse than infinitival *wh*-islands. So, although subjacency tells us that these examples are bad, it doesn't make any distinctions as to *degree* of ungrammaticality which is what we'd ideally like.

Next, let's see how subjacency can handle the Left Branch Condition. Here, again, is the typical example with the BCs highlighted:

(37) *Whose$_i$ did [$_{IP}$ you play [$_{DP}$ t$_i$ guitar]] ?

This example is straightforward; movement crosses IP and DP and subjacency is violated. In fact, subjacency works too well here, seemingly, since we saw that in a number of languages the LBC doesn't seem to hold (see examples (23 and 24)).

So far we've seen that subjacency does a pretty impressive job of unifying the island constraints we saw in the previous section. The CSC, however, is not so readily captured by subjacency as formulated in (32). Here are the relevant examples again:

(38a) *What$_i$ did [$_{IP}$ Bill buy [$_{DP}$ potatoes and t$_i$]] ?

(38b) *What$_i$ did [$_{IP}$ Bill buy [$_{DP}$ t$_i$ and potatoes]] ?

(38c) *Which guitar$_i$ does [$_{IP}$ Keith [$_{VP}$ [$_{VP}$ play t$_i$] and [$_{VP}$ sing madrigals]]] ?

(38d) *Which madrigals$_i$ does [$_{IP}$ Keith [$_{VP}$ [$_{VP}$ play the guitar] and [$_{VP}$ sing t$_i$]]] ?

I'm following the standard assumption that conjoined categories form a bigger category of the same kind: you 'ight note in passing that this means that coordinate structures don't fit the X'-schema that we gave in Chapter 1 – see Kayne (1994) for a recent proposal to reconcile coordination with X'-theory). Extraction out of a conjoined DP as in (38a and b) therefore violates subjacency, as you can see. However, extraction from a coordinated VP, as in (38c and d), doesn't. Here only one BC, IP, is crossed. To reconcile the CSC with subjacency, we'd need either to show that all coordinated categories were IPs or DPs or to say that the category dominating conjoined categories – whatever it is – is a BC. The former approach just won't wash empirically, and the latter amounts to restating the CSC. So the CSC poses a problem. And remember that 'across-the-board' extraction is allowed: see (16) above.

So far, we've seen that subjacency can give a unified analysis of the CNPC, the Subject Condition, some cases of the CSC, the LBC, and *wh*-islands. There are two basic kinds of problem. One is that subjacency can't handle a number of cases of the CSC; the other is that subjacency is a blanket condition that just bans certain kinds of movement, and because of this it can't distinguish extraction out of an infinitival *wh*-island from extraction out of a tensed *wh*-island. Nevertheless, it's clear that subjacency goes a good way towards giving a unified explanation for strong islands.

4.2.3 Parametric Variation

Before going on to the Empty Category Principle, I want to look at one interesting piece of apparent parametric variation in how subjacency works. Rizzi (1982) observed that, at first sight, Italian seems to allow subjacency to be violated in *wh*-islands:

(39) Tuo fratello, [a cui$_i$ mi domando [che storie$_j$ abbiano
Your brother, to whom I wonder which stories they-have
raccontato t$_i$ t$_j$], era molto preoccupato
told was very worried

These are tensed *wh*-islands, cases where subjacency gets it just about right for English by ruling extraction out altogether (Rizzi didn't discuss adjunct extraction, so we'll just leave that aside here). Why are they allowed in Italian?

One solution would be to say that subjacency just doesn't hold in Italian, but Rizzi shows that this isn't right. The CNPC is respected in Italian:

(40) *Tuo fratello, a cui$_i$ temo la possibilità che abbiano
Your brother, to whom I-fear the possibility that they-have
raccontato tutto t$_i$, ...
told everything, ...

Even more interestingly, Rizzi shows that the *wh*-islands of a more complicated sort do exist in Italian. Extraction from a *wh*-island inside another *wh*-island is impossible. Example (41a) gives the basic structure with one *wh*-island contained within another. This structure, like its English counterpart, is fine, since both the *wh*-movements that form the respective *wh*-islands are local, as the coindexed traces indicate. However, if we try to extract from the lowest island, ungrammaticality results in both English and Italian, as (41b) shows (in this example, I've only indicated the coindexed trace of the illicitly moved relative element):

(41a) Mi sto domandando [a chi$_i$ potrei chiedere t$_i$ [quando$_j$
I am wondering to whom I-may ask when
dovrò parlare di questo argomento t$_j$]
I'll have-to speak about this topic
'I'm wondering who I can ask when I'll have to speak about this topic'

(41b) *Questo argomento, [di cui$_k$ mi sto domandando [a chi
This topic of which I am wondering to whom
potrei chiedere [quando dovrò parlare t$_k$]]], mi
I-may ask when I'll-have-to speak to-me
sembra sempre più complicato
seems ever more complicated
'This topic, which I am wondering who I can ask when I'll have to talk about, seems more and more complicated to me'

Rizzi concludes that the only reasonable solution is that subjacency holds in Italian but that *Italian has different blocking categories from English.* Specifically, the blocking categories in Italian are CP and DP, rather than IP and DP. In this way, the above data are accounted for, as we can see in (42), where the Italian BCs are highlighted:

(42a) Tuo fratello, [a cui$_i$ mi domando [$_{CP}$ che storie$_j$
Your brother, to whom I wonder which stories
abbiano raccontato t_i t_j], era molto preoccupato
they-have told was very worried
((39). One Italian BC crossed)

(42b) *Tuo fratello, a cui$_i$ temo [$_{DP}$ la possibilità [$_{CP}$ che
Your brother, to whom I-fear the possibility that
abbiano raccontato tutto t_i, ...
they-have told everything, ...
((40). Two Italian BCs crossed)

(42c) *Questo argomento, [di cui$_k$ mi sto domandando [$_{CP}$a chi
This topic of which I am wondering to whom
potrei chiedere [$_{CP}$ quando dovrò parlare t_k]]], mi
I-may ask when I'll-have-to speak to-me
sembra sempre più complicato
seems ever more complicated
((41b). Two Italian BCs crossed)

And compare (39) (= (42a)) with its English counterpart: here the different BCs in the different languages are highlighted:

(43a) Tuo fratello, [a cui$_i$ mi domando [$_{CP}$ che storie$_j$ abbiano raccontato t_i t_j era molto preoccupato
(One Italian BC crossed)

(43b) *Your brother, who$_i$ [$_{IP}$ I wonder which stories [$_{IP}$ they've been telling t_i, was very worried
(Two English BCs crossed)

And so Rizzi showed that languages can differ according to which BCs they select. His analysis of the difference between English and Italian, although convincing in itself, raises another question: what is the range of choice among the BCs that UG makes available? Could we expect to find a language in which VP and AP were barriers? If not, why not? What we need is a more principled way of deciding what the possible BCs are, both in UG and at the level of the parametric choices made by different languages. This is going to be the principal topic of Section 4.4.

4.2.4 Conclusion

In this section we've seen the following points:

- the proposal that *wh*-movement is successive cyclic, and bounded by subjacency (see (32))
- the Strict Cycle Condition (see (34))
- how subjacency accounts for (most) island constraints
- Rizzi's proposal that Italian chooses different BCs from English.

In Section 4.4, we'll come back to the question of giving a more principled account of what the class of BCs is. Before doing that, however, I want to introduce some more data regarding *wh*-movement, and the other main locality principle: the Empty Category Principle.

4.3 The Empty Category Principle

In this section, I want to introduce the Empty Category Principle (ECP). This principle constrains Move-α by imposing an LF licensing requirement on traces. Although the ECP is intended to apply to all traces, including *wh*-traces, DP-traces and head-traces, in this section I'll restrict the discussion to *wh*-traces (we'll look at how it extends to other kinds of traces in 4.5). Since it's a condition on traces rather than on movement, the ECP can be made to distinguish different kinds of traces. In this way, as we'll see in 4.3.1, it can account for the phenomenon of **argument-adjunct asymmetries**. It can also handle the constraints on *wh*-movement in languages which appear to lack an overt version of this movement, as we'll see. In 4.3.2 I'll introduce another locality phenomenon, the **complementizer-trace effect**; we'll see that the ECP can handle this, and can give a very interesting analysis of the parametric variation associated with it. Finally, 4.3.3 looks at some extensions of the ECP proposed by Kayne; in this section, I'll also introduce the intriguing phenomenon of **parasitic gaps**.

4.3.1 Argument-Adjunct Asymmetries

4.3.1.1 Lexical Government and Antecedent Government

Up to now, I've presented the constraints on *wh*-movement, whether appearing as islands or unified under subjacency, as blanket constraints on any kind of movement. But in fact there are important differences between arguments of certain types and adjunct elements with regard to extraction. These differences emerge if we compare the extraction of a direct-object *wh*-element from a *wh*-island – seen in (44) (which is the repetition of (36)) – with extraction of an adverbial element from the same island, seen in (45):

(44a) ?[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ to fix t_i t_j] ?

(44b) ?*[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ you should fix t_i t_j] ?

(45a) *How$_j$ were you wondering [WHOSE CAR$_i$ TO FIX t_i t_j] ?

(45b) *How$_j$ were you wondering [WHOSE CAR$_i$ YOU SHOULD FIX t_i t_j] ?

The difference between (44) and (45) seems to be as follows: while the argument-extraction examples in (44) are very awkward, they are intelligible; in examples like (45), on the other hand, it is all but impossible to see the interpretation that is being looked for (with *how* interpreted as modifying the lower clause, looking for an answer like 'with a spanner' in each case). This suggests that the badness of adjunct-extraction has to do with an LF condition which prevents certain kinds of interpretations. This is where the ECP, which we can think of as an LF-condition on traces, comes in. As a first formulation, let's take the following:

ECP:

(46) Traces must be properly governed

'Proper government' here means a subspecies of government. Or, more precisely, it refers to two subspecies of government:

Proper Government:

(47) either: government by a lexical head (lexical government)
or: government by the moved category (antecedent government)

To see how the ECP accounts for these asymmetries, we need to look again at the definition of government that I gave in 2.2.2, (16):

Government:

α, a head, governs β if and only if

(48a) α c-commands β

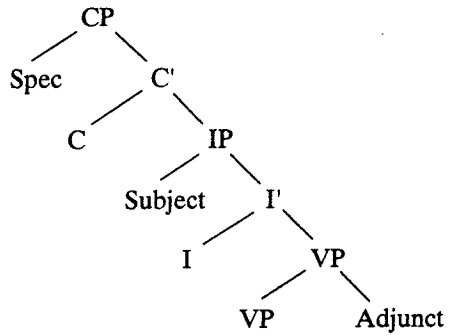
(48b) no barrier dominates β but not α

Barrier:

(49) Any XP except IP

Here's the structure of a simple clause (glossing over the 'split-Infl' structure once more):

(50)



I'm assuming that adjuncts are adjoined to VP. In that case, adjuncts are not governed by anything, given the definitions in (48) and (49). Complements, on the other hand, are always governed by the lexical head that selects them. So complements are always lexically governed and adjuncts never are. This means that traces in complement position always satisfy the ECP, while adjunct traces can only satisfy the ECP by being antecedent-governed. So now we need to look more closely at antecedent-government.

We'll elaborate our conception of antecedent-government steadily over Sections 4.4 and 4.5 – in fact, we'll see that it comes close to giving us the key unifying concept for the theory of locality in 4.5. For the moment, it's enough simply to state that antecedent-government is defined in terms of blocking categories just like subjacency:

Antecedent Government:

α, a moved category, antecedent-governs β if and only if

(51a) α c-commands β

(51b) no more than one blocking category dominates β but not α

For now, we retain from the previous section the idea that IP and DP are blocking categories. You might notice that the definition of antecedent-government in (51) isn't really much like the definition of government in (48). Unifying these definitions is one of the tasks of Section 4.4.

So, given the definition of antecedent-government in (51), the ECP basically requires that an adjunct trace be subjacent to its antecedent, otherwise it will not be licensed at LF. If the trace fails to be licensed at LF, then the interpretation of the antecedent-trace relation will not be available, and the effect of uninterpretability that we noticed in (45) will arise. So, an ECP violation and a subjacency violation give different kinds of ungrammaticality: subjacency violations give rise to syntactic awkwardness, while ECP violations are usually uninterpretable on the intended reading. Sometimes, as in the case of infinitival *wh*-islands, the subjacency violation appears to be rather mild, as we've seen.

The argument-adjunct distinction can be found in the other islands that we've seen. We can see this in the following examples:

(52a) CNPC:
*How$_i$ do you believe [the stories that [John fixed your car t_i]] ?

(52b) Subject Island:
*How$_i$ would [to fix your car t_i] be best ?

In each of these cases, in addition to the syntactic awkwardness created by the violation of subjacency, the intended interpretation (where *how* modifies the predicate inside the island, indicated by the position of the trace inside the brackets) is all but impossible to perceive. This is because the traces fail the ECP, since they are neither lexically governed nor antecedent-

governed. They are not lexically governed because they are adjuncts, and they are not antecedent-governed because their antecedent is separated from them by more than one BC (see the discussion of these island effects in the previous section).

In fact, calling these effects argument-adjunct asymmetries is slightly misleading. If you look again at the definition of lexical government, and at the clause structure in (50), you'll see that subjects are not lexically governed either. In (50), the subject position – Spec,IP – is governed by C, but C isn't lexical. So subjects ought to pattern with adjuncts as far as the asymmetries in (44) and (45) go. In fact this is basically true, as we can see from examples like (53):

(53) *Which band$_i$ were you wondering whether t_i will play that song ?

If we assume that *whether* is in SpecCP here, the dependency between *which band* and its trace crosses two IPs, that is the subject trace fails to be antecedent-governed. Since the subject is not lexically governed, the sentence is very bad. On the other hand, subjects that are in positions governed by a lexical head, typically a verb, are lexically governed and so we don't find the kind of violation seen in (53):

(54) Which band$_i$ did you consider t_i to be the best ?

Remember that complements to ECM Verbs like *consider* are AgrSPs: see 2.2.3. Leaving aside ECM subjects, we should really talk about 'complement/non-complement asymmetries' instead of talking about 'argument-adjunct asymmetries'. In fact, for now I'll use both terms synonymously. We'll come back to this and related points in the next subsection.

4.3.1.2 The ECP at LF: Comparative Evidence

There's very interesting and important comparative evidence that the ECP holds of LF, while subjacency holds of overt movement. This comes primarily from Huang's (1982) study of *wh*-movement and related phenomena in Chinese. As we saw in 2.6.3, *wh*-elements don't undergo overt movement in Chinese:

(55a) Zhangsan yiwei Lisi mai-le shenme?
Zhangsan think Lisi bought what
'What does Zhangsan think Lisi bought?'

(55b) Zhangsan xiang-zhidao Lisi mai-le shenme
Zhangsan wonders Lisi bought what
'Zhangsan wonders what Lisi bought'

As we mentioned in 2.6.3, it's natural to think that Chinese behaves this way because it has a weak *wh* feature (although we also mentioned that Watanabe (1992) proposes a different view).

What we're interested in here, though, is the fact that Huang shows that *covert wh-movement in Chinese is subject to the ECP*. In other words, we don't find the island constraints that affect movement of complements that we looked at using English data in 4.1:

(56a) CNPC (relative clause):
ni zui xihuan [SHEI MAI DE SHU] ?
you most like who buy Prt book
'Who is the *x* such that you like the books *x* bought?'

(56b) Subject Condition:
[WO MAI SHENME] zui hao ?
I buy what most good
'What is it best that I buy?'

(56c) Argument *wh*-Island:
ni xiang-zhidao [WO WEISHENME MAI SHENME] ?
you wonder I why buy what
'What is the *x* such that you wonder why I bought *x* ?'

But we do find that adjunct *wh*-elements inside *wh*-islands cannot be interpreted. Hence (56c) cannot have the interpretation 'What is the reason *x* such that you wonder what I bought for *x* ?' Huang interprets this fact as showing that adjuncts cannot move out of *wh*-islands at LF. In other words, adjunct-traces are subject to the ECP. Huang proposes that the ECP applies to the traces of covert movement, that is, that the ECP is an LF-condition on traces while subjacency only constrains overt movement. LF-movement of complement *wh*-elements as in (55) and (56) is fine, therefore, whether these elements are in islands or not. The traces of these movements are uniformly lexically governed. On the other hand, adjunct traces fail to be antecedent-governed where their antecedent is moved out of an island (see the definition of antecedent-government in (51)); the adjunct traces fail the ECP in examples like (57a), and the interpretation is unavailable. In a similar way, we account for the ungrammaticality of sentences with a *wh*-element inside other kinds of islands – these examples contrast with those in (56a and b) (the English translations indicate the interpretation of 'why' which is relevant here; these translations are of course also ungrammatical in English with the coindexation given):

(57a) CNPC (relative clause):
*ni zui xihuan [WEISHENME MAI SHU DE REN] ?
you most like why buy book Prt person
'Why$_i$ do you like [the man who bought the books t_i]]?'

(57b) Subject Condition:
*[WO WEISHENME MAI SHU] zui hao ?
I why buy book most like
'Why$_i$ is [that I buy the books t_i] best ?'

Huang's work gave a very nice cross-linguistic confirmation of the existence of covert movement and of the distinction between subjacency and the ECP. However, there are clearly connections between subjacency and the ECP, in particular in that antecedent-government, as we defined it in (51), seems to be very close to subjacency. We also have to look more closely at the notions of blocking category, relevant for subjacency and antecedent-government, and barrier, relevant for government and lexical government. These are the central issues addressed by the *Barriers* system (Chomsky (1986b)); we'll look at them in detail in 4.4.

4.3.2 Complementizer-trace Effects and the Null Subject Parameter Revisited

4.3.2.1 Complementizer-trace Effects

The original motivation for the ECP didn't in fact come from the asymmetries that we've been looking at, but from Complementizer-trace effects (henceforth C-t effects). The observation here is that extraction of a subject across a complementizer is not good:

(58) *Who$_i$ did you say that t$_i$ wrote this song?

For a while in the 1970s it was thought that examples like (58) were evidence that *wh*-traces were subject to Principle A of the binding theory. In this way, (58) is assimilated to examples like:

(59) *Mick$_i$ thinks that himself$_i$ is the greatest

However, *wh*-traces are fine in object position in this kind of example, unlike anaphors:

(60a) Who$_i$ did you say that Phil admires t$_i$?

(60b) *Mick$_i$ thinks that Marianne admires himself$_i$

Comparing (58) and (60a), we see another example of a complement/non-complement asymmetry (a subject-object asymmetry in this case), and so the ECP is the relevant principle. In (58a), the subject trace is not lexically governed, unlike the object trace in (60a). If the subject trace also fails to be antecedent-governed in (58), then we can rule it out by the ECP.

According to the definition of antecedent-government in (51), however, the subject trace in (58) would be acceptable. We can see this if we look more closely at the structure of (58), taking into account successive-cyclic movement and highlighting the BCs:

(58') *[$_{CP1}$ Who$_i$ did [$_{IP1}$ you say [$_{CP2}$ t'$_i$ that [$_{IP2}$ t$_i$ wrote this song]]]] ?

Who moves from the lower subject position (SpecIP$_2$) to SpecCP$_2$, crossing just the one BC, IP$_2$. The second step of movement takes it to SpecCP$_1$, crossing just IP$_1$. So there is no step of movement that crosses two BCs here. In fact, the movement appears to be exactly the same as that which the object *wh*-element undergoes in (60a), as you should be able to see.

An important clue to what's going on in (58) comes from the grammaticality of examples like (61), where *that* has been dropped (an option which is generally available in finite complements to verbs in English):

(61) Who$_i$ did you say t$_i$ wrote this song?

Example (61) presumably looks like (62') if we take successive cyclicity into account:

(62') [$_{CP1}$ Who$_i$ did [$_{IP1}$ you say [$_{CP2}$ t'$_i$ [$_{IP2}$ t$_i$ wrote this song]]]] ?

The contrasts between (58) and (60a) and between (58) and (61) show that subject-extraction is sensitive to the presence of a complementizer: the Complementizer-trace effect.

We can account for the C-t effect in terms of the ECP if we say that the presence of the complementizer in (58) blocks antecedent government of the subject trace. Since the subject isn't lexically governed, the presence of the complementizer leads to ungrammaticality. Complements, on the other hand, are quite indifferent to the presence of the complementizer, as (60a) and (63) show:

(63) Who$_i$ did you say Phil admires t$_i$?

However, we just pointed out that the definition of antecedent government in (51) allows the subject trace to be antecedent-governed in (58). More generally, the definition in (51) won't distinguish between the presence and the absence of a complementizer.

We can handle complementizer-trace effects by adding a notion of 'minimality' to the definition of antecedent government, as follows:

Antecedent Government (Revised):
α, a moved category, antecedent-governs β if and only if

(51a') α c-commands β

(51b') no more than one blocking category dominates β but not α

(51c') there is no filled C-position that minimally c-commands β and does not c-command α

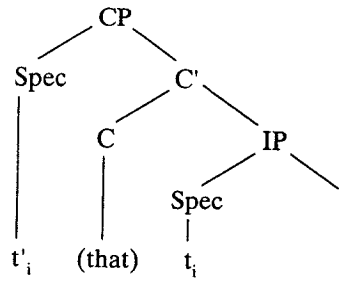
The definition of minimal c-command was given in 1.4.2 (68), as follows (and see the Appendix to Chapter 2):

Minimal C-command:

(64) α minimally c-commands β iff α c-commands β and there is no γ, such that γ both c-commands β and does not c-command α

Let's look at the relevant parts of (58) to see how (51') handles C-t effects:

(65)



We can see that C minimally c-commands the subject: it c-commands the subject and there is nothing else that c-commands the subject without also c-commanding C. By (51c'), then, when C is overt it will block antecedent-government between the trace in SpecCP and the trace in SpecIP. Since the trace in SpecIP is not lexically governed, it will violate the ECP when C is overt. This accounts for C-t effects. On the other hand, if the Complementizer is not overt, in other words if *that* is not present, (51c') allows the subject trace to be antecedent-governed by the trace in SpecCP (t'_i).

