7 Wh-Movement

Contents

Introduction and overview

1 Wh-movement: some examples

2 Wh-phrases

3 The landing site of wh-movement
   3.1 Long vs. short movement
   3.2 C-command
   3.3 Wh-movement and substitution
   3.4 The doubly filled COMP filter
   3.5 Adjunction
       3.5.1 General discussion
       3.5.2 Wh-movement as adjunction?
       3.5.3 A note on some alternative proposals
   3.6 Movement of maximal projections: A-movement vs. A'-movement

4 Traces and wh-movement
   4.1 Theta theory and the projection principle
   4.2 Agreement and binding
   4.3 Case
       4.3.1 Wh-pronouns and case
       4.3.2 Wh-trace vs. NP-trace: more contrasts
   4.4 Adjunct traces

5 Subject movement
   5.1 Vacuous movement
   5.2 The that-trace filter
6 Bounding Theory
6.1 Island constraints: the data
6.1.1 The complex NP constraint
6.1.2 Wh-islands
6.1.3 Islands
6.2 Subjacency
6.2.1 The subjacency condition on movement
6.2.2 The complex NP constraint
6.2.3 Wh-islands
6.3 Subjacency as a diagnostic for movement
6.3.1 Left dislocation: movement and copying?
6.3.2 Relative clauses and wh-movement
6.3.3 Relative clauses and resumptive pronouns
6.3.4 NP-movement
6.4 The subjacency parameter
7 Binding theory and traces of wh-movement
7.1 Typology of NPs
7.2 Crossover
8 Movement to the right in English
8.1 Heavy NP-shift
8.2 PP-extraposition from NP
8.3 Conclusion
9 Summary
10 Exercises

Introduction and Overview

In chapter 6, section 1 we gave a survey of various types of movement. Movement affects either heads or maximal projections. Chapter 6 discussed movement of NPs in passive and raising patterns. In this chapter we turn to wh-movement. We discuss the moved constituent, its landing site and the arguments for positing traces in the extraction site. We show that the subjacency condition imposes a constraint on the range of wh-movement and is subject to parametric variation. Using the subjacency condition as a diagnostic we show that English relative clauses are derived via wh-movement.

We continue to elaborate our typology of empty categories, adding wh-traces which are of the type [–anaphor, –pronominal]. From our discussion it follows that heavy NP-shift and PP-extraposition from NP are also instantiations of wh-movement.

In section 1 we illustrate wh-movement in questions. Section 2 concerns the moved constituent; section 3 the landing site. In section 4 we consider arguments for traces of wh-movement. In section 5 we describe some special properties of subject extraction. Section 6 deals with the subjacency condition on movement. In section 7 we turn to the typology of empty categories and in section 8 we discuss heavy NP-shift and PP-extraposition from NP.

1 Wh-movement: Some Examples

In chapter 6 section 1.2 we gave a brief analysis of wh-questions such as (1).

1 Whom will Lord Emsworth invite?

The wh-constituent whom is the internal argument of invite: it is VP-internal at D-structure. We also assume that the auxiliary will is base-generated under I.
In this chapter we do not deal with the movement of the auxiliary and we concentrate on the movement of elements such as whom.¹

2 Wh-Phrases

In this section we discuss some of the main properties of the constituent which undergoes wh-movement in interrogative sentences.² In each of the wh-questions in (3) a constituent is fronted:

3a [What] will Poirot eat?
3b [Which detective] will Lord Emsworth invite?
3c [Whose pig] must Wooster feed?
3d [Where] will Jeeves live?
3e [When] will the detective arrive at the castle?
3f [Why] must Wooster feed the pig?
3g [To whom] will the police inspector give the money?
3h [In which folder] does Maigret keep the letters?
3i [How] will Jeeves feed the pigs?
3j [How big] will the reward be?

From (3) it is clear that the moved constituent is a phrase. Various types of constituents can move: NPs (3a, b, c), adverb phrases (3d, e, f, i), PPs (3g, h) and APs (3j). As the reader can verify for himself, the moved element may be both an argument of the verb or an adjunct.³ The moved constituent in (3)-will be referred to as a wh-phrase or a wh-constituent. The motivation for the label is transparent in (3a)–(3h): the moved constituent either consists of, or contains a word beginning with wh-. In (3i) and (3j) the wh-questions are not introduced by a word which begins with wh-, but we can paraphrase the examples using a wh-phrase:

4a [In what way] will Jeeves feed the pigs?
4b [Of what size] will the reward be?