Rule (51c') doesn't make reference to complementizers as such, just to the presence of a filled C-position. We saw in 1.4.2.2 that in main clauses C can be filled by a moved auxiliary (I). Usually, in main-clause *wh*-questions in English, C must be filled by a moved auxiliary – in other words I-to-C movement is obligatory. However, when the subject is questioned, I-to-C movement is impossible. The following contrast shows this:

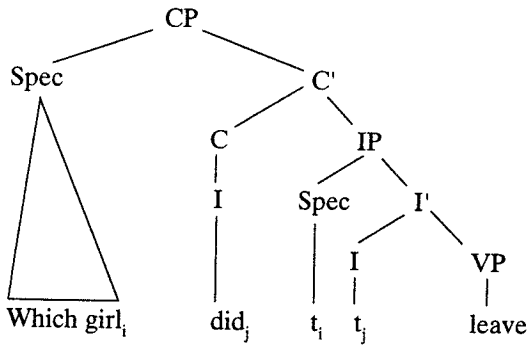
(66a) Which $girl_i$ did_j he t_j kiss t_i ?

(66b) *Which $girl_i$ did_j t_i t_j kiss him?

(66c) Which $girl_i$ t_i kissed him?

(51c') tells us why I-to-C movement (which in these examples is instantiated by movement of *do*) is impossible. Where I, containing the overt element *do*, raises to C, we have a structure like (67):

(67)



Here *did* acts like *that* in (58), and blocks the antecedent-government relation between *which girl* and its trace. Object traces are unaffected, because they are lexically governed (and here I-to-C is obligatory for reasons connected to the *wh*-criterion: see 2.6.3).

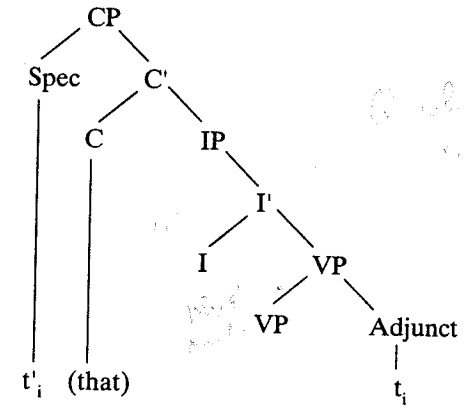
Adjuncts are not sensitive to C-t effects, as (68) shows:

(68) How$_i$ did you say [$_{CP}$ t'$_i$ (that) [$_{IP}$ he fixed your car t$_i$]] ?

The presence or absence of *that* has no effect on the grammaticality of (68). Now, adjunct traces are like subjects in not being lexically governed, so we have to see why the presence of *that* does not affect the ability of the adjunct trace t_i in (68) to be antecedent-governed.

Sticking to the idea that adjuncts are adjoined to VP, t_i in (68) appears in the following configuration:

(69)



The category which minimally c-commands the adjunct in (69) is I (the VP that the adjunct adjoins to doesn't c-command it – see the discussion of adjunction in 2.6.4). The other two clauses of (51) allow t'_i to antecedent-govern t_i: t'_i c-commands t_i and only one BC – IP – intervenes. We see then that it is clause (51'c) of the definition of antecedent-government that is central in accounting for C-t effects.

4.3.2.2 The Null-subject Parameter Again

C-t effects have been of great interest for comparative syntax. The basic observation was originally made by Perlmutter (1971), and has become known as **Perlmutter's generalization**. This generalization states that *null-subject languages do not show C-t phenomena* (on null-subject languages, see 3.3.3). This is illustrated by Italian sentences like (70):

(70) Chi hai detto che ha scritto questo libro?
Who have-you said that has written this book?
'Who did you say wrote this book?'

As you can see from the gloss and from the English translation (which, in

order to be grammatical, does not contain a translation of *che* = 'that'), the extracted category *chi* is the subject of the embedded clause. The complementizer is also present (this is obligatory in this context in Italian), and the sentence is fine.

One way to handle (70) is to say that (51'c) is switched off in some languages. However, this is rather an uninteresting 'solution' to the problem, and does not offer any direct way to capture the correlation with the availability of null subjects (the ability to license *pro* in subject position, as we saw in 3.3.3). A much more interesting approach was suggested by Rizzi (1982). He connected Perlmutter's generalization to the availability of 'free inversion' in null-subject languages. In 3.3.3.3, we saw that the subjects in examples like (71) are in a postverbal position (distinct from the direct-object position – see 2.3.2), while *pro* occupies the subject position:

(71a) Hanno telefonato molti studenti
Have-3Pl phoned many students (Pl)
'Many students have phoned'

(71b) Vinceremo noi
Will-win-1Pl we (1Pl)
'We will win'

Let's suppose that the postverbal subjects are in a position adjoined to VP. We can now account for Perlmutter's generalization by saying that C-t effects can be *apparently* violated in null-subject languages, since *subjects can be extracted from the adjunct-like postverbal position*. The definition of antecedent-government in (51') allows a trace in VP-adjoined position to be antecedent-governed by a trace in SpecCP whether or not C is overt. Extraction of the postverbal subject is rather like extraction of an adjunct in English, and as such is not subject to Complementizer-trace effects.

Rizzi's analysis implies that the representation of (70) is (70'a) rather than (70'b):

(70'a) $[_{CP1}$ Chi$_i$ $[_{IP1}$ *pro* hai detto $[_{CP2}$ t'$_i$ che $[_{IP2}$ *pro* ha $[_{VP}$ scritto questo libro] t$_i$]]]]]?

(70'b) $[_{CP1}$ Chi$_i$ $[_{IP1}$ *pro* hai detto $[_{CP2}$ t'$_i$ che $[_{IP2}$ t$_i$ ha $[_{VP}$ scritto questo libro]]]]] ?

In (70'a), *pro* in the main clause is an argumental pronoun – the silent version of *tu* ('you') – while *pro* in the embedded clause is an expletive. The expletive *pro* in the embedded clause is formally licensed by AgrS, as we saw in 3.3.3. This possibility is only open to null-subject languages, hence Perlmutter's generalization.

Rizzi's idea that the subject is extracted from postverbal position in finite clauses in null-subject languages receives direct support in certain Northern Italian dialects. In these dialects, a subject clitic appears in many persons of finite clauses as a kind of 'extra' marker of agreement; in fact, it is probably

an overt AgrS. The following examples from the Florentine dialect (from Brandi and Cordin (1989)) illustrate:

(72a) Mario E parla
Mario SCL speaks
'Mario speaks'

(72b) E parla
SCL speaks
'He speaks'

(72c) *Parla
speaks

(Brandi and Cordin (1989) and Rizzi (1986b) show that this element is not, despite appearances, a subject pronoun; for a detailed discussion of these clitics in many dialects of the region, see Poletto (1993).)

In 'free-inversion' sentences like (71), an expletive subject clitic appears in preverbal position in Florentine:

(73) Gl 'ha telefonato delle ragazze
SCL (MSG) has telephoned some girls (FPL)
'Some girls telephoned'

As you can see, the preverbal subject clitic that appears when there is free inversion does not agree with the postverbal subject. Instead, it and the verb are in a default, third-person-masculine singular form. We can think of this subject clitic as the one that licenses an expletive *pro*. In most languages, including English, expletives are third-person-masculine singular.

Now, the really interesting and nice fact is this: *when the subject is extracted from a finite clause with an overt complementizer, the default subject clitic appears.* This is shown in (74):

(74a) Quante ragazze tu credi che gli abbia parlato?
How-many girls you think that SCL(MSG) has (3SG) spoken
'How many girls do you think have spoken?'

(74b) *Quante ragazze tu credi che le abbiano parlato?
How-many girls you think that SCL(FPL) have (3PL) spoken

Example (74a) shows the clitic-agreement pattern typical of a free-inversion sentence like (73); so we conclude that the subject is extracted from the postverbal position. Example (74b) shows the preverbal clitic-agreement pattern (in other words, the subject clitic agrees with the verb). We can attribute the ungrammaticality of this sentence to the C-t effect, since the clitic-agreement pattern shows the subject is extracted from preverbal position. This sentence is then ruled out in exactly the same way as (58): the subject trace fails to be antecedent-governed due to (51'c). So this kind of example shows us that the C-t effect is operative in Italian dialects, the only way for extraction of the

subject of a finite clause to be grammatical is by extracting from the postverbal position. More generally, these facts directly support the connection that Rizzi made between free inversion and Perlmutter's generalization.

Putting the discussion here together with that in 3.3.3, we see that four properties characterize null-subject languages, as opposed to non-null-subject languages:

(75a) Possibility of phonologically empty referential subject pronouns

(75b) Impossibility of overt expletive pronouns

(75c) Possibility of free inversion

(75d) Apparent absence of complementizer-trace effects

If a language allows referential *pro* in subject position, it will have the other properties in (75) (all other things being equal); if it does not, it will not have those properties. The way in which the null-subject parameter ties together a number of apparently unconnected possibilities is very elegant, and is also important for the theory of parameters, as we'll see in Chapter 5.

4.3.2.3 Conclusion

In this section, I've introduced C-t effects, and shown how an addition to the definition of antecedent-government – (51'c) – can handle these. You'll have noticed that (51'c) is hardly an elegant definition. In fact, it is rather transparently designed precisely to handle C-t effects. In that sense, although it does the job, it is clearly unsatisfactory. We also saw how Perlmutter's generalization can be handled in a way that extends the empirical coverage of the null-subject parameter and allows us to maintain that, despite initial appearances, null-subject languages do have C-t effects. In other words, antecedent government works the same way in those languages as in non-null-subject languages. In the next section, we'll continue to develop the notion of antecedent-government.

4.3.3 Connectedness, Preposition-Stranding, and Parasitic Gaps

In this section I'll briefly describe some of the proposals made by Kayne in a series of papers dating from the early to mid-1980s (collected in Kayne (1984); see particularly chapters 3 and 8). These proposals primarily concern the nature of antecedent-government. To a certain extent, they also anticipate the *Barriers* framework which is the subject of the next section. In addition to their theoretical importance, Kayne's proposals are of considerable empirical interest.

4.3.3.1 Connectedness

Let's begin by formulating the definition of antecedent government in terms of 'g-projections', as follows:

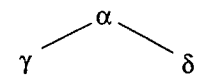
Antecedent Government (Second Revision);
α, a moved category, antecedent-governs β if and only if

(76a) α binds β

(76b) α is connected to a g-projection of γ, where γ canonically governs β

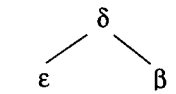
'Connected' here means 'forms a subtree of the whole tree'. G-projections (the name is intended to suggest 'government-projections') are defined in two steps. Consider first the configuration in (77):

(77)



By the definition in (48), γ governs δ. Moreover, γ is on the left of δ. We can call this combination of government and linear order the **canonical government relation** (for VO languages; Kayne suggests that OV languages may have a different canonical government configuration, but we'll leave that aside). In (77) α is either a g-projection of δ, or of some category β of which δ itself is a g-projection. The latter is true just where δ contains a canonical government configuration, as follows:

(78)



A further condition is that the lowest head in the sequence of g-projections must be a structural Case-assigner (in terms of the government-based Case theory which Kayne was assuming – see 2.2). Finally, the usual X' projections of X (that is, X' and XP) are also g-projections of X, and of whatever X is a g-projection of.

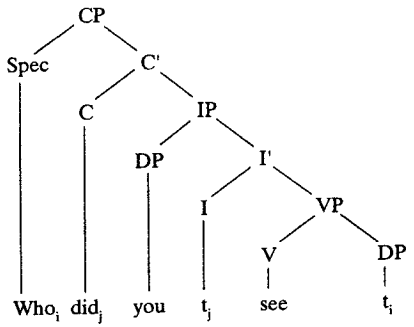
So g-projections 'start', as it were, from a structural Case-assigner and 'go up the tree', first following the X'-projection, and then, following the canonical government configuration in (77), going to the projection of 'the next governor up'. The fact that in (77) α can be the g-projection of something which δ is the g-projection of shows that g-projections can 'percolate up' a tree. In this way, the g-projection of a deeply embedded category can go all the way to the root of a tree.

As usual, this kind of abstract notion is much easier to understand in practice. So let's take a simple example of extraction of an object (ignoring for the moment the fact that object traces are always lexically governed):

(79) Who_i did you see t_i ?

Example (79) has the structure in (80):

(80)



The verb *see* is a structural Case-assigner. VP (and V', not indicated here) is a g-projection of V, since it is an X'-projection of V. I' is a g-projection of V, since I and VP are in a canonical government configuration. IP is also a g-projection of V, since it is a projection of I'. C' is a g-projection of V, since C and IP are in a canonical government configuration. And because C' is a g-projection of V, so is CP. So the g-projections of V go all the way to the root of the tree, CP. V canonically governs the trace, and the moved category *who* is connected to (forms a subtree with) a g-projection of V, namely CP. Hence, by the definition of antecedent-government given in (76), *who* antecedent-governs its trace.

We'll look at two main empirical domains that this system accounts for: cross-linguistic differences in Preposition-stranding, and the interaction of the Subject Condition with parasitic gaps.

4.3.3.2 Preposition-stranding

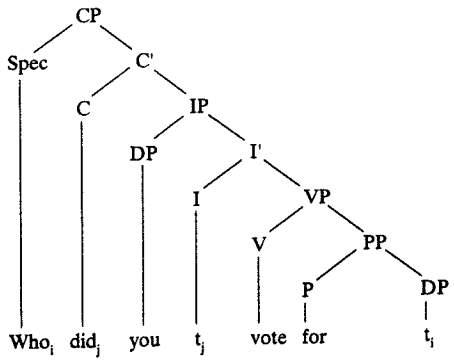
Preposition-stranding is of interest because of the cross-linguistic differences that are found. Compare the English and French sentences in (81):

(81a) Who$_i$ did you vote for t_i?

(81b) *Qui$_i$ as-tu voté pour t_i? (= (81a))

These sentences both have the structure in (82):

(82)



There is no independent reason to propose a structural difference between these two sentences, and yet the French one is bad while its English counterpart is good.

Let's begin by supposing that Prepositions are not lexical governors; this is not implausible, since Prepositions are in many ways like functional categories. So the trace in (82) has to be antecedent-governed in order to satisfy the ECP. In terms of the approach described above, we can see why the English sentence is good: P is a structural Case-assigner, PP is a projection and therefore a g-projection of P, and VP is a g-projection of P since V and PP are in a canonical government configuration. The rest is as described above for (79/80).

What about the French example? Here, Kayne's idea (which we mentioned briefly in Section 2.5.3) that English Prepositions are structural Case-assigners becomes crucial – or rather its inverse, namely that *French Prepositions are not structural Case-assigners*. Categories which are not structural Case-assigners are not able to project g-projections. Hence the trace of the stranded Preposition in (81b) cannot be antecedent-governed. If Prepositions are not lexical governors, then this trace fails the ECP altogether. So we are able to link the fact that English allows Preposition-stranding to our discussion of inherent Case in Section 2.5.3. This is an interesting comparative result.

4.3.3.3 Parasitic Gaps

Let's begin our discussion by looking at the Subject Condition. To see how Kayne's approach works here we have to put aside lexical government altogether. So let's just formulate the ECP as follows:

(83) Traces must be antecedent-governed

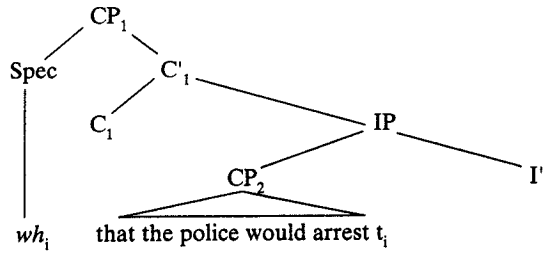
In later sections we'll return to the question of the status of lexical government. For the moment, we are entertaining the definition of antecedent-government given in (76) (this definition actually makes reference to the head that governs the trace, and so in a way subsumes lexical government). Now we can look at a typical Subject Condition violation:

(84a) * Which rock star$_i$ was that the police would arrest t_i expected?

(84b) ??Which rock star$_i$ were admirers of t_i arrested?

Looking first at (84a), you should be able to see that the subject CP *that the police would arrest t_i* is a g-projection of *arrest* – here the reasoning parallels what we saw for (79/80). But what about the rest of the structure? The relevant parts of the structure are as in (85):

(85)



Remember that there are basically two ways of forming a g-projection: either by normal X'-projection or by a canonical government configuration of the kind in (78). IP is clearly not an X'-projection of CP_2 in (85). And neither are CP_2 and I' in a canonical government configuration (because CP_2 is to the left of I'). Therefore the g-projection of *arrest* stops at CP_2. The moved *wh*-category isn't connected to CP_2, and so the trace fails to be antecedent-governed. Successive cyclic movement won't make any difference, as the intermediate trace of such movement in $SpecCP_2$ (see (35a)) is not antecedent-governed by the moved *wh*-element for the reason just given (although the intermediate trace antecedent-governs the trace in the complement position of *arrest*).

Example (84b) is treated in just the same way as (84a). The DP *admirers of* t_i is a g-projection of *of*, but the g-projection stops there. There is therefore no way for the moved *wh*-category to be connected to the governor of its trace, and so the trace fails to be antecedent-governed.

Now the really interesting aspect of Kayne's account concerns **parasitic gaps**. Parasitic gaps can be illustrated with the following kind of contrast:

(86a) ??Which book$_i$ did you write an essay before reading t_i ?

(86b) Which book did you buy t_i before reading t_i ?

Example (86a) is rather bad (in fact, it's an example of an Adjunct Island, a kind of island that we'll look at in more detail in the next section). But (86b) is much better. In it there are two gaps for a single *wh*-element: the object of *buy* and the object of *reading*. Each of these gaps is interpreted as a variable bound by *which book*, as the interpretation of (86b) clearly shows. The gap in object position of *buy* is usually referred to as the 'real gap', and the one in the object position of *reading* is known as the 'parasitic gap'. We'll write the gap – the one which is the trace of the actual movement – as *t*, and the parasitic gap as *e*.

Without going into an analysis of the Adjunct Condition, we can see from (86) that a parasitic gap is better than a real gap in an adjunct. Kayne shows that a parasitic gap in a subject inside an adjunct is not good:

(87a) ?How many books$_i$ have you read t_i [because you knew the authors of e_i] ?

(87b) *How many books$_i$ have you read t_i [because [$_{IP}$ the authors of e_i were famous]] ?

Let's suppose that the trace inside the adjunct can be antecedent-governed by *how many books* in (87a). The trace in (87b) fails to be antecedent-governed for the same reason as that in (84): IP is not a g-projection of the subject or of anything inside the subject. We can thus account for the ungrammaticality of (87b). Since the parasitic gap is not actually related to *how many books* by movement here, we should understand antecedent-government as a condition on the *representation* formed by movement, and not as a condition on movement, since we want both traces in (87) to be antecedent-governed by the moved *wh*-element, but only one of them is the actual trace of *wh*-movement.

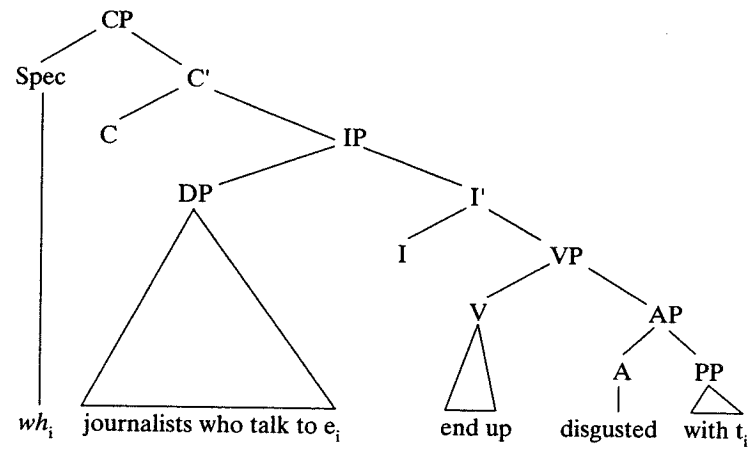
Kayne also discusses parasitic gaps inside subjects. These typically 'save' Subject Condition violations, as the following examples show:

(88a) ?Which rock star$_i$ do [$_{DP}$ journalists who talk to e_i] usually end up disgusted with t_i ?

(88b) *Which rock star$_i$ do [$_{DP}$ journalists who talk to e_i] usually have blackmail in mind ?

Example (88b) is a straightforward Subject Condition violation, and we have seen how Kayne handles this kind of case. Why is (88a) better? To see this, let's look a (simplified) tree for (88a):

(89)



One g-projection starts from *with*, and includes PP (by X'-theory), AP (canonical government), VP (canonical government), I' (canonical government), IP (X'-theory), C' (canonical government), and CP (X'-theory). So the *wh*-category is connected to *with*, the canonical governor of t_i. As for the parasitic gap, as long as we assume that the head of a relative canonically governs the relative clause, then the whole subject DP is a g-projection of *to*. Now, as usual, the g-projections of the category inside the subject stop here. However, there is already a subtree formed by the g-projections of *with* that the subject is connected to (forms a subtree with). And the g-projections of *with* are connected to the *wh* which binds the trace inside the subject. So, at one remove, the trace inside the subject is indeed connected to the *wh*-category. In a sense, then, the parasitic gap is truly parasitic, in that it depends on the real one in order to be antecedent-governed.

Finally, Subject Condition effects can be found with subject parasitic gaps:

(90a) ?Which rock star$_i$ do [teenagers that read articles about e_i] always try to imitate t_i ?

(90b) *Which rock star$_i$ do [teenagers to whom [stories about e_i] are told] always imitate t_i ?

The contrast here is clear. Example (90a) is another instance where the gap inside the subject is saved by the real gap, in the way described above. In (90b) the parasitic gap is beyond redemption, since its g-projections stop at the subject of the relative clause. As such they cannot be connected to the g-projections of the real gap, and so e_i fails to be antecedent-governed. The contrast in (90) (and between (90) and standard Subject Condition cases) is strong support for an approach like Kayne's.

4.3.4 Conclusion

In this section, we've introduced and elaborated the other major principle of locality, the ECP. The ECP requires that traces be properly governed, where proper government means either government by a lexical head other than P (lexical government) or government by the moved category (antecedent-government). Our definition of government is as follows:

Government:
α, a head, governs β iff:

(48a) α c-commands β

(48b) no barrier dominates β but not α

Barrier:
(49) any XP except IP

We've actually entertained two notions of antecedent-government, (51') and (76), which we repeat here:

Antecedent Government (Revised):
α, a moved category, antecedent-governs β iff:

(51a') α c-commands β

(51b') no more than one blocking category dominates β but not α

(51c') there is no filled C-position that minimally c-commands β and does not c-command α

Antecedent Government (Second Revision):
α, a moved category, **antecedent**-governs β iff:

(76a) α binds β

(76b) α is connected to a g-projection of γ, where γ canonically governs β

The distinction between lexical government and antecedent government is needed to account for argument-adjunct asymmetries of the kind we saw in 4.3.1. Condition (51b') is the same as subjacency (see 4.2). Condition (51c') is needed just for Complementizer-trace effects, as we saw in 4.3.2. Condition (76b) effectively imposes a requirement of government 'all the way along the path' from a moved element to its antecedent. Since the definition of government given in (48) refers to barriers, this implies that the notion of barrier may be relevant to the theory of locality.