¹ For a discussion of the movement of auxiliaries i.e. head-to-head movement, see chapters 10, 11 and 12.
² In sections 6 and 8 we shall see that the range of elements that undergo wh-movement includes non-interrogative elements.
³ In chapter 9 we shall see that the distinction between arguments and adjuncts is important. We return to this distinction in chapters 10 and 12.
The term *wh*-phrase will also be used to refer to the moved phrases in (3a) and (3j).

Let us consider the structure of the *wh*-phrases more carefully. In (3a), (3d), (3e), (3f) and (3i) the *wh*-element is itself the head of the moved phrase (cf. (5a)). In (3b), (3c) and (3j) the *wh*-element is the specifier of the moved phrase (5b). We turn to the PPs in (3g) and (3h) presently.

\[5a\] XP
\[5b\] XP
\[\begin{array}{l}
X' \\
\phantom{X} X \\
\phantom{X} [+WH] \\
\end{array}\]
\[\begin{array}{l}
\phantom{X'} \\
\phantom{X} Spec \\
\phantom{X} X' \\
\phantom{X} [+WH] \\
\end{array}\]
\[\begin{array}{l}
a. \text{what} \\
b. \text{which detective} \\
c. \text{whose pig} \\
d. \text{where} \\
e. \text{when} \\
f. \text{why} \\
g. \text{to whom} \\
h. \text{in which folder} \\
i. \text{how} \\
\end{array}\]

The presence of a *wh*-phrase as the complement of the preposition apparently suffices to allow the PP to undergo *wh*-movement. We might propose that the [+WH] feature of the NP percolates to the PP.

Compare (3g) and (3h) with the examples in (7):

\[7a\] [so] Whom will the police inspector give the money to?
\[7b\] [so] Which folder does Maigret keep the letters in?

In (7) the *wh*-phrase is moved out of the PP; the head of the PP is left behind. The phenomenon in which a preposition is left behind after its complement has been moved out is called preposition-stranding. The phenomenon where the preposition is moved along with the complement NP is referred to as pied-piping: the preposition is pied-piped with the NP in (3g) and (3h). In English pied-piping of prepositions is always legitimate. Preposition-stranding is restricted, as the following examples show:

\[8a\] ?Which party did Poirot meet Maigret after?
\[8b\] ?Whose office did the inspectors discuss the crime in?

Abbey (1987) proposes that NPs should be reinterpreted as projections of the determiner. If this analysis were to be adopted it would obviously also follow that a phrase whose determiner is [+WH] is itself characterized as [+WH] (cf. chapter 11). See Hornstein and Weinberg (1981) and Kayne (1984) for discussion.
The possibility of preposition-stranding is subject to cross-linguistic variation, as illustrated in the following examples:

9a **French**
*Qui as-tu parlé de?*  
who have-you talked about  
‘Who did you talk about?’ vs.  
De qui as-tu parlé?

9b **Italian**
*Cui hai parlato di?*  
who have-you talked about  
vs.  
Di cui hai parlato?

The question that should be answered is what explains the difference between languages with preposition-stranding, such as English, and languages without, such as Italian and French. Ideally the difference should be related to some difference in parameter setting between the languages. We do not discuss this issue here.

3 The Landing Site of Wh-Movement

In this section we discuss where the constituent which is affected by wh-movement is moved to, i.e. its landing site.

3.1 Long vs. Short Movement

Consider the following example:

10 Whom do you believe [CP that [Lord Emsworthk will invite]]?

(10) is a complex sentence. **Believe** takes two arguments: the external argument is realized by its subject you, the internal argument is clausal (CP). We have already seen that **invite** takes two arguments. The external argument is realized by the NP Lord Emsworth. It seems natural to say that whom is the internal argument. On this assumption, the D-structure of (10) should be (11a).