Although the notions of lexical government and antecedent government do a lot of empirical work and have some theoretical depth, we clearly need to sort all this out. There are three principal issues that we need to address:

(91a) What is the correct definition of antecedent-government?

(91b) What is the relation between antecedent-government and subjacency?

(91c) What is the relation between antecedent-government and lexical government?

The next three sections are devoted to providing answers to these questions.

4.4 Barriers

This section summarizes the principal proposals in Chomsky's (1986b) monograph, *Barriers*. This was the first serious attempt to unify the theory of locality and so to provide answers to the questions raised at the end of Section 4.3.4. As the title suggests, the central notion is that of 'barrier'. We'll deal first with the proposed approach to subjacency, second with the ECP, and finally look at how the model captures the locality of head-movement and DP-movement.

4.4.1 Barriers and Subjacency

4.4.1.1 Adjunct Islands and the Condition on Extraction Domains

The central question in this section concerns the definition of the blocking categories for subjacency. Up to now, we've been assuming that these are IP and DP. But why should that be? In fact, we saw in 4.2.3 that in Italian the blocking categories are CP and DP. What is the total range of permitted variation in blocking categories and why are some categories BCs but not others? These are the questions we address here.

Our starting point is the Adjunct Condition. This is another island constraint, first noticed by Huang (1982). The observation is that extraction from inside an adjunct phrase is not allowed:

(92a) ?*Which bottle of wine$_i$ was Mick annoyed [BECAUSE KEITH DRANK t_i] ?

(92b) ?*Which dignitary$_i$ did the band leave the stage [WITHOUT BOWING TO t_i] ?

Many speakers find Adjunct Condition violations slightly less bad than

Subject Condition violations (or other island violations), which I've indicated here by '?*', as opposed to '*'. Here's (92a) with the BCs highlighted:

(92a') ?*[$_{CP1}$ Which bottle of wine$_i$ was [$_{IP1}$ Mick annoyed [$_{XP}$ because [$_{IP2}$ Keith drank t_i] ?

How subjacency applies depends on the precise analysis of the adjunct clause: what is XP in (92a') and what does it dominate? All the examples of embedded finite IPs that we've seen so far have been dominated by CPs, and so we probably want to say that there's a CP dominating IP_2 here (see 1.4.1.1). One possibility is that *because* is a complementizer. In that case, XP in (92a') is a CP and its Spec is available for successive-cyclic movement. And then there is no violation of subjacency here, as each step of movement just crosses one IP. We could block this by saying that *because* is in SpecCP, but there's really no reason to say this (except to give the right result for subjacency here); all the SpecCP elements we've seen up to now have been *wh*-elements, except for the special case of fronted XPs in verb-second languages (see 1.4.2.4). It seems more likely that *because* is a Preposition that selects a CP. In that case, there is a SpecCP position available for successive-cyclic movement and subjacency (formulated as in 4.2) cannot account for the violation here.

The fact that both subjects and adjuncts are islands is significant, since we saw in 4.3.1 that subjects and adjuncts are not lexically governed categories. Huang (1982), in fact, unified the Subject Condition and the Adjunct Condition under the following constraint, which he called the Condition on Extraction Domains (or CED):

Condition on Extraction Domains:

(93) No category can be extracted from a category which is not lexically governed

As we saw in 4.3.1, subjects and adjuncts are not lexically governed, while complements always are.

The CED is important both in theoretical and in empirical terms. From a theoretical point of view, it brings out the relationship between classical island effects of the type generally handled by subjacency and lexical government, and therefore the ECP. This suggests that a theory of subjacency of the type presented in 4.2 is missing something, since government plays no role in it. From an empirical point of view, the CED is important because it does not make reference to specific *categories* as BCs, but rather to specific *configurations*. It implies that any adjunct, whatever its category, will be an island (the implication has no real consequences for subjects, as they may well be always DPs in any case). This seems to be correct, as contrasts like the following show:

(94a) *Who$_i$ did you meet John [$_{AP}$ angry at t_i] ?

(94b) Who$_i$ did you make John [$_{AP}$ angry at t_i] ?

The category which linearly follows *John* in both examples in (94) is arguably an AP. In (94a), it is an adjunct (a secondary predicate); in (94b) it is selected by *make*. As you can see, extraction from an adjunct AP is much worse than extraction from an argument AP. It seems, then, that the function or the configuration of a category is more important for determining its islandhood than its actual category.

4.4.1.2 Defining Barriers

In many ways, the CED provided the cue for the development of the *Barriers* system. In this system, a chain of definitions is set up which begins with **θ-government**. We can think of θ-government as government by an element which assigns a θ-role to its governee. We can then define the following notions:

L-marking:

(95) α L-marks β if and only if α θ-governs β

L-marking plays a role comparable to lexical government, in that it distinguishes complements (L-marked) from subjects and adjuncts (not L-marked). Notice also that, if we stick to the basic idea that functional categories don't assign θ-roles, then the structural complements of functional categories are not L-marked.

Blocking category:

(96) α, an XP, is a BC for β if and only if α is not L-marked and α dominates β

This definition replaces the characterization of BCs as IP and DP that we gave in 4.2.1. We can now define barrier:

Barrier:

α is a barrier for β if and only if
either:

(97a) α is the first maximal category that dominates γ, a BC for β
or:

(97b) α is a BC for β, and α is not IP

Again, this definition supersedes the one we've been working with up to now. Definition (97) says that there are two ways in which something can be a barrier: either by inheritance from a BC it dominates (97a), or by simply being a blocking category, (97b). As in the simpler definition we've been using up to now, IP is excluded from the class of intrinsic barriers (remember that 'IP' means all V-related functional categories). The definition in (97) says that barrier is a relational notion rather than an absolute one – α is a barrier for β if the conditions given obtain.

Finally, we redefine subjacency in terms of barriers:

(98) In the following structure, α and β cannot be related by movement:
... α ... [$_B$ ... [$_B$ ... β ....] ..]
where α and β are separated by more than one barrier B

So now we should go through the class of island constraints, including the Adjunct Condition, in order to see how this particular approach works.

Before looking at any bad examples of extraction, though, we have to make sure that we don't rule out any good cases. So let's take a very simple example, where nobody would doubt that extraction is possible:

(99) Who$_i$ does Phil think [$_{CP}$ t'$_i$ Loretta likes t$_i$] ?

Does the first step of movement, from t_i to t'_i, cross any barriers? In fact, yes. In (99), the direct object is extracted out of VP. Nothing θ-marks VP, and so, according to the definitions we've given, VP is not L-marked. Therefore VP is a BC for the object trace and, by (97b), a barrier for the object trace. IP, which is the first XP dominating VP, is a barrier by inheritance; see (97a). So two barriers intervene between t_i and t'_i, and subjacency should be violated. Exactly the same is true of the matrix clause; here too, VP is not L-marked and so is a BC and a barrier, making IP into a barrier and creating a subjacency violation. So we have a problem; the definitions proposed give the wrong result in some very simple, well-formed examples. The same problem arises if we assume a split-Infl clause structure, as you should be able to see.

It's clear that the problem here has to do with VP. Chomsky proposes two solutions to this. One is to assume that there is an abstract sense in which I (or some element of the I-system) θ-marks, and so θ-governs, VP. In that case, the problem is solved as VP is L-marked. The difficulty with this solution is that it extends the notion of θ-assignment rather far; it is not easy to relate whatever semantic relation may hold between I and VP to notions such as Cause and Patient (see the discussion of θ-roles in 2.1).

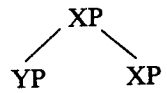
The other solution extends the notion of successive cyclicity one step further. Chomsky proposes the following two ideas:

(100a) Any category which doesn't receive a θ-role can be adjoined to

(100b) Only categories, not segments of categories, can be barriers

I briefly introduced the distinction between categories and segments earlier in 2.6.4. Let's recapitulate. In an adjunction structure like (101), the category adjoined to, XP, is divided into two 'segments'. Neither of these segments is the category itself, and therefore (100b) says that neither of them can be a barrier for YP:

(101)



If it's possible for YP to adjoin to XP, then, XP cannot count as a barrier for YP.

As we have seen, VP most probably is not θ-marked. Hence the object can adjoin to VP in (99). In that case, VP does not count as a barrier because neither of its segments can be a barrier. So, the true representation of (99) should be (99'):

(99') Who$_i$ does Phil [$_{VP}$ t'''$_i$ [$_{VP}$ think [$_{CP}$ t''$_i$ Loretta [$_{VP}$ t'$_i$ [$_{VP}$ likes t$_i$]]]]] ?

Here, each step of successive-cyclic movement is well-formed, as no barriers are crossed. Neither of the VPs are barriers because they can be adjoined to.

4.4.1.3 Barriers and the Island Constraints

Now we can look at the various islands. Let's begin with the CED. The CED bans extraction from categories that are not lexically governed: subjects and adjuncts. If we equate lexical government and θ-government, then we see that subjects and adjuncts are not L-marked, by (95). Since they are not L-marked, such categories are BCs, by (96). And hence, if they are not IP, they are barriers by (97b). More concretely, let's look again at the typical Subject Condition violation:

(102a) *Which rock star$_i$ was that the police would arrest t$_i$ expected ?

(102b) ??Which rock star$_i$ were admirers of t$_i$ arrested ?

Here, in each case the subject is not L-marked, and is therefore a BC (by (96)) and a barrier (by (97b) – subjects are never IPs). Because the subject is a barrier, the IP that immediately dominates it is a barrier by inheritance (97a). Therefore movement from within the subject to the matrix SpecCP crosses two barriers, and subjacency is violated. The option of adjunction to VP plays no role here, but we have to make sure that IP is not adjoined to. Since IP is not θ-marked, it could in principle be adjoined to, given what we said in (100a). To get around this, let's just assume that IP is never available as an adjunction site. In that case, there is no derivation of (102) that does not violate subjacency.

Next consider the Adjunct Condition. I repeat the example from (92a') here:

(92a') ?*[$_{CP1}$ Which bottle of wine$_i$ was [$_{IP1}$ Mick annoyed [$_{XP}$ because [$_{IP2}$ Keith drank t$_i$] ?

Here XP, whatever its category, is not L-marked, and hence is a BC and a barrier. We've been assuming that adjuncts are adjoined to VP. In that case, IP is a barrier by inheritance. So extraction from an adjunct violates subjacency. However, there remains the possibility of adjunction to XP. Since XP is not assigned a θ-role, this is possible, in which case XP will not be a barrier and IP will not be a barrier by inheritance. This is a problem for the approach Chomsky gives in *Barriers*.

Consider next *wh*-islands. Here we're only concerned with extraction of arguments; we'll look at extraction of adjuncts in the next section. The relevant examples were given in (36):

(36a) ?[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ to fix t_i t_j] ?

(36b) ?*[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ you should fix t_i t_j] ?

For the version of subjacency that defines IP and DP as barriers, the crucial problem with these examples is that *whose car* has to cross two IPs on its way to SpecCP$_1$, since *how* occupies SpecCP$_2$. On the other hand, the *Barriers* system, as I've presented it so far, is unable to rule out these examples. IP can only be a barrier by inheritance, and so, given the option of VP-adjunction as described above, neither IP is a barrier in (36). The full representation of (36a), taking VP-adjunction of *whose car* into account, is (36a'):

(36a') ?[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you [$_{VP1}$ t''$_i$ [$_{VP1}$ wondering [$_{CP2}$ how$_j$ [$_{IP2}$ to [$_{VP2}$ t'$_i$ [$_{VP2}$ fix t_i t_j]]]]]]] ?

Let's see what the barriers here are: the segments of VP$_2$ are not barriers by (100b); IP$_2$ is not a barrier because IP can only be a barrier by inheritance; CP$_2$ is not an intrinsic barrier because it's L-marked by *wonder*, but it is a barrier by inheritance from the non-L-marked IP$_2$. Finally, neither segment of VP$_1$ can be a barrier.

Chomsky concludes that this is not necessarily a bad result, given that the subjacency effect here is quite weak. Crossing one barrier gives a certain awkwardness, but not real ungrammaticality. To account for the contrast with (36b), Chomsky suggests that the most deeply embedded tensed IP can be an inherent barrier. Thus, if IP$_2$ is a barrier in (36b), and CP$_2$ is a barrier by inheritance, movement from the lower clause to the higher clause violates subjacency. It may be possible to account for Rizzi's evidence that Italian chooses different barriers (see 4.2.3) with the idea that tensed IP is never a barrier in Italian. In that case, *wh*-island configurations can be readily extracted from.

Next, the CNPC. Here, once more, are the typical examples:

(33a) *Which band$_i$ did you write [$_{DP}$ A SONG WHICH [$_{IP}$ WAS ABOUT t_i]] ?

(33b) *Which band$_i$ did you believe [$_{DP}$ THE CLAIM THAT [$_{IP}$ WE HAD SEEN t_i]] ? (see (9))

The relative-clause example in (33a) is straightforwardly handled by the *Barriers* system. The head of a relative clause, *a song* in this example, does not θ-govern the CP that modifies it. Therefore, this CP is not L-marked, and so is both a BC and a barrier. This CP is dominated either by NP or by DP (which one depends on exactly what analysis we give to relative clauses, and that's a question I don't want to go into here), which is thus a barrier by inheritance. And so movement out of a relative clause always violates subjacency.

In (33b) no barriers are crossed by movement out of the complement to *claim*. The CP *that we had seen* t_i is the complement of *claim*, and so is L-marked and therefore neither a BC nor a barrier. The entire complex DP is the complement of the main verb *believe*, and so is L-marked. The NP headed by *claim* can presumably be adjoined to, since it is not θ-marked. So, as with *wh*-islands, no subjacency violation is predicted here. Unlike *wh*-islands, however, the possibility of designating tensed IP as a barrier makes no difference, since the SpecCP of the complement is available for successive-cyclic movement.

The Left Branch Condition in its simplest form cannot be handled by the *Barriers* system. That is, there is no account for the ungrammaticality of examples like the following:

(18) *Whose$_i$ did you play [$_{DP}$ t_i guitar] ?

DP is L-marked here, and so is neither a BC nor a barrier. We suggested in 4.1 that the ungrammaticality of examples like (18) may be due to a condition on pied-piping that is peculiar to English, and so this is not really a problem – in fact, we want our theory of locality to allow the examples that we looked at from other languages, like (23) and (24).

On the other hand, extraction of the left branch of a left branch, as in (21a), is ruled out:

(21a) *Whose$_i$ did you play [$_{DP1}$ [$_{DP2}$ t_i friend] 's favourite guitar]] ?

Here DP$_2$ is not L-marked, and so is a BC and barrier, and DP$_1$ is a barrier by inheritance. So movement of *whose* violates subjacency here. (Again, adjunction to DP$_2$ might be a possibility, with the result that neither DP$_2$ nor DP$_1$ would be a barrier. However, it may be that possessor DPs are θ-marked – perhaps by *'s* – but not θ-governed by anything in DP$_1$ since they occupy Spec,DP; in that case, adjunction to possessors would be impossible.)

Finally, the Coordinate Structure Constraint can't really be handled by the Barriers system. Let's look again at the examples we saw earlier:

(15a) *What$_i$ did Bill buy POTATOES AND t_i ?

(15b) *What$_i$ did Bill buy t_i AND POTATOES ?

(15c) *Which guitar$_i$ does KEITH [PLAY t_i] AND [SING MADRIGALS] ?

(15d) *Which madrigals$_i$ does KEITH [PLAY THE GUITAR] AND [SING t_i] ?

We could in fact handle (15a) if we assume that *and* t_i is a non-L-marked XP (variants of this idea have been proposed by Munn (1993), Thiersch (1993), and Kayne (1994)). This XP would be a BC and a barrier, making the whole conjunct a barrier; and so subjacency would be violated here. Example (15b) cannot be accounted for along these lines, however. In (15c and d) it is natural to assume that the conjoined categories are not L-marked, and so are BCs and barriers. However, if they are not arguments, they can be adjoined to, and so the fact that they are barriers does not create a subjacency violation.

4.4.1.4 Conclusion

To summarize, in this section I've outlined the *Barriers* system, which consists of the definitions in (95–97), the formulation of subjacency in (98), and the further assumptions in (100). This system provides a deeper characterization of blocking categories than the earlier version of subjacency described in 4.2, and successfully accounts for the CED, the relative-clause case of the CNPC, and, with one further assumption, *wh*-islands. The possibility of adjoining to a potential barrier and moving on raises problems for the analysis of the Adjunct Condition and the relative-clause case of the CNPC. However, we have an alternative possibility for preventing VP from being a barrier in simple cases like (99): that of assuming that VP is θ-governed by I. If we assume this, we may be able to simply drop the idea that successive-cyclic adjunction can neutralize potential barriers, as examples like (99) are the main cases for which this idea is needed. For the moment, let's leave this issue open. We'll come back to it below.

4.4.2 Barriers and the ECP

In the previous subsection we saw the approach to subjacency and the characterization of blocking categories in the *Barriers* system. In this subsection, we turn to the ECP. Remember that what we're interested in is seeing how we can arrive at a unified characterization of subjacency and the ECP.

The version of the ECP that Chomsky adopts in *Barriers* is what we saw in 4.3.1.1. The ECP requires that traces be properly governed, where proper government is defined as either θ-government or antecedent-government. If we take θ-government to be broadly the same as lexical government, this corresponds to the definition of proper government that I gave in (47) of 4.3.1.1:

Proper Government:

(103) either: government by a lexical head (lexical government)
or: government by the moved category (antecedent-government)

Government is defined in terms of barriers, as follows:

Government (Revised):

(104) β governs α if and only if there is no barrier for α that dominates β but not α

You'll notice that this is slightly different from the definition of government we've been working with up to now. It follows from (104), combined with the definition of antecedent-government in (47) that antecedent-government is blocked by a single barrier. We see, then, that subjacency and antecedent-government are closely related: subjacency is blocked by more than one barrier; antecedent-government by a single barrier. So the approach to antecedent-government adopted in *Barriers* is like the one we assumed in (51) of 4.3.

Given these assumptions, the account of argument-adjunct asymmetries

that emerges is very similar to the one described in 4.3.1.1. Let's look again at the typical contrast:

(105a) ?*[$_{CP1}$ Whose car were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ you should fix t_i t_j] ?

(105b) *How$_j$ were you wondering [$_{CP}$ whose car$_i$ you should fix t_i t_j] ?

As we saw in 4.3.1.1, the contrast here is between the awkwardness of (105a) and the uninterpretability of (105b). Although both sentences are bad, the clear difference in their status suggests that different principles are at work in each case. In (105a), subjacency is violated by the extraction of *whose car* from the *wh*-island (allowing that the finite embedded IP_2 is a barrier – see subsection 4.4.1.3). The trace of *whose car* satisfies the ECP, however, since it is θ-governed by *fix*. In (105b), subjacency is violated in exactly the same way as in (105a); however, here the ECP is violated since the adjunct is not θ-governed, and fails to be antecedent-governed because two barriers intervene between *how* and its trace: CP_2 (a barrier by inheritance from IP_2) and IP_2 (an inherent barrier if tensed).

What about the non-finite counterpart of (105b)?

(106) *How$_j$ were you wondering [$_{CP2}$ whose car$_i$ [$_{IP2}$ to fix t_i t_j]] ?

Here IP_2 is non-finite, and so it is not an inherent barrier. Nevertheless, it is not L-marked and so is a BC for material it dominates. It is not a barrier, for the simple reason that we have made an exception for IP – IP can only be a barrier by inheritance. However, the fact that IP is a BC means that CP, the first XP dominating it, is also a barrier by inheritance. So CP_2 is a barrier, which means that *how* cannot antecedent-govern its trace here. This accounts for the ungrammaticality of (106).

Subjects are like adjuncts in typically not being θ-governed. And so an example like (53), repeated here, is ruled out because two barriers – tensed IP and CP – intervene between the trace and *which band*:

(53) *Which band$_i$ were you wondering whether t_i will play that song ?

As we saw in 4.3.2.1, there is a further complication with subjects: C-t effects, as in (58):

(58) *Who$_i$ did you say that t_i wrote this song?

We accounted for (58), and the contrast with the corresponding sentence without *that*, with clause (51'c) of the definition of antecedent-government:

α, a moved category, antecedent-governs β iff
(51c') there is no filled C-position that minimally c-commands β and does not c-command α

As I remarked in 4.3.2.1, although this does the job of accounting for C-t effects, it rather muddies the waters of the definition of antecedent-government.

The approach adopted in *Barriers* consists of defining a further kind of barrier, a 'minimality barrier':

Minimality Barrier:
(107) β is a minimality barrier for α if and only if:
β is an X' (other than I')
β dominates α
the head of β is lexical

Let's look more closely at (58):

(58') *Who$_i$ did you say [$_{CP}$ t'$_i$ [C' that t$_i$ wrote this song]] ?

C' fulfills the requirements for being a minimality barrier for t_i here: it is an X' (but not I'), it dominates t_i, and its head C is lexical. And so t'_i in SpecCP fails to antecedent-govern t_i, owing to the intervening minimality barrier C'. Since subjects are not θ-governed, t_i violates the ECP. Of course, if *that* is not present, then C' does not fulfil clause (c) of the definition of minimality barrier in (107) and t'_i can then antecedent-govern t_i. Minimality barriers are only relevant for government, not movement; hence the notion of 'barrier' that is relevant for subjacency does not include minimality barriers, while the one that applies to antecedent-government does, as we have just seen.

Remember that adjunct traces are not subject to C-t effects, as (68), repeated here, shows:

(68) How$_i$ did you say [$_{CP}$ t'$_i$ (that) [$_{IP}$ he fixed your car t$_i$]] ?

In *Barriers*, Chomsky takes over from Lasnik and Saito (1984) the assumption that the antecedent-government requirement applies to adjunct traces only at LF. Since *that* can delete at LF (as it has no semantic content) its overt presence makes no difference to the status of an adjunct trace. Example (68) is allowed with or without an overt *that* because *that* is in any case missing at LF. This account also extends to the absence of C-t effects in null-subject languages. If we follow Rizzi's proposal that in examples like (70) the subject is extracted from postverbal position, and assume that this position is an adjunct position, then (70) has the same status as (68):

(70) Chi hai detto che ha scritto questo libro?
Who have-you said that has written this book?
'Who did you say wrote this book?'