![Diagram of D-structure](image)

Under the subject-in-VP hypothesis, Lord Emsworth would originate in [Spec, VP] and move to its position in [Spec, IP]. We leave this aspect of the derivation aside for expository reasons.

At S-structure whom is moved from the lower clause to the [Spec, CP] of the higher clause, leaving a coindexed trace in its base-position (ti). In addition the auxiliary do is moved to C.

---

6 Kayne (1984), for example, relates the difference between English and other languages to the fact that in English prepositions assign structural ACCUSATIVE while in other languages prepositions assign an inherent case.
11b [o, Whom do [w, you, believe [cp, that [w, Lord Emsworth, will invite t]]]]

The difference between our earlier examples such as (1) and (3), where a wh-element moves to the [Spec, CP] position of the sentence in which it is theta-marked, and examples such as (11), where it moves beyond its own clause to the [Spec, CP] of a higher clause, is often made in terms of short vs. long movement.

We assume that the reader is familiar with the contrast between direct or root questions such as the ones discussed so far and indirect or embedded questions as illustrated in (12):

12a He wonders [if [Lord Emsworth will invite Poirot]].
12b I wonder [whom [Lord Emsworth will invite]].

One property that distinguishes root questions from embedded questions is that in the latter the auxiliary does not move: in (12) will has not inverted with the subject NP. The D-structure for (12a) will be as in (13). Wonder is a two-place predicate, which assigns its external theta role to be, the subject and the internal theta role to the interrogative clause which it governs. The realization of the arguments of invite is unproblematic: the external argument is the subject NP and the internal argument is the object NP. The S-structure of the sentence will also be as in (13) since no constituents are moved.

What about (12b)? Based on the preceding discussion we propose that the D-structure and the S-structure of (12b) are (14a) and (14b) respectively:
(14) exemplifies short *wh*-movement in embedded questions. The reader is invited to provide a description of the syntactic representation of (15), which is an example of long *wh*-movement in embedded questions:

15 I wonder whom they believe that Lord Emsworth will invite.

3.2 *C*-command

When we consider the configurational relation between the antecedent of *wh*-movement and the trace we see that, as in the case of NP-movement (cf.
chapter 6, section 2.3.2), the \( wh \)-antecedent c-commands its trace. We leave it to the reader to verify this in the preceding examples.

3.3 \( Wh \)-movement and Substitution

We have been assuming that the landing site of \( wh \)-movement is \([\text{Spec, CP}] \). The motivation for this proposal was discussed briefly in chapter 2.

A first and rather natural assumption is that an element moves into an unoccupied position. Remember that we adopted the structure preserving principle for transformations (chapter 6, section 4.1). Considering that \( wh \)-movement moves phrasal projections of different categories, it is not reasonable to claim that all these categorially distinct \( wh \)-constituents move to a position that is labelled for one specific category. Rather, the landing site for \( wh \)-movement must be a position which is not specified for the phrasal category. \([\text{Spec, CP}] \) is just such a position: the phrase structure rules allow us to project the position but they do not identify it for a specific phrasal category. A non-filled \([\text{Spec, CP}] \) can receive phrasal constituents of any syntactic category: NP, AP, etc. The proposal developed here treats \( wh \)-movement as substitution: the \( wh \)-phrase fills a hitherto unoccupied position. In this respect, \( wh \)-movement is like NP-movement.

3.4 The Doubly Filled COMP Filter

In chapter 2 we have already given empirical evidence for taking \([\text{Spec, CP}] \) to the position to the left of \( C \), as the landing site of \( wh \)-movement: in many languages we find sequences of a \( wh \)-word followed by an overt complementizer.

16a Dutch
Ik weet niet \( wie \) of Jan gezien heeft.
I know not whom whom Jan has seen has
'I don't know whom Jan has seen.'

16b Flemish (a dialect of Dutch)
Ik weet niet \( wie \) \( dat \) Jan gezien heeft.
I know not whom that Jan seen has
'I don't know whom Jan has seen.'

For discussion see Haegeman (1992).