In conclusion, the *Barriers* version of the ECP covers the principal phenomena that we looked at in 4.3.1 and 4.3.2 (it doesn't handle the 'connectedness' phenomena discussed in 4.3.3, however). There is a partial unification with subjacency: antecedent-government is partly a stricter version of subjacency, since it bans the crossing of even one barrier. The ECP also makes reference to θ-government, which is indirectly implicated in the definition of BCs and barriers. However, the ECP is also sensitive to minimality

barriers, while subjacency just is not; and this is the principal conceptual flaw in the attempted unification. I'll present a different version of minimality condition in 4.5, and return to the question of unifying it with subjacency in 4.6.

Before going on to those matters, however, let's reconsider the relations between *wh*-movement and the other kinds of movement, a topic that we have left to one side over the last few sections.

4.4.3 Barriers, DP-Movement, and Head-Movement

Here I want to take a brief look at how the *Barriers* system can account for the properties of DP-movement and head-movement.

4.4.3.1 DP-movement

Here we want to allow straightforward examples like those in (108) and rule out cases of 'super-raising' and 'super-passive' like (109):

(108a) John$_i$ seems t$_i$ to speak Chinese

(108b) The students$_i$ were beaten up t$_i$ (by the police)

(109a) *The train$_i$ seems that it is likely t$_i$ to be late

(109b) *The students$_i$ seem that it was told t$_i$ that there would be extra reading

In (108a), movement of *John* crosses the lower IP and the VP headed by *seems*. Here IP is L-marked by *seems*, since raising verbs select IP (see 2.3.3). VP, however, is a BC and a barrier since it is not L-marked. The subject of a complement infinitive is not θ-governed by the higher verb – we saw in 2.3.3 that raising verbs do not assign θ-roles to the subject of their infinitival complement. If t_i isn't θ-governed here, it must be antecedent-governed. But the VP barrier intervenes between the trace and its antecedent.

At this point, it's useful to compare (108a) with an example of extraction of the subject of an ECM infinitive, like (54):

(54) Which band$_i$ did you consider t$_i$ to be the best ?

In 4.3.1.1, we said that the trace here is lexically governed by *consider*. However, this is a case where the switch from lexical government to θ-government has consequences: we clearly don't want to say that *consider* θ-governs the trace here. So the trace in (54) must satisfy the ECP by being antecedent-governed. Example (54) is well-formed, which means that there are no barriers at all between *which band* and its trace here. The lower IP is not a barrier, since it is non-finite. The VP headed by *consider* is a barrier, but it can be adjoined to and neutralized in that way (see 4.4.1). The matrix IP is not a barrier. So we correctly allow (54) without assuming that the trace is θ-governed. (Here and elsewhere we have to ignore the possibility that V' is a minimality barrier. Chomsky's assumption was that V' was simply

not present; since this work predated the VP-internal subject hypothesis (see 2.3.4), this was tenable. I'm just going to leave this question to one side.)

Now look again at (108a). The structure of the example is broadly comparable to (54), and we want to see here too how the trace is antecedent-governed. Again, the lower IP poses no problem. But what about VP? The option of VP-adjunction is not available for DP-movement. Remember that VP-adjoined positions are A' (or non-L-related) positions (see 3.3.2.3). Movement from an A-position to an A'-position and back to an A-position is not allowed – this is referred to as 'improper movement'. We can rule out improper movement with Principle C of the binding theory if we suppose that traces in A'-positions are subject to this Principle. If we adjoin to VP and move on, a trace is left in the adjoined position. But since DP-movement moves to a Case position such as the matrix SpecIP, and since Case positions are A-positions, then the VP-adjoined trace will be A-bound. In that case, it will violate Principle C. In other words, movement from A-positions to either A- or A'-positions is allowed, and movement from A'-positions to other A'-positions is allowed, but movement from A'-positions to A-positions is NOT allowed.

If DP-movement in (108a) cannot adjoin to VP, then VP should be a barrier. The fact that the example is well-formed tells us that this can't be right. To solve this problem, Chomsky makes two proposals, both of which we have already seen in slightly different guises. First, there is a relation between V and I which we can think of as either overt or covert movement (see 1.4 and 2.6.5). Second, there is a relation between I and SpecIP: Spec-head agreement which facilitates feature-checking and/or Case-assignment (see Chapter 2, especially 2.6). We can indicate each of these relations by coindexing, and this gives the following representation for (108a):

(108a') John$_i$ I$_i$ seems$_i$ t$_i$ to speak Chinese

The coindexed elements here form a chain (see 3.5: each element c-commands and is coindexed with the next). Moreover, by the definition of government given in (104), *John* governs I and *seems* governs t_i. Allowing that there is a 'special relationship' between I and V which facilitates government (this seems legitimate, since we know that V-to-I movement is possible in many languages: see 1.4.1.1), then we have a chain where each element governs the next. We can in fact think of this as antecedent-government, although this is an extension of the strictly movement-based definition that we gave in (47). If we see this chain-relation as a species of antecedent government, then it follows that the trace is antecedent-governed in (108a). The crucial step in all of this involves the V–I relation. As I've already mentioned, at the price of an implausible extension of the intuitive content of θ-theory, we can simplify this picture by saying that I θ-marks VP, and therefore that VP is not a BC or a barrier.

We have no problems allowing simple passive sentences like (108b). The

trace in the direct-object position is θ-governed by the verb, and so satisfies the ECP. Passives of ECM verbs like (110) are just like raising structures:

(110) George$_i$ is widely believed t$_i$ to be best

Here the considerations just raised in connection with (108) come in.

What about (109)? In (109a), we want to prevent the formation of a chain of antecedent-governors comparable to the one that allows (108a). This can be done if we take it that the coindexing that indicates Spec-head agreement and the head–head relation between I and V can only arise when those relations hold: that there is no such thing as 'accidental coindexation'. In that case, there is no way for a chain to be formed which connects the train in the matrix SpecIP with its trace in the lowest SpecIP. More concretely, we have the following relations in (109a):

(109a') *The train$_i$ I$_i$ seems$_i$ that it$_j$ is$_j$ likely$_{(j)}$ t$_i$ to be late

The train is coindexed under Spec-head agreement with I. I is coindexed with V, by their special relationship, and of course the trace of *the train* bears the same index. However, the trace is too far from *seems* to be governed by it. Hence it fails to be antecedent-governed, and, since it is not θ-governed, violates the ECP. There is another chain formed by the elements bearing the j-index here, but, even if *likely* shares that index with *is* (which is perhaps doubtful), it is of no help to the trace, since the trace hears a distinct index and there is no way of establishing any kind of equivalence between them.

Exactly the same considerations apply in (109b), as you should be able to see. So the trace of *the students* is not antecedent-governed here. However, this trace should be θ-governed since it is a direct object, and so there should be no problem. The ungrammaticality of (109b) implies that θ-government may not always be enough for well-formedness: the antecedent-government requirement seems to hold anyway. We'll come back to this point in the next section.

4.4.3.2 Head-movement

Turning now to head-movement, what we'd ideally like to do is to derive the Head Movement Constraint (HMC: see 1.4.2.3(67)):

The Head Movement Constraint

(111) A head X can only move to the most local c-commanding head-position

As we saw in 1.4.2.3, the HMC prevents V from moving over I to C in one step, and so it rules out examples like:

(112) *Have$_i$ he could t$_i$ done it ?

Here I'm assuming that *have* heads its own VP, and takes another VP headed by *done* as its complement. This is where it becomes crucial not to assume

that I L-marks VP. If it does, then neither VP nor IP are barriers, and so (112) should be allowed. If I doesn't L-mark VP, on the other hand, VP is a BC and a barrier. IP is then a barrier by inheritance, and (112) violates both antecedent-government and subjacency (and it's safe to assume that V is not θ-marked by I). The account of (112) also appears to rule out the possibility of V-to-I movement, which is amply attested crosslinguistically (see 1.4.1.1). In the discussion of DP-movement above, I appealed to a 'special relationship' between V and I. Chomsky's proposal is that I θ-marks and therefore θ-governs VP (but not V), but that I does not L-mark VP because only lexical elements can L-mark. So we change the definition of L-marking as follows:

L-marking:

(95') α L-marks β if and only if α is a lexical category that θ-governs β

This prevents I from L-marking VP, and so VP is a BC and a barrier, as usual. However, the operation of raising V to I creates a lexical I, and such an I can then L-mark VP. Hence, V-to-I movement licenses itself, as it were. This does not affect our conclusions above regarding (112), assuming that a modal like *could* is unable to L-mark VP.

In general, I-to-C movement of the kind that creates inversion and verb-second structures (see 1.4) is allowed, since IP is not an inherent barrier. Movement from I to a position higher than C is ruled out, since CP will inherit barrierhood from IP. (Remember that throughout all of this I'm using 'I', 'IP', and so on as shorthand for the full functional structure that makes up the clause. All these categories – AgrSP, TP, AgrOP, and the others – are presumably 'defective' as barriers in the way IP is, and are similarly unable to L-mark their complement unless V moves into them. Adopting a split-Infl structure doesn't really alter anything of substance in what's gone before, but it does mean that the reality is (even) more complex than what we have seen.)

In a very important work, Baker (1988) showed that incorporation structures of the kind found in many (usually non-Indo-European) languages involve head-movement. Mohawk Noun-incorporation is an example:

(113a) Yao -wir -a?a ye- -nuhwe? -s ne ka -NUHS -a?
PRE -baby- SUF 3SgF/3N -like -ASP the PRE -house -SUF
'The baby likes the house'

(113b) Yao -wir -a?a ye- -NUHS -nuhwe? -s?
PRE-baby -SUF 3SgF/3N -house -like -ASP
'The baby likes the house.'

If incorporation is head-movement, it must be subject to the conditions on movement, in particular the ECP. We can therefore predict that incorporation from subjects and from adjuncts is impossible, since such incorporation would involve crossing a barrier, with the result that the trace would not be antecedent-governed and the ECP would consequently be violated (it seems

reasonable to think that heads are never θ-marked and so never θ-governed). In a survey of a wide array of languages Baker showed that there are no cases of noun-incorporation of subjects or adjuncts. In other words, we don't find examples of incorporation that would look like the following:

(114) *Ye -WIR -nuhwe? -s ne ka -nuhs -a? (Mohawk; see (113))
3Sgf/3N -baby -like the PRE-house -SUF
'The baby likes the house'

(115) *Nofo ana a ia (he) (Niuean)
live cave ABS he (in)
'He cave-lives (in)'

Baker extends his theory of abstract incorporation to many other construction-types that are found in the world's languages. In each case, what he shows is that the head-to-head dependency can only hold between a head and its immediate complement. For example, Baker argues that one kind of causative construction involves V-to-V incorporation. This kind of causative is found in Romance, so we have:

(116) J'ai fait laver la voiture à Jean (French)
I've made wash the car to John
'I made John wash the car'

According to Baker, there is a complex verb *faire-laver* here. This complex verb is rather like verbs such as 'give' in having a direct object (*la voiture*) and an indirect object (*à Jean*). Most importantly, *faire* can only form a complex verb with a verb in its immediate complement. Again, the ECP explains this pattern: head-movement from out of a subject or an adjunct crosses a barrier and so leads to a violation of the ECP. Baker's results have made possible the development of a theory of incorporation, causatives, and many other phenomena which is fully integrated into linguistic theory. Most of the predictions that this theory makes come from the ECP.

In this section we've seen that, once certain points are clarified, the *Barriers* system can capture the basic locality constraints on DP-movement and head-movement. For DP-movement, we need to extend the notion of antecedent-government so that it relates to chains of various kinds, not just movement chains. We also saw that θ-government may not be enough for DP-traces. For head-movement, we see that the ECP subsumes the HMC, but we need a special account of why ordinary V-to-I raising is possible.

4.4.4 Conclusion

The *Barriers* framework was the first attempt to unify fully the theory of locality. We have looked at the system in some detail, and we can see what its strong and weak points are. The principal strength of the system is the central role played by the notion of 'barrier': this notion is crucial for defining government, the ECP, and subjacency. Empirically, most of the coverage of the earlier formulation of subjacency is retained (except for *wh*-islands and

the complement case of the CNPC); similarly, leaving aside the 'connectedness' effects discussed in 4.3.3, *Barriers* successfully handles most of what the ECP did. As we've just seen, the approach also extends to DP-movement and head-movement. In the case of DP-movement, there are some difficulties. In the case of head-movement, the approach makes Baker's important crosslinguistic results possible.

Conceptually, however, a number of questions remain:

- What is the precise nature of the minimality condition, and why is it restricted to government and not movement relations?
- Why is θ-government seemingly 'not enough' for DP-traces?
- Why is the 'I-system' defective, in only being a barrier by inheritance rather than inherently?
- Why are adjunct traces only subject to the ECP at LF, exempting them from the full effects of the minimality condition?

Over the next two sections, answers to most of these questions will emerge. The next section focuses almost exclusively on the ECP, in that it reviews an influential alternative version of the minimality condition: Rizzi's Relativized Minimality.

4.5 Relativized Minimality

In this section I'll summarize the main points of Rizzi's (1990) theory of relativized minimality. As its name implies, the central idea is that the minimality condition shouldn't be seen as an absolute condition, but rather that each kind of movement relation – *wh*-movement, DP-movement, and head-movement – is associated with its own kind of minimality condition and is blind to the others. The relativized minimality condition is the central facet of the theory of antecedent-government. In addition to this, Rizzi proposes that θ-government can be altogether dispensed with and replaced with a simpler head-government condition. We'll look first at the proposal for antecedent-government (4.5.1) and then at the proposal for head-government (4.5.2).

4.5.1 Relativized Minimality and Antecedent-Government

4.5.1.1 A'-specifiers block A'-movement

Let's begin by looking, yet again, at the argument-adjunct asymmetries that are found in *wh*-islands:

(105a) ?*[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ you should fix t_i t_j] ?

(105b) *How$_j$ were you wondering [whose car$_i$ you should fix t_i t_j] ?

Only subjacency is violated in (105a), while the ECP is also violated in (105b) since the adjunct trace must be antecedent-governed and fails to be so. Rizzi retains the account of argument extraction: complement traces are θ-governed and therefore pass the ECP – the relatively mild ill-formedness of (105a) is due to subjacency. Rizzi's observation is that in (105b) a distinct *wh*-element, *whose car*, intervenes between the moved adjunct *wh*-element and its trace. In an obvious sense, like is interfering with like here.

As we have seen, alternative accounts have been proposed for these argument-adjunct asymmetries. However, there are other cases where like appears to block like which are much harder to account for in terms of what we have seen up to now. One such case is the phenomenon of pseudo-opacity found in French. French allows certain kinds of quantifiers that are found some distance to the left of a direct object to be interpreted as quantifying that direct object, as in the following pair:

(117a) Jean a consulté beaucoup de livres
John has consulted much books.
'John has consulted many books'

(117b) Jean a beaucoup consulté de livres
John has much consulted books.
'John has consulted many books'

As we saw in 4.1, French also has one quantifier that can optionally violate the Left Branch Condition, *combien* ('how many'). This element allows either left-branch extraction or pied-piping of the quantified DP:

(118a) Combien$_i$ a-t-il consulté t_i de livres?
How-many has-he consulted of books?
'How many books has he consulted?'

(118b) Combien de livres$_i$ a-t-il consultés t_i ?
How-many of books has-he consulted?
'How many books has he consulted?'

In (118b), *combien de livres* is the direct object, and so is extractable like any argument. It's reasonable to think that *combien* alone is an adjunct, since it's not θ-marked by anything. Now, if we combine the 'leftward' *beaucoup* of (117b) with extraction of *combien*, we see that only the pied-piping variant is allowed:

(119a) Combien de livres$_i$ a-t-il beaucoup consultés t_i ?
How-many of books has-he much consulted?
'How many books has he consulted a lot?'

(119b) *Combien$_i$ a-t-il beaucoup consulté t_i de livres?
How-many has-he much consulted of books?

If we take it that *beaucoup* occupies an A'-position, presumably adjoined to VP, in these examples, then we see that, as in the case of *wh*-islands, the presence of an A'-element blocks extraction of an adjunct to an A'-position.

The third case is a further kind of island phenomenon that was originally discussed by Ross, but which we haven't introduced yet: negative islands. The basic observation is that clausal negation blocks extraction of adjuncts:

(120a) Which car$_i$ didn't he fix t$_i$? (Answer: The blue GTi)

(120b) *?How$_i$ didn't HE FIX YOUR CAR t$_i$? (*Answer: With a spanner)

The argument-adjunct asymmetry is very clear here. Rizzi also gives examples which show that the adjunct must be in the scope of negation in order for the island effect to apply. First we need to look briefly at some scope ambiguities involving negation (the concepts of 'scope' and 'scope ambiguity' and their relevance for syntactic theory are explained in the Appendix to this chapter). For example, think about the following sentence:

(121) They don't exploit their fans [because they love them]

Example (121) is ambiguous in at least two ways (in fact, there are more possible interpretations, but they are less obvious and rather harder to see, so I'll just leave them aside). It can either mean 'Because they love them, they don't exploit their fans': only the main clause is negated, not the adjunct clause. Call this the 'main-clause negation' reading. This is probably the most natural one for (121), given the lexical items and ordinary pragmatic considerations. However, (121) also has the interpretation which we can paraphrase as 'It's not because they love them that they exploit their fans (but for some other reason)'. Here the main predicate isn't negated, just the adjunct. So call this the 'adjunct-negation' reading. So we see that, at least with *because*-type adjuncts, clausal negation allows at least two different scopes.

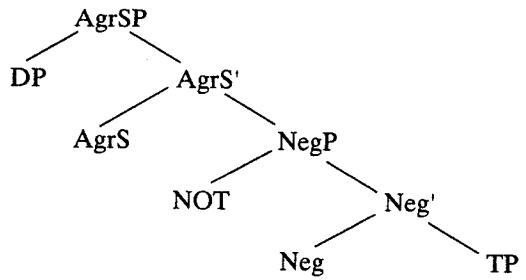
Now let's look at what happens if we extract a *because*-adjunct (using *why*, which naturally invites an answer with *because*):

(122) Why$_i$ don't they exploit their fans t$_i$?

Example (122) only allows the main-clause negation interpretation. In other words, it can be answered only with something like 'Because they love them'. To put it another way, the trace of *why* can't be interpreted as being in the scope of negation. This is due to the negative island: negation blocks clausal extraction. The usual assumption about scope is that it is determined by c-command (see the Appendix to this chapter), so what the scope restriction on the interpretation of (122) tells us is that negation blocks the extraction of adjuncts within its c-command domain.

When I introduced the split-Infl hypothesis in 1.4.3, I briefly mentioned NegP. Suppose that negative clauses feature a NegP projection that is situated above TP and below AgrSP (there is some evidence that the position varies across languages, but I'll gloss over this):

(123)



(Here, NOT refers to whatever element carries clausal negation: *not, pas, nicht*, and so forth.) It is natural to think that the position that NOT occupies isn't a GF-position. In other words, it is a non-L-related or A'-position. In that case, the negative-island effect that we saw in (120b) and (122) is a third instance of an A'-element blocking the extraction of a *wh*-adjunct.

The generalization that unifies these three cases – *wh*-islands, pseudo-opacity in French, and negative islands – is:

(124) A filled A'-position specifier α blocks antecedent government between an A'-position β that c-commands α and an adjunct trace that α c-commands.

To put it schematically, α prevents β from antecedent-governing *t* in (125), where *t* is an adjunct trace and each element c-commands the next:

(125) [... β ... [α [.... t ...]]]

In a *wh*-island case like (105b), α is a *wh*-element; in pseudo-opacity examples like (119b), α is an adverb like *beaucoup*, and in negative islands α is NOT.

4.5.1.2 A-specifiers block A-movement

Now let's look at the basic DP-movement configuration that violates the ECP – super-raising and super-passive:

(109a) *The train$_i$ seems that it is likely t$_i$ to be late

(109b) *The students$_i$ seem that it was told t$_i$ that there would be extra rations

Since this is DP-movement, these examples involve movement to an L-related A-position; in fact, in both of those examples the landing-site is a subject position. Now, in between the moved DP and its trace, there is another subject position occupied by *it*. Movement to this position will give well-formed results:

(126a) It seems that the train$_i$ is likely t$_i$ to be late

(126b) It seems that the students$_i$ were told t$_i$ that there would be extra rations

What goes wrong in (109), then, can be thought of as an element in an A-position blocking movement of another element to a higher A-position. In other words:

(127) A filled A-position specifier α blocks antecedent government between an A-position β that c-commands α and a DP-trace that α c-commands

The similarity between (127) and (124) should be clear. Before attempting to unify them into a single principle, however, let's look at head-movement.

4.5.1.3 Heads Block Head Movement

The typical Head Movement Constraint violation looks like (112) from the previous section:

(112) *Have$_i$ he could t$_i$ done it ?

We can rephrase the Head Movement Constraint as follows:

(128) A head α blocks antecedent government between a head β that c-commands α and a head-trace that α c-commands

Statement (128) is little more than a restatement of the HMC as given in (67) of Chapter 1 and in (111) above. Again, there is an obvious formal similarity with the generalizations in (124) and (127).

4.5.1.4 Relativized Minimality

Relativized minimality unifies (124), (127), and (128), as follows:

Relativized Minimality:
X antecedent-governs Y only if there is no Z such that:

(129a) Z is a typical potential antecedent governor for Y

(129b) Z c-commands Y and does not c-command X

Now we need to know what a typical potential antecedent governor is:

Typical Potential Antecedent Governor:

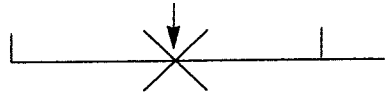
(130a) Z is a typical potential antecedent governor for Y, Y in an A-chain = Z is an A-specifier c-commanding Y

(130b) Z is a typical potential antecedent governor for Y, Y in an A'-chain = Z is an A'-specifier c-commanding Y

(130c) Z is a typical potential antecedent governor for Y, Y in a head-chain = Z is a head c-commanding Y

Given the notion of typical potential antecedent governor as defined here, relativized minimality says that antecedent-government will be blocked in the following configuration:

(131) [.... X$_i$ .... [.... Z .... [... Y$_i$ ....]]]



C = (X$_i$, Y$_i$)

Where C is an chain of the same type (A, A', head) as the position occupied by Z, Z will block antecedent-government. If Y is an adjunct *wh*-trace, a DP-trace, or a head-trace, the result will be a violation of the ECP, since these elements all require antecedent-government. If Y is an argument *wh*-trace (as in (105a)), the structure will be well formed unless subjacency is also violated.

Relativized minimality is intended to replace the 'absolute minimality' condition of the *Barriers* system that we saw in (107). More precisely, the definition of antecedent government that is adopted is the following:

Antecedent Government (Third Revision):
X antecedent-governs Y iff:

(132a) X and Y are coindexed

(132b) X c-commands Y

(132c) no barrier intervenes

(132d) relativized minimality is respected

Principles (132a–c) are essentially as in *Barriers*, except that here 'barrier' means only 'barrier for movement' (defined in (97)). Principle (133d) replaces the 'absolute' minimality condition that I gave in (107).

According to relativized minimality, each type of chain is 'on its own track'. What is an intervener for one kind of chain has no effect on the others. So, DP-movement can cross intervening elements that block antecedent-government of *wh*-traces, such as negation:

(133) The students$_i$ weren't arrested t$_i$

And, of course, it can cross intervening heads – in fact, all cases of DP-movement must do this, since it is movement to a Specifier position. Conversely, head-movement can cross intervening Specifiers of both kinds – and, indeed, it must do so in order to reach higher head-positions.

Similarly, *wh*-movement is unaffected by the presence of an intervening subject position (remember that DP-movement is blocked by an intervening subject position):

(134) How$_i$ did he fix your car t$_i$?

Moreover, according to relativized minimality, *wh*-movement is not affected by intervening heads. This can also be seen in (134), where adjunct movement crosses the intervening heads I and C (at least). Here we see a major difference with the 'absolute' minimality condition of (107): remember that

(107) is intended to account for C-t effects by blocking antecedent-government across a filled C-position. However, (107) doesn't really work very well: we have to prevent both I' and V' from being minimality barriers. The fact that I' isn't a minimality barrier seems to be an instance of the general 'defectivity' of the I-system in *Barriers*, but of course this is in any case a problem. And we had no real suggestion to make about V'. Conceptually, relativized minimality seems preferable to absolute minimality in that it avoids these problems. We simply drop (107) and replace it with the definition of relativized minimality in (129). At the same time, we have a unified notion of 'barrier', relevant for both subjacency and the ECP: that in (97). And relativized minimality accounts for pseudo-opacity and negative islands – two phenomena that aren't handled in the *Barriers* system.

So relativized minimality seems like a conceptual and empirical improvement on the *Barriers* approach. Except for one thing: so far there is no account of C-t effects. Without this, there is an empirical case for retaining something like the absolute minimality condition, whatever its other flaws. The next section deals with Rizzi's account of C-t effects.

4.5.2 Head Government and θ-Government

4.5.2.1 A Conjunctive ECP

As we saw at the end of the last subsection, adopting relativized rather than absolute minimality means that we need an alternative account of C-t effects. To see what needs to be done, let's look again at a typical example:

(58') *Who$_i$ did you say [$_{CP}$ t'$_i$ [$_{C'}$ that t$_i$ wrote this song]] ?

According to relativized minimality, t'_i antecedent-governs t_i. No barrier intervenes (since we are now disregarding minimality barriers in the sense defined in (107)), and relativized minimality is respected, as no typical potential antecedent governor in an A'-specifier intervenes (see (129) and (130)). The subject trace is not θ-governed, as usual, but this is of no importance since it is antecedent-governed. So what's wrong with (58')?

Up to now, we've been working with a 'disjunctive' ECP: one that imposes an 'either–or' requirement of antecedent government or θ/lexical government. Rizzi proposes that what's needed is a 'conjunctive' ECP, as follows:

Traces must be both:

(135a) properly head-governed, and

(135b) either antecedent-governed or θ-governed

You can see that the disjunction of the previous formulation of the ECP is retained in (134b). Antecedent government is defined as in (132), and θ-government still distinguishes complements from non-complements. The new thing here is 'proper head-government'. We can define this notion in a way which parallels our definition of antecedent government, as follows:

Proper Head Government
X properly head-governs Y iff:

(136a) X is a lexical head, Agr or T

(136b) X c-commands Y

(136c) no barrier intervenes

(136d) relativized minimality is respected

In (58'), C doesn't properly head-govern t_i because C is not included in the class of proper head-governors. Since all traces must be properly head-governed according to Rizzi's conjunctive ECP, (58') is ruled out.

Rizzi gives a number of arguments for the conjunctive ECP. One is that objects, but not subjects, can undergo 'Heavy DP Shift', an operation that apparently places 'heavy' DPs to the right of a clause:

(137a) I would like to introduce t$_i$ to Mannie [$_i$ all the teenagers who can play the drums]

(137b) *t$_i$ are talented [$_i$ all the teenagers who can play the drums]

The contrast here has nothing to do with the presence of complementizers, and yet can be reduced to proper head-government. In (137a) the trace is properly head-governed by V, while in (137b) there is no proper head-governor for the subject trace present at all. It is important to see that here, as in (58'), I (AgrS) doesn't properly head-govern the trace because it doesn't c-command it. Also, if we take it that Heavy DP-Shift, rather like extraposition (see 4.2), adjoins the DP to the right of the clause, the moved DP antecedent-governs the subject trace as no barriers or A'-specifiers intervene (here it's important to bear in mind that 'intervening' is a hierarchical notion defined in terms of c-command, not a linear notion). Since antecedent-government is satisfied, the ECP must impose a further requirement – proper head-government does the job both here and in the case of Complementizer-trace effects.

The conjunctive ECP in (135), combined with the definition of proper head-government in (136), gives us the desired alternative account of C-t effects. However, we also have to account for the fact that dropping the complementizer makes examples like (58) good again. In other words, why is (62) good?

(62) [$_{CP1}$ Who$_i$ did [$_{IP1}$ you say [$_{CP2}$ t'$_i$ [$_{IP2}$ t$_i$ wrote this song]]]] ?

Rizzi proposes that the null complementizer is a kind of Agr in English. Because of this, the subject trace is properly head-governed, as Agr is defined as belonging to the class of proper head-governors in (136). The true representation for (62) thus looks more like this:

(62') [$_{CP1}$ Who$_i$ did [$_{IP1}$ you say [$_{CP2}$ t'$_i$ Agr [$_{IP2}$ t$_i$ wrote this song]]]] ?

Just looking at English, this might seem like a rather arbitrary thing to say. However, there is quite a bit of crosslinguistic evidence that agreeing complementizers exist and that they facilitate extraction of the subject.

4.5.2.2 Cross-Linguistic Evidence for Agr in C

One piece of evidence comes from French. French shows C-t effects that are very similar to those we've seen in English:

(138) *Qui$_i$ as -tu dit qu' t$_i$ a écrit ce livre ?
Who have you said that has written this book ?

However, the option of deleting the complementizer (or having a null complementizer) doesn't exist in French. Instead, sentences like (138) can be 'saved' by changing *que* (which in (138) undergoes a regular phonological reduction to *qu'*) to *qui*:

(139) Qui$_i$ as -tu dit qui t$_i$ a écrit ce livre ?
Who did you say QUI has written this book ?
'Who did you say wrote this book?'

The morphological change from *que* to *qui* can be thought of as a reflex of the presence of Agr in C. In other words, *qui* is *que* + Agr, and so able to act as a proper head-governor for the subject trace. (If you try to apply the *Barriers* account of C-t phenomena that I described 4.4.2. to the French examples in (138) and (139) you'll see that both examples come out as ungrammatical.)

West Flemish has a similar alternation between *da* and *die*:

(140a) Den vent$_i$ da Pol peinst [$_{CP}$ t'$_i$ DA [Marie t$_i$ getrokken heet]]
The man that Pol thinks that Marie photographed has

'The man that Pol thinks that Marie has photographed'

(140b) Den vent$_i$ da Pol peinst [$_{CP}$ t'$_i$ DIE [t$_i$ gekommen ist]]
The man that Pol thinks DIE come is
'The man that Pol thinks has come'

As with French *qui*, we can think of West Flemish *die* as *da* + Agr; hence it is able to properly head-govern the subject trace.

Some languages have very elaborate systems of agreement in Comp independently of subject extraction. One example, again given by Rizzi, is that of the Bantu language Kinande. Kinande is typical of Bantu languages in having a rich system of noun classes, which trigger various kinds of agreement processes. Among these processes is one whereby a fronted *wh*-element agrees in class with its complementizer (the Roman numerals indicate the noun classes):

(141a) IyondI yO kambale alangIra
who-I that-I Kambale saw

(141b) aBahI Bo kambale alangIra
who-II that-II Kambale saw

(141c) EkIhI kyO kambale alangIra
what-VII that-VII Kambale saw

(141d) EBIhI ByO kambale alangIra
what-VIII that-VIII Kambale saw
'What/who did Kambale see?'

So there is good reason to think both that C can bear agreement features and that the presence of these features can facilitate subject extraction by making C into a proper head-governor.

In our discussion of Complementizer-trace phenomena in 4.3.2.1 we also mentioned the fact that I-to-C movement is not allowed when the subject is extracted:

(66a) Which girl$_i$ did$_j$ he t$_j$ kiss t$_i$?

(66b) *Which girl$_i$ did t$_i$ t$_j$ kiss him?

(66c) Which girl$_i$ t$_i$ kissed him?

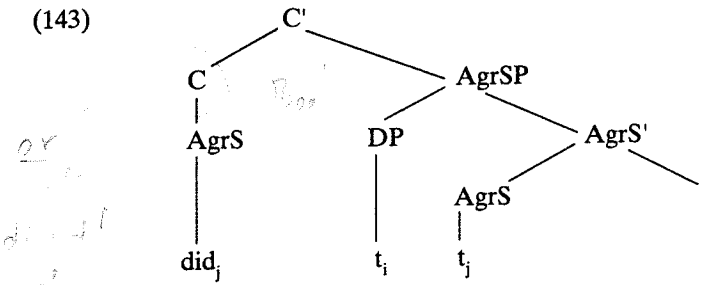
In (66b) we appear to have Agr in C, since the auxiliary is dominated by AgrS and is raised to C. Why, then, is this ruled out? This suggests a further restriction on the definition of proper head government. What we want to say, in order to capture the difference between (66b) and (68), is that it is not really C that c-commands the subject trace in (66b), but AgrS. Let's look again at our definition of c-command:

C-command:

(142) α c-commands β iff every category dominating α also dominates β

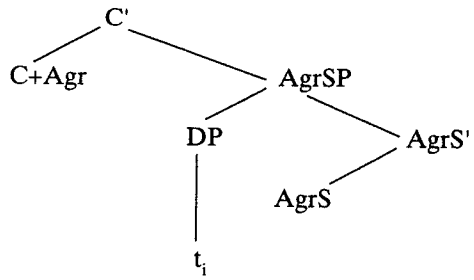
The relevant parts of (66b) look like this:

(143)



Did is an AgrS element, and the first node dominating AgrS is C, and C does not dominate the subject trace. On the other hand, in well-formed examples like (66c) and (68), we have a structure like (144):

(144)



Here C+Agr c-commands the subject trace, since the first dominating node it (C') dominates the subject trace.

In 1.4.2.4 I introduced the verb-second (V2) phenomenon of the Germanic languages. We saw that one XP moves to SpecCP and the verb (plus the I-material) moves to C. In subject-initial clauses, then, we actually have subject extraction:

(145) $[_{CP}$ Ich_i $[_{C'}$ las $[_{IP}$ t_i schon letztes Jahr diesen Roman]]]
I read already last year this book
'I read this book last year already'

(This analysis of subject-initial V2 clauses has been called into question by Travis (1984) and Zwart (1993).) Similarly, extraction of the subject from embedded V2 clauses is possible:

(146) Wer_i hat sie gesagt $[_{CP}$ t'_i ist $[_{IP}$ t_i gekommen]] ?
Who has she said is come?
'Who did she say came?'

This implies that C in a V2 clause automatically has an Agr associated with it, and so is able to act as a proper head-governor for the subject trace. It is possible that the presence of this Agr is also connected to verb-movement to C in these constructions (this has been suggested by a number of researchers).

So we see that Rizzi offers an account of Complementizer-trace phenomena which relies on the notion of proper head-government given in (136). This approach has various crosslinguistic implications, some of which we have looked at here.

4.5.2.3 Against θ-Government

Rizzi's approach to Complementizer-trace phenomena seems to do well empirically, as we have seen. It certainly seems to do better than the notion of absolute minimality given in (107). However, the ECP now looks quite complicated, since it contains a conjunction, one of whose clauses is itself a disjunction. As a last topic under the heading of relativized minimality, let's see how this can be simplified.

The first step is to see that the complement/non-complement asymmetries that θ-government is meant to account for aren't really complement/non-complement asymmetries, but instead involve the idea of 'referentiality'. We

can see this if we look at adverbial and idiomatic complements. These elements are clearly complements, and yet they behave like adjuncts for extraction (I've dropped the indices here, for reasons that will be explained below):

(147a) ??What project are you wondering [how [to make headway on t t]] ?

(147b) *What headway are you wondering [how [to make t on this project]] ?

A noun like *headway* in an idiom like *make headway on* has no reference: there's no particular thing called 'headway' that's being made; instead, the whole idiom means something like 'advance'. According to any of the various versions of the ECP that we've seen up to now, the trace of *headway* is θ-governed in (147b). But the extraction is very bad, rather like adjunct-extraction. So Rizzi suggests that what underlies complement/non-complement distinctions is the property of being referential, rather than the property of being a complement. Since, with the exception of pieces of idioms like *headway* and some adverbial or adjectival complements, complements are typically referential, most of the time the distinction can't be seen.

Instead of θ-government, then, Rizzi proposes that what underlies the complement/non-complement asymmetries the ECP deals with is the ability of referential categories to bear an index and thus be bound by their antecedent. Non-referential categories have no index and so can't be bound (see the definition of binding given in (7) of Chapter 3); this is why they're subject to a more rigorous locality condition. In Rizzi's view, aside from subjacency, there is no well-formedness condition on the relation between argument traces and their antecedents beyond what, right at the beginning of this chapter, we called the central property of Move-α: the formation of a binding relation between the moved category α and its trace. So the second clause of the ECP can be dropped for argument traces, leaving just the proper head-government requirement.

Subjects are usually referential (except for expletives, which can't be extracted). So the suggestion just made implies that subjects should pattern with objects as far as asymmetries are concerned. In fact, the only clear examples of subjects patterning with adjuncts can come under the rubric of complementizer-trace effects. This is true of (53), for example:

(53) *Which $band_i$ were you wondering whether t_i will play that song ?

And conversely, it's quite clear that subjects pattern with objects rather than with adjuncts in negative islands. Compare (148) with the discussion of (122) above:

(148a) Who_i don't you think we can help t_i ?

(148b) ?Who_i don't you think t_i can help us ?

When they are more deeply embedded in *wh*-islands, subjects pattern with objects and against adjuncts:

(149a) ??Who$_i$ do you wonder whether we believe t_i can help us ?

(149b) ?Who$_i$ do you wonder whether we believe we can help t_i ?

(149c) *How do you wonder whether we believe [we can help Bill t] ?

(Remember that we're only interested in the lowest construal of *how* in (149c); I'm not coindexing the trace as adjuncts lack referential indices according to the theory being put forward here.) The data in (148) and (149) confirm that subjects and objects – as typically referential categories – pattern together, distinct from adjuncts.

Getting rid of θ-government is a good move as regards our treatment of DP-traces. Remember that in 4.4.3.1 we saw that DP-traces appear to require antecedent-government as well as θ-government. This question of why θ-government isn't enough partly disappears if we get rid of this notion. However, the question now becomes: why do DP-traces act like adjuncts as regards movement? DP-traces presumably bear indices, and so why isn't forming a binding relation with the antecedent enough? Rizzi's answer to this is that DP-movement always forms a θ-chain: the head of the chain must be in a Case-position and the foot of it in a θ-position. Rizzi proposes that *the locality condition on chain links is antecedent-government*, defined in terms of relativized minimality as in (132). DP-traces must be in a well-formed θ-chain, and so must be antecedent-governed by their antecedent.

All of this just leaves adjuncts, or more precisely non-referential categories. Why are such categories required to be antecedent-governed? We've just seen that antecedent-government can be viewed as a condition on chain links. Moreover, non-referential categories lack indices. So the only way they can be connected to their antecedents is by antecedent-government. In the last analysis, then, it is the non-referential nature of certain traces – typically, but not always, adjuncts – which makes them subject to a stricter locality requirement than referential traces. Chain-formation by antecedent government is a kind of last-resort strategy for connecting these traces to their antecedents. If this idea is to work, we cannot define antecedent-government itself in terms of coindexation, as we did above. So let's make a minor reformulation, as follows:

Antecedent Government (Fourth Revision):
X antecedent-governs Y iff:

(132a') X and Y are non-distinct

(132b') X c-commands Y

(132c') no barrier intervenes

(132d') relativized minimality is respected

Statement (132') differs from (132) only in that 'coindexed' has been replaced by 'non-distinct'. This allows us to maintain that non-referential traces lack indices.

Rizzi's conclusion is that (132') is *not* part of the ECP, but part of the definition of chain-formation. It is relevant for non-referential traces by default, since, being non-referential, such traces can't be bound. This is presumably why it's also relevant for heads. And it's relevant for DP-traces since such traces must be in θ-chains. It is not relevant for referential traces, since they can be bound by their antecedents, and are subject only to the proper head-government requirement of the ECP (like all traces).

4.5.3 Conclusion

Relativized minimality really leads to two main conclusions. First, antecedent-government, defined as (132'), is relevant for chain-formation. Second, the ECP consists only of a rather simple proper head-government requirement as in (136). You might have noticed that proper head-government is a slightly stipulative requirement in that it simply lists those heads that head-govern. No reason emerges for why some functional heads and not others are head-governors. Also, although Rizzi presents a very interesting theory of the ECP, the *Barriers* account of subjacency remains, along with the notion of barrier as defined in (97). In a sense, this is another result: what emerges with particular clarity from Rizzi's work is that head-movement, DP-movement, and movement of non-referential/adjunct *wh*-elements have important properties in common: all three movements are highly local, and, allowing for successive cyclicity, we can say that they always move to the nearest appropriate c-commanding position. In fact, this was part of our informal characterization of head-movement and DP-movement at the beginning of this chapter. Movement of referential/argumental *wh*-elements is much freer, on the other hand, and appears to be subject only to subjacency and the proper head-government requirement. So there are two broad types of movement, with *wh*-movement taken as a whole straddling the distinction.

We're now basically in the position of having sorted out the role of antecedent-government: as defined in (132'), antecedent-government is the condition on chain-formation. We have a residual ECP (proper head-government as in (136)) and we have subjacency, defined as in *Barriers* (see 98)). The question that remains is: can subjacency and the ECP be unified under a single locality condition? This is the question that Manzini (1992) addresses and to which we should now turn.

4.6 Locality

Manzini's system is, at the time of writing, the most comprehensive and up-to-date set of proposals on the nature of Move-α and the theory of locality. In this section, we'll come back to the questions set out at the beginning of this chapter: what we're really after is a unified, conceptually simple theory

of movement relations. It should be clear from the previous section we're almost there, but not quite.

4.6.1 Some Definitions

A fundamental component of Manzini's theory is her use of two types of index. A distinction is made between 'categorial indices' and 'addresses'. Categorial indices are basically the indices we've been using all along to indicate anaphoric relations, including, of course, relations holding between traces and their antecedents. Addresses, on the other hand (the term and to some extent the idea are taken from Vergnaud (1985)) differ from categorial indices in being relational. That is, a category is addressed by being in a certain kind of relation with another category. The usual (but not the only) way that categories get addressed is by Case-assignment: a Case-assigner gives an address to the category it Case-marks (of course, it is possible to think of this in terms of Case-checking, as discussed in 2.6).

All movement dependencies involve sharing categorial indices between the moved category and its trace – again, there is nothing new here; this is precisely what we've been assuming all through this chapter (except for 4.5.2.3). The differences arise in connection with addressing. Let's look at each kind of movement in turn, distinguishing adjunct *wh*-movement from argument *wh*-movement (I'll revert to referring to the two kinds of *wh*-movement in this way, even though Rizzi argues that this is not really correct, as we have seen).

First, head-movement, as in (150) (I'm ignoring the split-Infl structure again):

(150) Jean $[_I$ embrasse$_i$] souvent $[_{VP}$ t_i Marie]
Jean kisses often Marie
'Jean often kisses Marie'

Here neither the trace nor the antecedent is in an addressed position, and so the dependency must be categorial rather than address-based. In general, since heads aren't Case-marked, head-movement cannot be an address-based dependency.

Similarly, adjunct *wh*-movement moves from a Caseless position (an adjunct) to a Caseless position (SpecCP):

(151) How$_i$ did you fix the car t_i ?

Third, we've defined DP-movement as being movement in search of Case (see 2.3). Therefore, this cannot be an address-based dependency since only the antecedent has an address.

So we see that the inherently more local kinds of movement – head-movement, DP-movement, and adjunct *wh*-movement – do not form address-based dependencies. With these types of movement the dependency is based purely on a shared categorial index.

On the other hand, argument *wh*-traces occupy Case-marked positions.

We saw in 2.4 that this is in fact required, as the ungrammaticality of examples like (152) shows (see 2.4 (70a and b)):

(152a) *Who$_i$ does it seem [t_i to speak Chinese] ?

(152b) *Who$_i$ was it believed [t_i to speak Chinese] ?

In general, properties of traces are transmitted to their antecedents. So, since the trace has an address, we can regard the argument *wh*-movement dependency as an address-based dependency. The difference in the ability to form address-based dependencies corresponds to the notion of referentiality in Rizzi's work (see 4.5.2.3): this is how we distinguish the more local dependencies (adjunct *wh*-movement, DP-movement, and head-movement) from the less local argument *wh*-movement.

One more notion is needed before we can give the general locality constraint. This is the notion of a 'sequence', a neutral term intended to cover both categorial and address-based dependencies:

Sequence:

(153) $(\alpha_1 \ldots \alpha_n)$ is a sequence iff every element in the sequence α_i is co-indexed with and c-commands the next

In other words, both types of dependencies require that each link between the antecedent and the trace c-commands the next. Now for the locality principle:

Locality:

(154) α is a dependent element iff there is an antecedent β for α and a sequence $(\beta \ldots \alpha)$ where no link of the sequence crosses a barrier

Definition (154) is very close to Chomsky's definition of antecedent-government in *Barriers* (see (104)), and to part of Rizzi's definition: see (132').