For discussion of the Bavarian data, see Bayer (1984a, 1984b). The doubly filled COMP filter was first formulated in Chomsky and Lasnik (1977). There are a number of alternative analyses for the structure of the data in (16). (18) is only one example. We leave these divergencies out of the discussion here since they have become obsolete.
Following the convention in the current literature, we maintain the label doubly filled COMP filter here and reinterpret it according to the CP-analysis of clauses as suggested in (17). 11

The hypothesis that the landing site of wh-movement is the position [Spec, CP], leads us to expect that only one element can be moved to occupy this position:

19a John wondered which book Bill bought for whom.
19b John wondered for whom Bill bought which book.
19c *John wondered for whom which book Bill bought.
19d *John wondered which book for whom Bill bought.

(19c) and (19d) will be ruled out by our grammar since two phrases would have to be moved to [Spec, CP]. Recall that we cannot move a phrase under C because C is a head position. Multiple movement is not possible in English.

However, there are examples from Polish which are problematic (Lasnik and Saito, 1984: 280):

20 Maria zastanawiała się, kto co przyniesie.
       Maria wondered who what would-bring
       'Maria wondered who would bring what.'

(20) illustrates multiple wh movement: in this case two wh phrases have been moved to sentence-initial position. In the next section we shall discuss one possible derivation for (20).

3.5 Adjunction

3.5.1 GENERAL DISCUSSION

Recall, first of all, our discussion of the structure preserving principle in chapter 6, section 4.1. We have seen that the structure preserving principle does not allow us to destroy existing structure by movement operations, but that it does not exclude that structure be added as long as the resulting representations are compatible with the principles of our grammar. We briefly discussed one example where new structure is generated: free subject inversion.

\[11\] As it stands (17) is non-explanatory. It would, of course, be preferable if it could be derived from some general principle.
In (22b) a new node XP, identified for convenience sake as XP2, is created, dominating the original XP, XP1, and the moved element is attached there. As mentioned briefly in chapter 6, section 4.1, this operation is referred to as adjunction. Adjunction respects our phrase structure theory: the new constituent XP is headed by X. The node XP created by adjunction is binary branching, etc.

Let us go into the relation between XP and ZP more carefully. There are two nodes XP. XP1 is the original maximal projection. It is sometimes called the base maximal projection. XP2 dominates the base maximal projection XP1 and the adjoined ZP. ZP is dominated by the topmost maximal projection XP2, but it is not dominated by the base maximal projection XP1. YP is in contrast is dominated both by XP1 and by XP2. In a way, YP is completely inside the projection of X, YP is included in the projection of X; ZP is only partly inside the projection of X, it is not fully part of the projection of X.

This proposal advocated by Chomsky (1986b: 7) based on May (1985) to define the notion dominance:

23 Dominance
A is dominated by B only if A is dominated by every segment of B.

A is ZP in our example, B is the maximal projection of X. The idea, informally, is that in (22b) the maximal projection of X is the combination of XP1 and XP2. ZP is not dominated by every segment of the maximal projection of X: ZP is dominated by the topmost XP2, but it is not dominated by the base maximal projection XP1.

Even though ZP is not dominated by the maximal projection of X, it is not entirely outside the maximal projection of X, being dominated by the topmost segment XP2. Because ZP is dominated by a segment of XP, we say that ZP is not excluded from XP. Following Chomsky (1986b) exclusion is defined as follows:

24 Exclusion
B excludes A if no segment of B dominates A.

Speaking metaphorically, we could say that a position created by adjunction is like a balcony: when on a balcony you are neither completely outside the
room nor completely inside. You may, for instance, easily participate in conversations going on inside while at the same time get dripping wet if it is raining outside.

A restriction imposed on adjunction by Chomsky (1986b) is that phrases can only be adjoined to maximal projections and that adjunction can only be to non-arguments.

We have gone in some detail into the notion adjunction because it will be relevant also for section 8 below and for subsequent chapters.

3.5.2 WH-MOVEMENT AS ADJUNCTION?

Let us return to multiple movement in the Polish example (20) repeated here as (25a):

25a Maria zastanawia się, [kto co przyniesie].
   Maria wondered who what would-bring
   ‘Maria wondered who would bring what.’

The bracketed indirect question is introduced by two *ub*-phrases co (*what*) and *kto* (*who*). Consider also (26a) (from Lasnik and Saito, 1984:234 (11)).