4.6.2 Weak Islands

Given the definitions in (153) and (154), weak islands (islands from which adjuncts can't be extracted but arguments can) can accounted for. Let's consider (for the last time) argument-adjunct asymmetries in infinitival *wh*-islands.

(155a) ?$[_{CP1}$ Whose car$_{i/K}$ were $[_{IP1}$ you wondering $[_{CP2}$ how$_j$ $[_{IP2}$ to fix $t_{i/K}$ t_j] ?

(155b) *How$_j$ were you wondering [whose car$_{i/K}$ to fix $t_{i/K}$ t_j] ?

Here the argument trace has both the index *i* and the address *K*, while the adjunct trace just has the categorial index *j*. For the purposes of this illustration, we'll just assume that CP is a barrier – we'll come to back to how barriers are determined in 4.6.4. In (155b), then, there is no sequence with index *j* where each link governs the next: CP intervenes and breaks the sequence, and

so locality isn't satisfied. Again, this account should seem quite familiar since it is broadly similar to both Chomsky's and Rizzi's. However, the treatment of (155a) is quite different. Here, the categorial dependency is blocked in the same way as for the adjunct. However, the argument has the capacity to form an address-based dependency, which involves assigning the address *K* to all the heads intervening between $wh_{i,K}$ and the trace $t_{i,K}$. So the sequence (wh_i, C, I, V, C, I, V, t_i), all bearing the address *K*, is formed and the locality condition is satisfied. In this way, argument-adjunct asymmetries can be handled in terms of the single notion of locality given in (154).

There are other syntactic contexts which act like weak islands in the sense described above. One such context that we haven't mentioned up to now is the complement to factive verbs. These were discussed by Cinque (1991). 'Factive' predicates are predicates which presuppose the truth of their complement. Even if you negate the main clause, the complement clause is still taken to be true (this is one of the classic diagnostics for presupposition). Compare *regret* with *believe* in the following examples:

(156a) I don't believe that Sheena is a punk rocker

(156b) I don't regret that Sheena is a punk rocker

In (156a), there is no commitment to the truth or falsity of the proposition expressed by the complement clause: Sheena may or may not in fact be a punk rocker. In (156b), on the other hand, the fact that Sheena is a punk rocker is presupposed, taken for granted as true. Verbs that are like *regret* in this respect are known as factive verbs, while *believe* and similar verbs are non-factive.

The relevant phenomenon in the present connection is that the complements of factive verbs are weak islands, as the following sentences show:

(157a) What$_i$ do you regret [$_{CP}$ THAT YOU FIXED t_i] ?
(157b) *How$_i$ do you regret [$_{CP}$ THAT I FIXED YOUR CAR t_i] ?

As usual, *how* in (157b) should be interpreted as related to an adjunct of the *embedded* clause, not the main clause; the position of the trace inside the brackets marking the lower clause indicates this in (157b). This interpretation is completely unavailable. Manzini proposes that there is an empty *wh*-like operator in SpecCP of factive clauses, and that this is the element that makes them factive. This operator makes factive clauses structurally equivalent to *wh*-islands, and so we find the selective blocking of extraction that we find in *wh*-islands.

This kind of approach extends to both pseudo-opacity and negative islands:

(119b) *Combien$_i$ a-t-il beaucoup consulté t_i de livres?
How-many has-he much consulted of books?

(120b) *?How$_i$ didn't he fix your car t_i ?

If we assume that only one Specifier position is available for a given category, and that *beaucoup* occupies a (non-L-related) Specifier structurally close to VP – we'll discuss the precise nature of this position in 4.6.4 – while *not* occupies SpecNegP (although in (120b) it has attached to the auxiliary and raised with it while presumably, its trace still occupies SpecNegP), then each of these examples creates an analogous configuration to *wh*-islands, and the account given for (155) carries over. Here, too, it should be clear that Manzini's theory is very similar to Rizzi's.

4.6.3 Strong Islands

An important advantage of the idea of address-based dependencies and their link to Case-assignment is that it can capture the fact that certain types of NPs (or DPs) block extraction, while VPs never do. We saw in 4.4.1.3 that the *Barriers* approach can't handle the complement case of the CNPC:

(33b) *Which band$_i$ did you believe [$_{DP}$ the claim that [$_{IP}$ we had seen t_i]] ? (see (9))

Complex DPs are strong islands, in the sense that they block extraction of both argument and adjunct *wh*-elements. Extraction of adjuncts is blocked by the fact that DP is a barrier (again, I'm just stating this for the moment – we'll come back to it in 4.6.4) and, unlike CP, its Specifier isn't a position that can be moved through successive-cyclically. Extraction of arguments is blocked by the fact that N has an address of its own, and so formation of an address-based sequence that passes through N is impossible. This is why complex NPs (or DPs) are islands.

Another interesting aspect of the address-based dependencies is that they can handle the fact that *wh*-islands are sensitive to tense. As I commented in 4.1, non-finite *wh*-islands more or less allow argument extraction, while tensed *wh*-islands do not. In other words, non-finite *wh*-islands are weak islands, while finite *wh*-islands are strong. Here, again, are the relevant examples:

(36a) ?[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ to fix t_i t_j] ?

(36b) ?*[$_{CP1}$ Whose car$_i$ were [$_{IP1}$ you wondering [$_{CP2}$ how$_j$ [$_{IP2}$ you should fix t_i t_j] ?

Manzini suggests that this contrast can be captured by saying that finite Tense (remember that 'IP' is a cover term for TP and the other functional categories associated with V) has its own address. In that case, the address-based dependency that is required to link an argument *wh*-trace to its antecedent across a *wh*-island cannot be formed.

Similar reasoning holds for another kind of island that we briefly alluded to in 4.1. Definite DPs seem to form strong islands:

(158a) ??Which band$_i$ did [$_{IP}$ you write [$_{DP}$ that song about t_i]] ?

(158b) *Which band$_i$ did [$_{IP}$ you sing [$_{DP}$ Mick's song about t_i]] ?

(158c) Which band$_i$ did [$_{IP}$ you write [$_{DP}$ a song about t_i]] ?

Manzini rules out (158a) and (158b) by saying that a definite D has its own address while an indefinite one does not. Because of this, address-based dependencies cannot be formed in (158a) and (158b), while they can in (158c). (Again, we have to assume that DP is a barrier here.) So the system makes the right distinctions, to a fair degree of approximation, among these examples.

As I said above, Manzini's approach captures argument-adjunct asymmetries in terms of the idea that only argument traces can form address-based dependencies. This effectively factors out the role played by θ-government in the *Barriers* version of the ECP. The definition of locality in (154) looks very like antecedent-government – in fact, for non-addressed traces (traces of adjuncts, DP-movement, and head-movement) (154) functions almost exactly like antecedent government in both *Barriers* and relativized minimality. The major difference between Manzini's system and the *Barriers* system is the proposal that (154) also accounts for subjacency. Now let's look at how this is done.

4.6.4 Subjacency Again

The locality principle requires that the sequence linking the antecedent to the trace be such that each link governs the next. For argument traces, as we have seen, this requirement relates to an address-based dependency. So when subjacency is violated, one link of an address-based dependency must be separated from the next by a barrier (as you can see, Manzini assumes a one-barrier definition of subjacency rather than the usual two-barrier one). In order to see in detail how this works, we need to see how barriers are defined. Barriers are defined in terms of 'g-marking', as follows:

Barrier:
β is a barrier for α iff

(159a) β is a maximal projection

(159b) β dominates α

(159c) if α is g-marked, β dominates the g-marker of α

So a barrier for α is a maximal projection dominating both α and its g-marker. G-marking is a concept similar to L-marking, although it makes reference to sisterhood rather than θ-government (and is not restricted to lexical heads, as L-marking is in (95')):

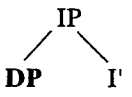
G-Marking:
β g-marks α iff β is a head and:

(160a) β is a sister to α, or

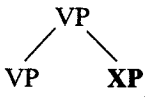
(160b) β is a sister to a category that agrees with α

Subjects, adjuncts, and the CP that modifies the head of a relative clause are not sisters to a head, and therefore not g-marked. The configurations are as follows:

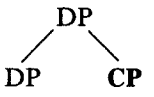
(161a) Subject:



(161b) Adjunct:



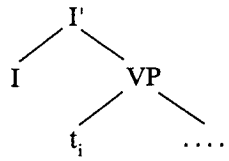
(161c) Relative:



(Again, if we follow Kayne's proposals about how phrase structure should be organized (see 1.3.4), we'll have to come up with different proposals for the positions of adjuncts and relatives; I'll leave this matter to one side here.) Since these domains are not g-marked, only clauses (159a and b) of the definition of barrier apply, and so the first maximal projection dominating the domains is a barrier. This means that no sequence of any kind can be formed between an element inside one of these domains and an element outside it. So extraction of anything, including an argument *wh*-element, from inside one of these domains will violate the locality principle.

The other principal strong islands that are discussed are the complement CNPC, tensed *wh*-islands, and definite DP-islands. We discussed these cases in the previous subsection, assuming without explanation that DP and CP are barriers. Now let's look at why these categories are barriers. In fact, the concept of barrier defined in (159) is extremely general: any maximal projection dominating a trace and its g-marker can be a barrier. This means that the assumption made above, that CP and DP are barriers, is unproblematic. The proposal is very strong, though. It means, for example, that where a direct object is extracted, VP becomes a barrier, as it dominates both the trace and the g-marker of the trace, V. However, where an argument is extracted, there are two options: either an address-based dependency can be formed, as we have seen, or extraction can pass through the A'-specifier position. Manzini assumes that all non-argument categories have an A'-specifier (for lexical categories, this would entail a complication of the notion of L-relatedness in terms of which we defined A- and A'-positions in 3.3.2.3; however, I'll leave this question to one side). This position is able to form an 'escape hatch' for extraction thanks to the second clause of the definition of g-marking given in (160). Consider the configuration where a category is successive-cyclically moved through the A'-Specifier of VP:

(162)



Assuming that a specifier always agrees with its head (see the discussion of this relation in 2.6), we see that I g-marks the trace here, since I is the sister of VP, a category that agrees with the trace. Since VP does not dominate I, it cannot be a barrier for the trace, because it doesn't dominate both the trace and its g-marker. However, VP is a barrier for any extraction which does not pass through its A'-specifier. This is why VP-adverbs can form weak islands, as in the case of pseudo-opacity (see (119b) above). As we saw above, argument *wh*-elements can form an address-based dependency in this case. More generally, if the Specifier position is filled in (162) (or a similar configuration in another non-argumental XP), then the only option for extraction is the formation of an address-based dependency. This is how the weak-island configurations that we looked at in 4.6.2 are created: factive islands, negative islands, and *wh*-islands.

CP, like other maximal projections, is a barrier for material it contains. Hence the only way out of CP is by formation of an address-based dependency (this option being available for argument *wh*-traces only) or by passing through the A'-specifier, SpecCP. In *wh*-islands, the A'-specifier is filled, and so all *wh*-islands block adjunct extraction. Argument extraction via an address-based dependency is possible unless some independent factor such as Tense (see 4.6.3) intervenes.

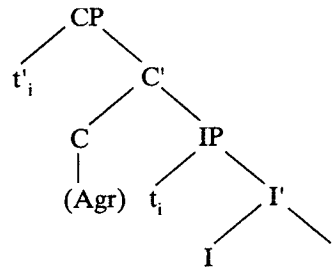
DPs are also barriers for the material they contain. Moreover, DPs are typically arguments. Developing Chomsky's proposal in *Barriers* that adjunction to arguments is impossible (see 100a), Manzini proposes that argument XPs cannot have A'-specifier positions. So the configuration in (164) never obtains with DP. For this reason DPs are always weak islands, and can become strong islands either through an intrinsic property of D (definiteness, as we saw above) or by containing a Case-marked N, as in the case of the complement CNPC. In both cases, the address-based dependency is blocked.

So we see that subjacency too can be derived from the locality principle in (154), combined with the definitions of barrier and g-marking in (159) and (160). As we saw at the end of the previous section, the other principal element of the theory of locality is proper head-government. Here the main facts to account for are C-t effects. So now let's see how Manzini's theory can deal with these, and, in particular, whether they can be unified with (154).

4.6.5 C-t Effects Again

The configuration for the Complementizer-trace effects is (163), as we have seen:

(163)



As we saw in 4.5.2.1, Rizzi argues that C must contain an abstract Agr in order for t_i to be properly head-governed. Manzini follows this basic idea, but reduces the requirement to the general locality principle of (154). In other words, the proper head-government requirement is derived, not primitive.

The basic idea is that the presence of an agreement relation is necessary in order for the subject trace to have an address. The agreement relation is optional in itself. If it does not hold, then the subject must form a categorial dependency. Since CP is a barrier for the subject trace (you can see this if you try the definitions of barrier and g-marking in (159) and (160)), SpecCP must function as an escape hatch. So, if SpecCP is filled by another *wh*-element, the sentence is bad. This is what we see in examples like (53):

(53) *Which band$_i$ were you wondering whether t_i will play that song ?

This example violates locality because the trace and the antecedent are separated by the barrier CP.

So, where the agreement-in-C option is not taken, the subject must move through SpecCP. Now, SpecCP agrees with C by Specifier-head agreement, and SpecIP agrees with I by Specifier-head agreement. By transitivity of agreement, C and I also agree. So, we have agreement in C. This is manifested in English by the null alternant of *that* and in French by the presence of *qui*, as we saw in 4.5.2.1.

If, on the other hand, the agreement-in-C option is taken, then the subject trace can form an address-based dependency. Here we automatically get the null form of *that* and *qui* instead of *que*. When the subject forms an address-based dependency, it can escape from *wh*-islands and negative islands, as in (148b) and (149a):

(148b) ?Who$_i$ don't you think t_i can help us ?

(149a) ??Who$_i$ do you wonder whether we believe t_i can help us ?

As long as we assume that agreement in C is related to the formation of address-based dependencies, we can reduce the proper head-government requirement to the general locality principle.

4.6.6 The Nature of Move-α

So we see that Manzini is able to unify subjacency, antecedent-government,

and proper head-government under a single principle. Remember that this principle also applies to DP-movement and head-movement. So the theory of Move-α reduces to (5b) – the statement that movement creates binding relations – and the locality principle in (154). In fact, since binding involves the formation of one kind of sequence (one based on categorial indices) we can collapse (5b) and (154) as follows:

Move-α (Final Version):

(164) a dependency such that there is an antecedent β for α and a sequence (β .... α) where no link of the sequence crosses a barrier

You might notice that this definition doesn't really say much about movement or about the creation of syntactic derivations. Following Rizzi's proposal for antecedent-government, it is natural to think of (164) as a condition on the formation of chains:

Chain:

(164') a dependency such that there is an antecedent β for α and a sequence (β .... α) where no link of the sequence crosses a barrier

We've seen that there are various kinds of chains: θ-chains formed by DP-movement and A-chains formed by anaphors of certain kinds (see 3.5.3); (164') can be taken as defining the class of chains which is then subdivided in various ways.

So it may turn out that we don't need Move-α at all, but simply the notion of chain given in (164'). This is an important issue, because if we don't need Move-α but can make do just with chains, then we don't really have derivations, but instead we can posit a single level of syntactic representation which contains all relevant semantic, phonological, and lexical information. The issue of whether syntax is really derivational or is instead 'representational' has been debated for over 10 years now, and is still not resolved. After all we have seen in this chapter, it should be reasonably clear that it is at best extremely hard to tell whether Move-α has an existence independently of chains.

4.7 Conclusion

This chapter has dealt with one of the most important areas of research in generative grammar. It is also worth pointing out that the phenomena discussed here, unlike those discussed in the other chapters, were completely unknown 40 years ago and have received no interesting treatment in non-generative theories. I've presented a general theory of locality, following Manzini's work, but after all the technicalities it's perhaps useful to take stock of what is in fact known about locality. I'll try to list these points below:

- despite appearances, *wh*-movement is not unbounded but local, cyclic and sensitive to islands (4.1, 4.2.1, 4.2.2)

- adjunct *wh*-movement and argument *wh*-movement show differing behaviour, with adjunct *wh*-movement being much more constrained than argument *wh*-movement (weak *vs* strong islands) (4.3.1)
- adjunct *wh*-movement patterns with DP-movement and head-movement in being blocked by a potential landing-site (relativized minimality; 4.5.1)
- argument *wh*-movement obeys the CED, the CNPC, and is sensitive to Tense and definiteness inside otherwise weak islands (4.6.3, 4.6.4)
- all movement is subject to a proper head-government requirement (4.3.2; 4.5.2; 4.6.5).

Manzini's approach captures all these observations under a single locality principle, as we saw in 4.6. This locality principle is formulated in terms of barriers, which are, in turn, defined in terms of g-marking, as we saw. The notions of barrier and g-marking are configurational notions, in that they define certain types of structures in a purely geometric way.

In recent work that sketches out the minimalist programme (see 2.6), Chomsky has suggested a seemingly rather different approach. According to this view, the essence of locality should be that any operation of movement is as 'short' as possible, in that movement should always target the closest available landing-site. In fact, relativized minimality effectively states this, since it says that a category β of the same type (A, A', head) as α – a potential landing site for α – prevents α being moved past β. So one can readily think of a 'shortest move' constraint as applying to adjunct *wh*-movement, DP-movement, and head-movement. It is not clear how this idea should extend to argument *wh*-movement, though. Here we presumably need to refer to actual rather than potential landing-sites, with the class of actual landing sites defined by subjacency in some way. At the time of writing, no proposal has been made for dealing with the locality constraints on argument *wh*-movement that fits with the minimalist conception of shortest move.

This chapter has not had much to say about comparative syntax, on the face of things at least. I've mentioned Rizzi's work on *wh*-islands in Italian, Huang's work on adjunct *wh*-elements in Chinese, and Kayne's work on Preposition-stranding. However, I hope you can see that the kinds of phenomena that are being dealt with here are fundamental to linguistic theory. The tacit assumption throughout this chapter has been that the principles that are being put forward are UG principles, subject to a small degree of parametrization. Now that you have a reasonably good idea what a number of those principles and some of the associated parameters are, it's time to take a wider view and look at the theory of parameters.

Parameters Discussed in this Chapter

As just mentioned, parametric variation has not been a central focus of this chapter. None the less, a number of parameters have been discussed to varying levels of detail. I'll now summarize and elaborate on what we've seen.

1. Some languages, like Russian and Latin, appear systematically able to violate the Left Branch Condition, while languages like English cannot (see 4.1). We suggested that this might be connected to the absence of overt D-elements in these languages. French can apparently violate the LBC in the sole instance of *combien*, a quite mysterious fact.
2. In 4.3.1.2 we saw that there is variation as to the level at which *wh*-movement can take place. In Chinese, *wh*-movement is always covert (but see Watanabe (1992), Aoun and Li (1994)). Many languages pattern like Chinese: Japanese, Korean, Armenian, and others. In French, main-clause *wh*-movement is optionally overt or covert, depending on register, while *wh*-movement in embedded [*+wh*] clauses is always obligatory:

(165a) Quelle fille a-t-il embrassée? (Standard/literary French)
Which girl has he kissed?

(165b) Il a embrassé quelle fille? (Colloquial French)
He has kissed which girl?
Which girl did he kiss?'

(165c) Je me demande quelle fille il a embrassée (Both registers)
'I wonder which girl he has kissed'

(165d) *Je me demande il a embrassé quelle fille
I wonder he kissed which girl

In English, a single *wh*-element moves to the Specifier of a [*+wh*] CP; where there is more than one *wh*-element in a clause, the others do not move overtly:

(166) Who$_i$ t$_i$ said what ?

As mentioned in the Appendix to Chapter 3, however, the interpretation of the *wh*-element *in situ* (*what* in (168)) as a quantifier binding a variable implies that we should assume that it moves covertly to Spec,CP. If so, the LF of (166) might look like (166'):

(166') Who$_i$ what$_j$ t$_i$ said t$_j$?

Some languages, including most of the Slavic languages, overtly form multiple questions that look like (166'). This is shown in the Russian example in (167):

(167) Kto čto kogda skazal?
who what when said?
'Who said what when?'

There are interesting differences among these languages as regards the multiple *wh* structures: see Rudin (1988) for discussion.

3. In 4.2.3 we saw Rizzi's evidence that Italian has slightly different BCs from English. Sportiche (1981) showed that French patterns like Italian. You might have noticed that we didn't see a way to integrate this variation into Manzini's version of subjacency in 4.6.4.
4. In (75), we saw the full effects of the null subject parameter, as described by Rizzi (1982). We recapitulate these here:

(75a) Possibility of phonologically empty referential subject pronouns

(75b) Impossibility of overt expletive pronouns

(75c) Possibility of free inversion

(75d) Apparent absence of Complementizer-trace effects

In 3.3.3, we suggested that the trigger for the null-subject parameter was the presence of verbal inflection that permits the identification of the referential properties of subject *pro*; we also discussed some problems for this idea. A number of researchers, notably Safir (1985), have denied that the properties listed in (75) correlate. Certainly, a superficial survey of languages will show that they do not. However, (75) is not a statement of an implicational relation of the Greenbergian kind that we discussed in Chapter 1 (see 1.3.3), but a statement of properties that follow from the availability of referential *pro* in a system of a given kind. If we look across all the languages of the world, we are likely to bring a number of extra unknown variables into play that may well disturb these correlations; but this would not affect their theoretical validity.
5. English allows Preposition-stranding; French doesn't. Most of the Scandinavian languages pattern like English (although Icelandic is a special case: see Kayne (1984)). In Dutch and German, the complement of a Preposition can only be extracted if the *wh*-element takes on a particular form into which the Preposition appears to be incorporated (see van Riemsdijk (1978)); archaic English shows the same phenomenon, as in *the person **whereof** I spoke*. In Welsh, the Preposition must agree with the *wh*-trace, giving something like *the person I spoke of+3sg* (see Hendrick (1988) – although Hendrick does not treat these cases as extraction). We saw in 4.3.3.2 that the difference between English and French can be related to the differing status of Prepositions as Case-assigners in these languages: see 2.5.3.

Further Reading

In this chapter more than the others, I've mentioned the central readings in the text; so here there'll be rather more repetition than in the other Further Reading sections.

wh-islands are discussed briefly and rather inconclusively in Chomsky (1964). Otherwise, the reference for island phenomena is Ross (1967), published in almost unaltered form as Ross (1986). This is a classic text of

generative grammar, and contains a wealth of insight and information. It remains a very influential piece of work. A good up-to-date discussion of island phenomena, including the division into weak and strong islands, is Cinque (1991). The first chapter of Manzini (1992) is an excellent overview, and is perhaps the best thing to read after this chapter.

Subjacency was first formulated in Chomsky (1973), an article which, as I said in the Further Reading to Chapter 3, is difficult but worth looking at. Successive cyclicity and the Strict Cycle Condition are also introduced here. The other major works on subjacency are Chomsky (1977), Rizzi (1982, ch. 2) (on the parametric variation between English and Italian; in this article, originally circulated in 1977, the idea of parametric variation appears for the first time), Huang (1982) (showing that it doesn't apply to covert *wh*-movement; argument–adjunct asymmetries and the Condition on Extraction Domains are first discussed here), Lasnik and Saito (1984, 1992), Chomsky (1986b) (*Barriers*), Cinque (1991), and Manzini (1992).

The Empty Category Principle was first proposed in Chomsky (1981). Several of the chapters of Kayne (1984) develop the idea in various ways; here the account of Preposition-stranding and the theory of connectedness are proposed. An approach comparable to connectedness is put forward by Pesetsky (1982). In addition to the references given under subjacency in the previous paragraph, all of which also deal with the ECP, important work on the ECP is found in Aoun (1985, 1986), Aoun, Hornstein, Lightfoot and Weinberg (1987) – Aoun's theory of 'generalized binding' largely anticipates relativized minimality.

Parasitic gaps were first discussed by Taraldsen (1979). Chomsky (1982, 1986b) discusses them in detail, and proposes quite different theories. Other important works on this phenomenon, in addition to Kayne (1984), are Engdahl (1983, 1985), Bennis and Hoekstra (1984), Longobardi (1985), and Frampton (1989).

The remaining topics of the chapter – barriers, relativized minimality, and locality – are discussed in the eponymous monographs by Chomsky (1986b), Rizzi (1990), and Manzini (1992) respectively. After reading this chapter, you should be able to tackle that material: indeed, I hope much of it will seem familiar. The important recent works on the theory of movement that I haven't gone into detail about here are Cinque (1991) and Lasnik and Saito (1992).

The question of whether Move-α has an existence independently of chain-formation was first raised in Chomsky (1981, ch. 6). The idea that Move-α should be reduced to chain-formation is argued for notably by Sportiche (1983), Brody (1985, 1995), and Rizzi (1986c). Moreover, both Rizzi (1990) and Manzini (1992) present theories that are broadly compatible with this view, as we have seen. The opposite view, namely that Move-α exists independently of properties of chains, is espoused by Lasnik and Saito (1984, 1992) and Chomsky (1986b). The minimalist organization of the grammar (see 2.6.2) appears to be 'derivationalist', but see Brody (1995) for a 'representationalist' version of minimalism.

Exercises

Exercise 1

We haven't said much in this chapter about extraction of APs, although we mentioned in 2.4 that *wh*-APs exist and can be moved. However, there are some surprising constraints on AP-extraction. Here is a range of different kinds of APs:

1. John considers Bill [$_{AP}$ intelligent] (epistemic)
2. John made Mary [$_{AP}$ angry] (causative)
3. We hammered the metal [$_{AP}$ flat] (resultative)
4. Keith took the stage [$_{AP}$ drunk] (circumstantial)
5. Kazuo ate the fish [$_{AP}$ raw] (depictive)

Try extracting these APs (with *how*) in this environment and in weak-island contexts. Try also to see whether the different kinds of APs are selected (by seeing whether the verbs are grammatical, or have the same meaning, without them). What generalizations can you arrive at here?

Exercise 2

Chomsky (1977) showed that a range of constructions, including *easy-to-please* constructions as in (1) and comparatives as in (2), involve movement of a null *wh*-phrase (a null operator: see 3.2.2.2):

1. John is easy to please
2. Ruth is stranger than Richard is

Chomsky used three main diagnostics for *wh*-movement:

3. It leaves a gap
4. It is apparently unbounded
5. It obeys subjacency

Show that these diagnostics apply to the operations in (1) and (2). This will entail showing the positions both of the gap and of the empty operator. (Hint regarding (2): many dialects of English express (2) as *Ruth is stranger than what Richard is.)*

Why can't we analyse *easy-to-please* constructions as involving the following movement dependency (possibly mediated by cyclicity)?

6. John$_i$ is easy to please t$_i$

Finally, following on from our brief discussion of control in 3.2.2.2, what does the interpretation of the PRO subject of *to please* tell us here?

Exercise 3

Consider the following parasitic-gap data:

This is the book that I filed t . . .

1. . . . without reading *e*

2. *... before hearing the rumour that the author had plagiarized *e*
3. *... after wondering whether I should throw away *e*
4. *... after the author of *e* had come to dinner
5. *... after going on holiday without reading *e*

What does this tell you about parasitic gaps? How best might they be analysed, especially in the light of Exercise 2?

Exercise 4
Look again at the contrast between (27) and (28), which illustrates the Right Roof Constraint:

(27) *The proof that the claim t_i was made by the Greeks was given in 1492 [$_{CPi}$ that the world was round]

(28) The proof that the claim t_i was made [$_{CPi}$ that the world was round] by the Greeks was given in 1492

There's a simple explanation for the rightward-boundedness of movement in terms of subjacency and successive cyclicity. Or is there? Think about it.

Exercise 5
In Modern Greek, factive islands are strong islands. That is, they block extraction of both arguments and adjuncts (this was discovered by Roussou (1993)):

1. *Pjon$_i$ thimase pu sinandises t_i ?
   Who you-remember that you-met?
   'Who do you remember that you met?'
2. *Pote$_i$ thimase pu sinandises ti Maria t_i ?
   When you-remember that you-met the Maria?
   'When do you remember that you met Maria?'

How might we integrate this fact in a theory like Manzini's? These factive complements have a special complementizer (*pu*); might this fact be significant?

Exercise 6
We saw in 4.4.3.2 that noun-incorporation cannot move the head of a subject DP into the verb. In that case, what do you make of the following Onandaga example?

1. Ka- hi- hw- i ne? o- HSAHE? T-a?
   3n- spill- CAUSE- ASP the PRE- bean-SUF
   'The beans spilled'

2. Ka- HSAHE? T- ahi- hw- i
   3n- bean- spill- CAUSE- ASP
   'The beans spilled'

The solution to this problem should become apparent if you re-read Section 2.3.2.

Exercise 7
Anaphors can occur inside DPs, as we have seen. These DPs can, under the right circumstances, undergo *wh*-movement, giving sentences like (1):

1. [$_{DPj}$ Which songs about himself$_i$] does Mick$_i$ particularly like t_j ?

To bring (1) into line with the binding theory, we have to propose that there is an LF operation of reconstruction which 'puts the *wh*-DP back' in the position of the trace. Once in the position of the trace, the reflexive can be bound in conformity with Principle A, as you should be able to see.

Barss (1986) discussed more complicated cases of reconstruction, such as the following:

2. [$_{DPj}$ Which songs about himself$_i$] did Mick$_{(i)}$ say that Keith$_{(i)}$ likes t_j ?

As the indexing indicates, either Mick or Keith can be the antecedent of himself in (2). How can we account for this using successive cyclicity?

A further complication arises where a VP is fronted. Here only the lowest interpretation is available. We can see this if we disambiguate the possible antecedents as in (3):

3. *[$_{VPj}$ Talk about himself$_i$], Mick$_i$ said that Sheena never did t_j
4. [$_{VPj}$ Talk about himself$_i$], Sheena said that Mick$_i$ never did t_j

Can you think of a way to exploit the VP-internal subject hypothesis (2.3.4) and the binding theory in order to account for (3) and (4) as well as the contrast with (2)? See Huang (1993).

Appendix: Syntactic Scope and Logical Scope

In the Appendix to the last chapter, I introduced the operation of Quantifier Raising (QR) and the idea that there is covert *wh*-movement in the mapping to LF. I also introduced the idea of logical scope. Here I want to follow up that discussion and introduce some ideas about the relation between syntactic scope and logical scope.

We saw in the last Appendix that in predicate logic the scope of a quantifier is the contents of the parenthesis to its right in the formula. We also saw that logical variable-binding corresponds to an A'-binding relation (in the sense of c-command and coindexation) between a quantifier and its trace (the trace is technically a *wh*-trace in all instances, subject to Principle C of the

binding theory). A natural extension of this is to say that *the scope of a quantifier is its c-command domain at LF.*

Now, clauses can contain two quantifiers which can show relative scope ambiguities. Here is an example:

(A1) Someone loves everyone

The two interpretations of (A1) are: (i) that there is a single maximally philanthropic individual, i.e. one lover for the whole world, and (ii) that everyone has a lover. The difference is this: in reading (i) everyone has the same lover, while in reading (ii) everyone may have a different lover. (You might notice that reading (i) entails reading (ii), that is, reading (ii) will be true whenever reading (i) is true; but reading (ii) does not entail reading (i), as reading (ii) can be true where people have different lovers, but reading (i) is false on this interpretation.)

In predicate logic, we can represent the different readings as follows (translating 'someone' and 'everyone' slightly inaccurately as 'pure' quantifiers):

(A2) (i) $\exists x\ (\forall y\ (\text{Love}\ (x, y)))$
(ii) $\forall y\ (\exists x\ (\text{Love}\ (x, y)))$

The relative scope of the quantifiers is clearly indicated by their order in the formulae. The rules of interpretation will guarantee that the entailment relations come out correctly, something that we don't need to go into here (see Allwood, Andersson and Dahl (1977), for example).

If, in our LF-representations, scope is determined by c-command, then relative scope should be determined by relative c-command. The ambiguity of (A1) implies that there must be two distinct LF representations, one in which *someone* c-commands *everyone*, and one in which the relations are the other way around. The earliest theory of LF was May (1977), and he proposed that these c-command asymmetries arose as the result of different orders of adjunction to IP. So, on this view, the two readings of (A1) correspond to the LFs in (A3):

(A3) (i) $[_{IP}\ \text{someone}_i\ [_{IP}\ \text{everyone}_j\ [_{IP}\ t_i\ \text{loves}\ t_j\ ]]]$
(ii) $[_{IP}\ \text{everyone}_j\ [_{IP}\ \text{someone}_i\ [_{IP}\ t_i\ \text{loves}\ t_j\ ]]]$

It's clear that where one quantifier Q_i has wider relative scope than another quantifier Q_j, Q_i c-commands Q_j – and conversely. C-command relations thus feed into the semantic interpretation rules at the LF interface, so as to determine the different entailment relations.

More recently, Hornstein (1995) has developed a variant theory which exploits the same basic idea – that scope relations are determined by c-command relations – in a minimalist framework. Hornstein adopts the analysis of transitive clauses that we sketched in 2.6; subjects are base-generated in Spec,VP and objects in the complement of V. By LF, the subject raises to Spec,AgrSP and the object to Spec,AgrOP. Hornstein adopts one further minimalist assumption that I haven't mentioned up to now: that traces are really

copies of moved categories (this is an idea with potentially far-reaching implications, as you can probably see). In both PF and LF, one copy of a moved element must be deleted. In PF, this is the one we've been calling a trace – the one you don't hear. *But you don't have to delete the same one in LF.* To see how this can be made into an account of relative scope, let's look at the representation of a sentence containing two quantifiers prior to copy-deletion:

(A4) $[_{\text{SpecAgrSP}}$ someone ... $[_{\text{SpecAgrOP}}$ everyone $[_{VP}$ someone loves everyone]]]

Now you must delete one copy of each quantifier, and at LF you're free to decide which. If you delete the upper copy of *someone* but the lower copy of *everyone*, then you get an interface representation like (A5):

(A5) $[_{\text{SpecAgrSP}}$ ... $[_{\text{SpecAgrOP}}$ everyone $[_{VP}$ someone loves]]]

This gives us reading (ii). If you delete the lower copy of *someone*, you get reading (i) whatever you do with *everyone*, and the same result ensues if you delete the upper occurrence of *everyone*. The crucial point, however, is that the copy + deletion approach gives us the possibility of two different LFs corresponding to the different relative scopes. And we retain the idea that relative scope is determined by c-command relations. Hornstein's approach does not appeal to a special rule like QR; the DPs are raised for checking reasons (this implies that they are in A-positions; combined with the fact they don't have traces in the sense we've seen up to now, this entails that a different approach to variable-binding has to be adopted: see Hornstein (1995) for details).

wh-elements can interact scopally with other quantifiers, as in (A6):

(A6) What did everyone buy for Bill?

Here *what* can be either inside or outside the scope of *everyone*. That is, the sentence can be either asking for a single answer ('everyone got together and bought him a gold watch') or a 'pair-list' answer ('Mick bought him a yacht; Keith bought him a spoon; Charlie bought him a record token', etc.). On the first interpretation, *what* is outside the scope of *everyone* – so we have one answer for everyone. On the second interpretation, the scope relations are reversed, and so everyone has a (potentially) different answer. There has been quite a bit of work on these interactions, starting with May (1985).

When two *wh*-quantifiers appear together in a single clause, they typically require a pair-list reading:

(A7) To whom did Johnny dedicate which song?

The natural answer to this question consists of pairs of people and songs. Higginbotham and May (1981) proposed that what happens here is that a complex *wh*-quantifier is created at LF by the absorption of the two *wh*-elements in Spec,CP at LF:

(A8) $[_{CP}$ {To whom$_i$, which song$_j$} [did Johnny dedicate t_j t_i]] ?

(I'm avoiding the technical details of absorption here.)

Finally, *wh*-movement at LF appears to show C-t effects. This is shown by the phenomenon of superiority, illustrated by the following contrast:

(A9) Who_i t_i drank what_j?

(A10) *What_j did who_i drink t_j ?

If *who* has to raise at LF, we can attribute the ungrammaticality of (A10) to the general impossibility of extracting a subject following a filled complementizer: the C-t effect discussed at length in this chapter. Note that we must then assume that there is a null *wh* complementizer in embedded clauses in order to account for superiority effects in embedded clauses (see 1.4.2.2):

(A11) *I can't remember what_j who_i drank t_j

If we adopt the Rizzi-Manzini account of C-t effects discussed in 4.5.2 and 4.6.5, one that relies on Spec-head agreement in CP, then this requirement must extend to LF.

The proposal that relative scope is determined by c-command is not restricted to quantifiers, but extends to negation, for example. We see this if we look again at an example of ambiguous scope of negation, as in (121), repeated from 4.5.1:

(121) They don't exploit their fans [because they love them]

As we noticed in 4.5.1, negation here can have main-clause scope (giving the interpretation that they love their fans, and therefore don't exploit them), or adjunct scope (giving the interpretation that they exploit their fans for some reason other than love). Now, we've assumed throughout our discussion of adjuncts in this chapter that they are adjoined to VP. We've also assumed that negation is situated outside VP (see 1.4.1 in particular), perhaps in NegP. Since NegP c-commands VP and VP-adjuncts, the scope ambiguity of (121), and in particular the fact that extraction of the adjunct can only give an interpretation where the adjunct is outside the scope of negation (as we saw in (122)), indicate that the adjunct can appear higher up than VP. It seems that we must allow different levels of attachment of adjuncts, in order to capture the scope ambiguities with negation that we observe. A natural possibility would be to allow adjunction to AgrSP (this would be a problem for the *Barriers* account of the CED (see 4.4.1.3), but not for Manzini's – you should be able to see this if you look again at 4.6.4). Something further needs to be said so as to allow negation to apply *only* to the adjunct when it has adjunct scope. I'll leave this as an open question.

As you might imagine, I haven't covered all there is to say about LF here. My excuse for this is that LF is an invariant level, and the focus of this book is on comparative syntax – hence things that vary across languages. The most up-to-date book on LF is Hornstein (1995).

5

Principles, Parameters, and Language Acquisition

5.0 Introduction

In this chapter, I don't intend to present any more of the mechanics of syntactic theory. Instead, I want to look at how the things we've seen in the earlier chapters relate to language acquisition and language learnability. In 5.1, I'll go through the argument for the existence of an innate language faculty that comes from the poverty of the stimulus to language acquisition. In 5.2 we'll look at how the 'principles-and-parameters' approach to language universals and language acquisition can give us a handle on this problem. Here we'll briefly outline the implications of the principles-and-parameters approach for language acquisition. Finally, in 5.3, I'll say something about the implications of the principles-and-parameters approach for the study of historical change in language.

5.1 The Argument from the Poverty of the Stimulus

5.1.1 The Nature of the Final State

In this section I review the well-known argument for the innateness of the language faculty that stems from the observed poverty of the stimulus to language acquisition. The idea is that the stimulus to language acquisition is deficient in two ways, which we will refer to as the *noisiness* and the *incompleteness* of the input to language acquisition.

Before considering the argument in detail, it is perhaps worthwhile to recall the nature of adult competence, or the kind of knowledge that makes up the adult I-language. Of course, the entire book up to now has been about this. The detailed, intricate complexity of the language faculty should be very apparent to you by now. It's only when one is aware of this complexity that one can fully appreciate the force of the argument from the poverty of the stimulus.

So let's look at some of the things an adult native speaker of English knows. Speakers of English can distinguish sentences with the grammatical SVO order from those with other imaginable orders, as illustrated by the following sentences:

Bibliography

Abney, S. (1987) *The English Noun Phrase in its Sentential Aspect.* PhD dissertation, MIT.

Aldridge, M. (1988) *The Acquisition of INFL*, University College of North Wales Monographs in Linguistics, 1.

Allwood, J., Andersson, L.-G. and Dahl, Ö. (1977) *Logic in Linguistics.* Cambridge: Cambridge University Press.

Anderson, J. (1971) *The Grammar of Case: Towards a Localistic Theory.* Cambridge: Cambridge University Press.

Anderson, S. (1982) 'Where's Morphology?' *Linguistic Inquiry* 13: 571–612.

Aoun, J. (1985) *A Grammar of Anaphora.* Cambridge, Mass.: MIT Press.

Aoun, J. (1986) *Generalized Binding.* Dordrecht: Foris.

Aoun, J., Hornstein, N., Lightfoot, D. and Weinberg, A. (1987) 'Two Types of Locality', *Linguistic Inquiry* 18: 537–77.

Aoun, J. and A. Li (1994) '*Wh*-in-Situ: Syntax of LF', *Linguistic Inquiry* 24: 199–238.

Atkinson, M. (1992) *Children's Syntax.* Oxford: Blackwell.

Baker, L. (1991) 'The Syntax of English *Not*: The Limits of Core Grammar', *Linguistic Inquiry* 22: 387–430.

Baker, M. (1988) *Incorporation: A Theory of Grammatical-Function Changing.* Chicago: Chicago University Press.

Baker, M., Johnson, K. and Roberts, I. (1989) 'Passive Arguments Raised', *Linguistic Inquiry* 20: 219–251.

Barss, A. (1986) *Chains and Anaphoric Dependence.* PhD dissertation, MIT.

Barss, A. and H. Lasnik (1986) 'A Note on Anaphora and Double Objects', *Linguistic Inquiry* 17: 347–54.

Belletti, A. (1988) 'The Case of Unaccusatives', *Linguistic Inquiry* 19: 1–35.

Belletti, A. (1990) *Generalized Verb Movement.* Turin: Rosenberg and Sellier.

Belletti, A. and L. Rizzi (1988) 'Psych Verbs and θ-Theory', *Natural Language and Linguistic Theory* 6: 291–352.

Benmamoun, E. (1991) 'Negation and Verb Movement', in T. Sherer (ed.), *Proceedings of NELS* 21: 17–31.

Bennis, H. and T. Hoekstra (1984) 'Gaps and Parasitic Gaps', *The Linguistic Review* 4: 29–87.

Berwick, R. (1985) *The Acquisition of Syntactic Knowledge.* Cambridge, Mass.: MIT Press.

den Besten, H. (1983) 'On the Interaction of Root Transformations and Lexical Deletive Rules', in W. Abraham (ed.), *On the Formal Syntax of the Westgermania.* Amsterdam: John Benjamins.

Bickerton, D. (1984) 'The Language Bioprogram Hypothesis', *Behavioral and Brain Sciences* 7: 173–221.

Blake, B. (1994) *Case.* Cambridge: Cambridge University Press.

Borer, H. (1984) *Parametric Syntax.* Dordrecht: Foris.

Borer, H. (1989) 'Anaphoric Agr', in O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter.* Dordrecht: Kluwer, 69–110.

Borer, H. and K. Wexler (1987) 'The Maturation of Syntax', in T. Roeper and E. Williams (eds.), *Parameter Setting.* Dordrecht: Kluwer, 123–72.

Borsley, R. D. (1991) *Syntactic Theory: A Unified Approach.* London: Edward Arnold.

Borsley, R. D. (forthcoming) *Modern Phrase Structure Grammar.* Oxford: Blackwell.

Borsley, R. D. and I. Roberts (eds.) (1996), *The Syntax of the Celtic Languages.* Cambridge: Cambridge University Press.

Bouchard, D. (1984) *On the Content of Empty Categories.* Dordrecht: Foris.

Brandi, L. and P. Cordin (1989) 'Two Italian Dialects and the Null Subject Parameter', in O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter.* Dordrecht: Kluwer, 111–42.

Bresnan, J. (1972) *Theory of Complementation in English.* PhD dissertation, MIT.

Bresnan, J. (1982a) 'Control and Complementation', *Linguistic Inquiry* 13: 343–434.

Bresnan, J. (1982b) *The Mental Representation of Grammatical Relations.* Cambridge, Mass.: MIT Press.

Brody, M. (1985) 'On the Complementary Distribution of Empty Categories', *Linguistic Inquiry* 16: 505–46.

Brody, M. (1995) *Lexico-Logical Form.* Cambridge, Mass.: MIT Press.

Brown, R. and C. Hanlon (1970) 'Derivational Complexity and the Order of Acquisition in Child Speech', in J. R. Hayes (ed.), *Cognition and the Development of Language.* New York: Wiley, 155–207.

Burzio., L. (1986) *Italian Syntax: A Government-Binding Approach.* Kluwer: Dordrecht.

Chomsky, N. (1955) *The Logical Structure of Linguistic Theory.* New York: Plenum (published in 1975).

Chomsky, N. (1957) *Syntactic Structures.* The Hague: Mouton.

Chomsky, N. (1964) *Current Issues in Linguistic Theory.* The Hague: Mouton.

Chomsky, N. (1965) *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

Chomsky, N. (1970) 'Remarks on Nominalizations', in R. Jacobs and P. S. Rosenbaum (eds.), *English Transformational Grammar.* Waltham, Mass.: Ginn, 184–221.

Chomsky, N. (1973) 'Conditions on Transformations', in S. Anderson and P.

Kiparsky (eds.), *A Festschrift for Morris Halle.* New York: Holt, Rinehart and Winston, 232–86.
Chomsky, N. (1976) 'Conditions on Rules of Grammar', *Linguistic Analysis* 2(4).
Chomsky, N. (1977) 'On WH-Movement', in A. Akmajian, P. Culicover, and T. Wasow (eds.), *Formal Syntax.* New York: Academic Press.
Chomsky, N. (1980) *Rules and Representations.* New York: Columbia University Press.
Chomsky, N. (1981) *Lectures on Government and Binding.* Dordrecht: Foris.
Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding.* Cambridge, Mass.: MIT Press.
Chomsky, N. (1986a) *Knowledge of Language: Its Nature, Origin and Use.* New York: Praeger.
Chomsky, N. (1986b) *Barriers.* Cambridge, Mass.: MIT Press.
Chomsky, N. (1991) 'Some Notes on Economy of Derivation and Representation', in R. Friedin (ed.), *Principles and Parameters in Comparative Grammar,* 417–54.
Chomsky, N. (1993) 'A Minimalist Program for Linguistic Theory', in K. Hale and S. J. Keyser (eds.), *The View from Building 20.* Cambridge, Mass.: MIT Press, 1–52.
Chomsky, N. (1995) 'Categories and Transformations', ms. MIT.
Chomsky, N. and H. Lasnik (1993) 'Principles and Parameters Theory', in J. Jacobs, A. von Stechow, W. Sternefeld, and T. Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research.* Berlin: de Gruyter.
Cinque, G. (1991) *Types of A'-Dependencies.* Cambridge, Mass.: MIT Press.
Clements, N. (1975) 'The Logophoric Pronoun in Ewe: Its Role in Discourse', *Journal of West African Languages,* 10: 141–77.
Comrie, B. (1981) *Language Universals and Linguistic Typology.* Chicago: Chicago University Press.
Croft, W. (1990) *Universals and Typology.* Cambridge: Cambridge University Press.
Duffield, N. (1995) *Particles and Projections in Irish Syntax.* Dordrecht: Kluwer.
Emonds, J. (1976) *A Transformational Approach to English Syntax: Root, Structure-Preserving and Local Transformations.* New York: Academic Press.
Emonds, J. (1978) 'The Verbal Complex V-V' in French', *Linguistic Inquiry* 9:151–75.
Emonds, J. (1980) 'Word Order in Generative Grammar', *Journal of Linguistic Research* 1:33–54.
Engdahl, E. (1983) 'Parasitic Gaps', *Linguistics and Philosophy* 6: 534.
Engdahl, E. (1985) 'Parasitic Gaps, Resumptive Pronouns and Subject Extractions', *Linguistics* 23: 3–44.
Everaert, M. (1986) *The Syntax of Reflexivization.* Dordrecht: Foris.
Faltz, L. (1977) *Reflexivization: A Study in Universal Syntax.* PhD dissertation, UC Berkeley.
Fiengo, R. and J. Higginbotham (1981) 'Opacity in NP', *Linguistic Analysis* 7: 395–422.

Fillmore, C. (1968) 'The Case for Case', in E. Bach and J. Harms (eds.), *Universals in Linguistic Theory,* 1–88.
Frampton, J. (1990) 'Parasitic Gaps and the Theory of *Wh*-Chains', *Linguistic Inquiry* 21: 49–78.
Fukui, N. and P. Speas (1986) 'Specifiers and Projections', in N. Fukui, T. Rapoport, and E. Sagey (eds.), *Papers in Theoretical Linguistics: MIT Working Papers in Linguistics,* 128–172.
Gazdar, G., Klein, E., Pullum, G. and Sag, I. (1985) *Generalized Phrase Structure Grammar.* Cambridge, Mass.: Harvard University Press.
Giorgi, A. (1984) 'Towards a Theory of Long-Distance Anaphora: A GB Approach', *The Linguistic Review* 3: 307–359.
Giorgi, A. and G. Longobardi (1991) *The Syntax of Noun Phrases.* Cambridge: Cambridge University Press.
Gold, E. (1967) 'Language Identification in the Limit', *Information and Control* 16: 447–74.
deGraff, M. and A. Pierce (1996) *Creoles, Diachrony and Language Acquisition.* Cambridge, Mass.: MIT Press.
Greenberg, J. (1963) 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements', in J. Greenberg (ed.), *Universals of Language.* Cambridge, Mass.: MIT Press.
Grimshaw, J. (1990) *Argument Structure.* Cambridge, Mass.: MIT Press.
Gruber, J. (1965) *Studies in Lexical Relations.* PhD dissertation, MIT.
Haegeman, L. (1995) *The Syntax of Negation.* Cambridge: Cambridge University Press.
Hale, K., Jeanne, L. M. and Platero, P. (1977) 'Three Cases of Overgeneration', in A. Akmajian, P. Culicover and T. Wasow (eds.), *Formal Syntax.* New York: Academic Press.
Hale, K. and S. J. Keyser (1993) 'On Argument Structure and the Lexical Expression of Syntactic Relations', in K. Hale and S.J. Keyser (eds.), *The View from Building 20.* Cambridge, Mass.: MIT Press, 53–109.
Hawkins, J. (1983) *Word Order Universals.* New York: Academic Press.
Hawkins, J. (ed) (1988) *Explaining Language Universals.* Oxford: Blackwell.
Heim, I., Lasnik, H. and May, R. (1991) 'Reciprocity and Plurality', *Linguistic Inquiry* 22: 63–101.
Hendrick, R. (1988) *Anaphora in Celtic and Universal Grammar.* Dordrecht: Kluwer.
Higginbotham, J. (1980) 'Pronouns and Bound Variables', *Linguistic Inquiry* 11: 679–708.
Higginbotham, J. (1981) 'Reciprocal Interpretation', *Journal of Linguistic Research* 1: 97–117.
Higginbotham, J. (1983) 'Logical Form, Binding and Nominals', *Linguistic Inquiry* 16: 547–93.
Higginbotham, J. (1985) 'On Semantics', *Linguistic Inquiry* 16: 547–593.
Higginbotham, J. and R. May (1981) 'Questions, Quantifiers and Crossing', *The Linguistic Review* 1: 41–79.

Hoekstra, T. and B. Schwartz (eds.), (1994) *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins.
Hornstein, N. (1985) *Logic as Grammar*. Cambridge, Mass.: MIT Press.
Hornstein, N. (1995) *Logical Form*. Oxford: Blackwell.
Hornstein, N. and D. Lightfoot (1981) *Explanation in Linguistics*. London: Longman.
Huang, C.T.J. (1982) *Logical Relations in Chinese and the Theory of Grammar*. PhD dissertation, MIT.
Huang, C.T.J. (1984) 'On the Distribution and Reference of Empty Pronouns', *Linguistic Inquiry* 15: 531–74.
Huang, C.T.J. (1993) 'Reconstruction and the Structure of VP: Some Theoretical Consequences', *Linguistic Inquiry* 24.
Hurford, J. (1994) *Grammar: A Student's Guide*. Cambridge: Cambridge University Press.
Hyams, N. (1986) *Language Acquisition and the Theory of Parameters*. Dordrecht: Kluwer.
Iatridou, S. (1990) 'About AGR(P)', *Linguistic Inquiry* 21:551–77.
Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. Cambridge, Mass.: MIT Press.
Jackendoff, R. (1977) *X'-Theory*. Cambridge, Mass.: MIT Press.
Jaeggli, O. (1982) *Topics in Romance Syntax*. Dordrecht: Foris.
Jaeggli, O. (1986) 'Passive' *Linguistic Inquiry* 17: 587–633.
Jaeggli, O. and K. Safir (eds.), (1989) *The Null Subject Parameter*. Dordrecht: Kluwer.
Jones, R. and A. Thomas (1977) *The Welsh Language: Studies in its Syntax and Semantics*. Cardiff: University of Wales Press.
Kayne, R. (1975) *French Syntax*. Cambridge, Mass.: MIT Press.
Kayne, R. (1984) *Connectedness and Binary Branching*. Dordrecht: Foris.
Kayne, R. (1989a) 'Notes on English Agreement', *CIEFL Bulletin*, Hyderabad, India.
Kayne, R. (1989b) 'Facets of Romance Past Participle Agreement,' in P. Benincà (ed.), *Dialect Variation on the Theory of Grammar*. Dordrecht: Foris, 85–104.
Kayne, R. (1991) 'Romance Clitics, Verb Movement and PRO', *Linguistic Inquiry* 22: 647–86.
Kayne, R. (1993) 'The Antisymmetry of Syntax', ms. CUNY.
Kayne, R. (1994) *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
Keenan, E. (1988) *Universal Grammar: 15 Essays*. London: Croom Helm.
Keyser, S. J. and T. Roeper (1984) 'On the Middle and Ergative Constructions in English', *Linguistic Inquiry* 15: 381–416.
Kitagawa, Y. (1986) *Subjects in Japanese and English*. PhD dissertation, University of Massachusetts, Amherst.
Koopman, H. (1984) *Verb Movement and Universal Grammar*. Dordrecht: Foris.
Koopman, H. and D. Sportiche (1982) 'Variables and the Bijection Principle', *The Linguistic Review* 2: 139–60.
Koopmann, H. and D. Sportiche (1991) 'The Position of Subjects', *Lingua* 85: 211–58.

Koster, J. (1984) 'On Binding and Control', *Linguistic Inquiry* 15: 417–59.
Koster, J. (1985) 'Reflexives in Dutch', in J. Guéron, H. G. Obenauer and J.-Y. Pollock (eds.), *Grammatical Representation*. Dordrecht: Foris.
Koster, J. (1987) *Domains and Dynasties: The Radical Autonomy of Syntax*. Dordrecht: Foris.
Koster, J. and E. Reuland (eds.) (1991) *Long Distance Anaphora*. Cambridge: Cambridge University Press.
Kroch, A. (1989) 'Reflexes of Grammar in Patterns of Language Change', *Journal of Language Variation and Change* 1: 199–244.
Kuroda, Y. (1988) 'Whether We Agree or Not: A Comparative Syntax of English and Japanese,' in W. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*. Center for the Study of Language and Information, Stanford University.
Labov, W. (1972) *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.
Lasnik, H. (1976) 'Remarks on Coreference', *Linguistic Analysis* 2: 122.
Lasnik, H. (1989) *Essays on Anaphora*. Dordrecht: Kluwer.
Lasnik, H. and M. Saito (1984) 'On the Nature of Proper Government', *Linguistic Inquiry* 14: 235–89.
Lasnik, H. and M. Saito (1991) 'On the Subject of Infinitives', *Proceedings of the Chicago Linguistics Society*, 27: 324–43.
Lasnik, H. and M. Saito (1992) *Move-α*. Cambridge, Mass.: MIT Press.
Lebeaux, D. (1983) 'A Distributional Difference between Reciprocals and Reflexives', *Linguistic Inquiry* 14: 723–30.
Lenneberg, E. (1967) *Biological Foundations of Language*. New York: Wiley.
Levin, B. and M. Rappaport (1986) 'The Formation of Adjectival Passives', *Linguistic Inquiry* 17: 623–62.
Levin, B. and M. Rappaport-Hotav (1995) *Unaccusativity*. Cambridge, Mass.: MIT Press.
Lightfoot, D. (1991) *How to Set Parameters*. Cambridge, Mass.: MIT Press.
Lightfoot, D. and N. Hornstein (1994) *Verb Movement*. Cambridge: Cambridge University Press.
Longobardi, G. (1985) 'Connectedness, Scope and C-command', *Linguistic Inquiry* 16: 163–92.
McCloskey, J. (1983) 'A VP in a VSO language', in G. Gazdar, E. Klein, and G. K. Pullum (eds.), *Order, Concord and Constituency*. Dordrecht: Foris.
McCloskey, J. (1991), 'Clause structure, ellipsis and proper government in Irish', *Lingua* 85: 259–302.
McCloskey, J. (1996), 'On the scope of verb-movement in Irish', *Natural Language and Linguistic Theory* 14: 47–104.
McLuhan, M. (1964) *Understanding Media: The Extensions of Man*. New York: McGraw-Hill.
McNeill, D. (1966) 'Developmental Psycholinguistics,' in F. Smith and G. Miller (eds.), *The Genesis of Language: A Psycholinguistic Approach*. Cambridge,

Mass.: MIT Press, 15–84.
Manzini, M.-R. (1983) 'On Control and Control Theory', *Linguistic Inquiry* 14: 421–46.
Manzini, M.-R. (1992) *Locality: A Theory and Some of its Empirical Consequences*. Cambridge, Mass.: MIT Press.
Marantz, A. (1984) *On the Nature of Grammatical Relations*. Cambridge, Mass.: MIT Press.
May, R. (1977) *The Grammar of Quantification*. PhD dissertation, MIT.
May, R. (1985) *Logical Form: Its Structure and Derivation*. Cambridge, Mass.: MIT Press.
Mohammad, M. (1988) *The Sentential Structure of Arabic*. PhD dissertation, University of Southern California.
Munn, A. (1993) *Topics in the Syntax and Semantics of Coordinate Structures*. PhD dissertation, University of Maryland.
Newport, E., Gleitman, L. and Gleitman, H. (1977) ' "Mother, I'd Rather Do It Myself": Some Effects and Non-effects of Maternal Speech Style', in C. E Snow and C.A. Ferguson (eds.), *Talking to Children, Language Input and Acquisition*. Cambridge: Cambridge University Press.
Osherson, D., Stob, M. and Weinstein, S. (1982) 'Learning Strategies', *Information and Control* 53: 32–51.
Osherson, D., Stob, M. and Weinstein, S. (1984) 'Learning Theory and Natural Language', *Cognition* 17:128.
Osherson, D., Stob, M. and Weinstein, S. (1986) *Systems That Learn*. Cambridge, Mass.: MIT Press.
Ouhalla, J. (1994) *Introducing Transformational Grammar*. London: Edward Arnold.
Perlmutter, D. (1971) *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.
Perlmutter, D. (1978) 'Impersonal Passives and the Unaccusative Hypothesis', *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, 157–89.
Perlmutter, D. (ed.) (1983) *Studies in Relational Grammar I*. Chicago: University of Chicago Press.
Perlmutter, D. and P. Postal (1984) 'The 1-Advancement Exclusiveness Law', in D. Perlmutter and C. Rosen (eds.), *Studies in Relational Grammar II*. Chicago: Chicago University Press.
D. Perlmutter and C. Rosen (eds.) (1984), *Studies in Relational Grammar II*. Chicago: Chicago University Press.
Pesetsky, D. (1982) *Paths and Categories*. PhD dissertation, MIT.
Pica, P. (1987) 'On the Nature of the Reflexivization Cycle', *Proceedings of NELS 17* 2: 483–99.
Pierce, A. (1992) *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Dordrecht: Kluwer.
Pierce, A. and V. Deprez (1993) 'Negation and Functional Projections in Early Grammar', *Linguistic Inquiry* 24: 25–68.

Poletto, C. (1993) *La Sintassi dei Clitici Soggetto nei Dialetti Italiani Settentrionali*, Doctoral Dissertation, Universities of Venice and Padua.
Pollard, C. and I. Sag (1994) *Head Driven Phrase Structure Grammar*. Chicago: Chicago University Press.
Pollock, J.-Y. (1989) 'Verb Movement, Universal Grammar and the Structure of IP', *Linguistic Inquiry* 20: 365–424.
Postal, P. (1969) 'On the So-Called "Pronouns" in English' in D. Reibel and S. Schane (eds.), *Modern Studies in English*. Englewood Cliffs, NJ: Prentice-Hall, 201–24.
Postal, P. (1971) *Crossover Phenomena*. New York: Holt, Rinehart and Winston.
Postal, P. (1974) *On Raising*. Cambridge, Mass.: MIT Press.
Postal, P. (1986) *Studies in Passive Clauses*. Albany, N.Y.: State University of New York Press.
Radford, A. (1986) 'Small Children's Small Clauses', *Bangor Research Papers in Linguistics*, 1. Bangor: University of Wales.
Radford, A. (1990) *Syntactic Theory and the Acquisition of English*. Oxford: Blackwell.
Reinhart, T. and E. Reuland (1991) 'Anaphors and Logophors: An Argument Structure Perspective', in Koster, J. and E. Reuland (eds.) (1991), *Long Distance Anaphora*. Cambridge: Cambridge University Press, 283–321.
Reinhart, T. and E. Reuland (1993) 'Reflexivity', *Linguistic Inquiry* 24.
van Riemsdijk, H. (1978) *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*. Dordrecht: Foris.
Rizzi, L. (1982) *Issues in Italian Syntax*. Dordrecht: Foris.
Rizzi, L. (1986a) 'Null Objects in Italian and the Theory of *pro*', *Linguistic Inquiry* 17: 501–57.
Rizzi, L. (1986b) 'On the Status of Subject Clitics in Romance,' In O. Jaeggli and C. Silva-Corvalán (eds.), *Studies in Romance Syntax*. Dordrecht: Foris, 391–419.
Rizzi, L. (1986c) 'On Chain Formation', in H. Borer (ed.), *The Syntax of Pronominal Clitics: Syntax and Semantics*, 19. New York: Academic Press.
Rizzi, L. (1988) 'The Structural Uniformity of Syntactic Categories', in *Proceedings of the Conference on the Basque Language*, Eusko Jaurlaritzaren Argitalpen-Zerbitzu Nagusia/Servicio Central de Publicaciones del Gobierno Vasco, Vitoria/Gasteiz.
Rizzi, L. (1990) *Relativized Minimality*. Cambridge, Mass.: MIT Press.
Rizzi, L. (1991) 'Residual Verb Second and the WH-Criterion', *University of Geneva Technical Reports in Formal and Computational Linguistics*, 2.
Rizzi, L. (1996) 'Some Notes on Linguistic Theory and Language Development: The Case of Root Infinitives', *Language Acquisition* 6.
Rizzi, L. and I. Roberts (1989) 'Complex Inversion in French', *Probus* 1: 1–39.
Roberts, I. (1987) *The Representation of Implicit and Dethematized Subjects*. Dordrecht: Foris.
Roberts, I. and U. Shlonsky (1996) 'Pronominal Enclisis and VSO Languages', in R. D. Borsley and I. Roberts (eds.), *Comparative Celtic Syntax*. Cambridge: Cambridge University Press, 171–99.

Rosenbaum, P. (1967) *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass.: MIT Press.
Ross, J. R. (1967) *Constraints on Variables in Syntax*. PhD dissertation, MIT.
Ross, J.R. (1986) *Infinite Syntax!* Norwood, NJ: Ablex.
Roussou, A. (1993) 'Factivity, Factive Islands and the *that-t* Filter', in *Proceedings of CONSOLE* 1.
Rouveret, A. and J.-R. Vergnaud (1980) 'Specifying Reference to the Subject: French Causatives and Conditions on Representations', *Linguistic Inquiry* 11: 97–202.
Rudin, J. (1988) 'On Multiple Questions and Multiple WH-Fronting', *Natural Language and Linguistic Theory* 6: 445–502.
Safir, K. (1985) *Syntactic Chains*. Cambridge: Cambridge University Press.
Shopen, T. (1985) *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press.
Sportiche, D. (1981) 'Bounding Nodes in French', *The Linguistic Review* 1: 219–46.
Sportiche, D. (1983) *Structural Invariance and Asymmetry in Syntax*. PhD dissertation, MIT.
Sproat, R. (1985) 'Welsh Syntax and VSO Structure', *Natural Language and Linguistic Theory* 3: 173–216.
Stowell, T. (1981) *The Origins of Phrase Structure*. PhD dissertation, MIT.
Taraldsen, T. (1979) 'The Theoretical Interpretation of a Class of Marked Extractions', in A. Belleti, L. Brandi and L. Rizzi (eds.), *Theory of Markedness in Generative Grammar*. Pisa: Scuola Normale.
Thiersch, C. (1993) 'On the Formal Properties of Constituent Coordination', *GLOW Newsletter* 30: 70–71.
Tomaselli, A. (1989) *La sintassi del verbo finito nelle lingue germaniche*. PhD dissertation, University of Pavia.
Travis, L. (1984) *Parameters and Effects of Word Order Variation*. PhD dissertation, MIT.
Travis, L. (1991) 'Parameters of Phrase Structure and Verb-Second Phenomena', in R. Friedin (ed.), *Principles and Parameters in Comparative Grammar*. Cambridge, Mass.: MIT Press, 339–364.
Vergnaud, J.-R. (1985) *Dépendances et niveaux de représentation en syntaxe*. Amsterdam: John Benjamins.
Vikner, S. (1994) *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford/New York: Oxford University Press.
Wasow, T. (1972) *Anaphoric Relations in English*. PhD dissertation, MIT Press.
Wasow, T. (1979) *Anaphora in Generative Grammar*. Ghent: E. Story-Scientia.
Watanabe, A. (1992) '*Wh*-in-situ, Subjacency and Chain Formation', *MIT Occasional Papers in Linguistics*, 2.
Watanabe, A. (1993) *Case Absorption*. PhD dissertation, MIT.
Webelhuth, G. (ed.) (1995), *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell.
Wexler, K. (1994) 'Finiteness and Head Movement in Early Child Grammars', in D. Lightfoot and N. Hornstein (eds.), *Verb Movement*. Cambridge: Cambridge University Press, 305–50.
Wexler, K. and M.-R. Manzini (1987) 'Parameters, Learnability and Binding Theory', *Linguistic Inquiry* 18: 413–44.
Williams, E. (1980) 'Predication', *Linguistic Inquiry* 11: 203–38.
Williams, E. (1981) 'Argument Structure and Morphology', *The Linguistic Review* 1: 81–114.
Williams, E. (1994) 'A Reinterpretation of the Evidence for Verb Movement in French', in D. Lightfoot and N. Hornstein (eds.), *Verb Movement*. Cambridge: Cambridge University Press, 189–206.
Zagona, K. (1992) *Verb Phrase Syntax*. Dordrecht: Kluwer.
Zubizarreta, M.-L. (1987) *Levels of Representation in the Lexicon and Syntax*. Dordrecht: Foris.
Zwart, J.-W. (1993) *Dutch Syntax*. PhD dissertation, University of Groningen.