26a Maria myśli [że co [ip Janek kupi tę]].
   Maria thinks that what Janek bought
   ‘What does Maria think that Janek bought?’

Since *co* follows the complementizer *że* and precedes the subject *Janek* in (26a), it obviously is not in [Spec, CP]. Let us assume that *co* is adjoined to IP. Adjoined positions are A’-positions. The relevant part of the structure of (26a) would be (26b).

\[
\begin{aligned}
\text{Spec} & \rightarrow \text{C} \\
\text{C} & \rightarrow \text{IP} \\
\text{IP} & \rightarrow \text{NP} \\
\text{NP} & \rightarrow \text{IP} \\
\text{IP} & \rightarrow \text{ze co Janek} \\
\end{aligned}
\]

For (25a) we propose that *co* is also adjoined to IP, and that *kto*, which precedes it, has been moved to [Spec, CP]:

\[
\begin{aligned}
\text{Spec} & \rightarrow \text{C} \\
\text{C} & \rightarrow \text{IP} \\
\text{IP} & \rightarrow \text{NP} \\
\text{NP} & \rightarrow \text{IP} \\
\text{IP} & \rightarrow \text{Kto co} \\
\end{aligned}
\]

Note that neither (25b) nor (26b) violates the doubly filled COMP filter: in (26a) the position dominated by C is occupied, but [Spec, CP] is not; in (25b) [Spec, CP] dominates overt material but C does not.

An adjunction analysis of *ub*-movement in the Polish examples (25) and (26) allows us to derive sentences with multiple movement.12 Polish differs crucially from English in that the former, though not the latter, allows for multiple movement:

The Polish data are discussed for the purpose of exemplification. It is quite conceivable that the analysis proposed here is not the optimal analysis. For discussion of *ub*-movement in the Slavic languages the reader is referred to Rudin (1989), Tooman (1981), and Wachowicz (1974).
27a *I wonder what to whom John gave.
27b *I wonder to whom what John gave.

In order to exclude examples of multiple movement at S-structure in English (and similar languages) we assume that in English wh-movement is done by substitution: the moved phrase moves into [Spec, CP].\footnote{In earlier versions of the theory an adjunction analysis had also been adopted for wh-movement in English (see Chomsky, 1980, 1981a) but such an analysis has been abandoned. Lasnik and Saito (1984), which is to a large extent the basis of chapter 9 of this book, still assume an adjunction analysis for English. In chapter 9 their discussion is reinterpreted in terms of substitution.} We could then argue that whether wh-movement operates through adjunction or not is a matter of parametric variation.

3.5.3 A NOTE ON SOME ALTERNATIVE PROPOSALS

Adjunction has played an important role in a number of recent developments of the theory, as we shall also see in chapters 9 and 10. However, more recently proposals have been put forward to constrain adjunction possibilities severely (cf. Kayne, 1993). We do not go into these developments in detail. Let us just return for one moment to the problem raised in (22a), repeated here as (28a). The question we addressed above was where ZP could move to in the structure (28a). Let us assume that KP occupies [Spec, XP].

28a

\[
\begin{array}{c}
\text{XP} \\
\text{KP} \quad \text{X'} \\
\text{X} \quad \text{YP} \\
\text{Spec} \quad \text{Y'} \\
\text{Y} \\
\text{ZP} \\
\text{Spec} \quad \text{Z'} \\
\text{Z}
\end{array}
\]

In (28b) we create an adjunction structure on KP, which would have the same properties as the adjunction structure discussed in section 3.5.1. Notably, ZP would not be dominated by KP, because it is only dominated by one segment of KP, but, on the other hand, KP does not exclude ZP, since one segment of KP dominates ZP. Structures such as (28b) have also been used in the literature.

A final more radical approach would be to propose that when we move ZP it must always move to a specifier position, and that specifier-head relations be unique, i.e. there is one specifier to one head. If such a constraint is imposed on our grammar then the leftward movement of ZP in (28a) would not simply create a segment of a maximal projection to which the constituent can adjoin, but rather to create a full projection, i.e. a maximal projection with its own head. The maximal projection which would have
to be created would not have a lexical head. If there had been a lexical head (say, V, or P) available, then the maximal projection would have been projected at D-structure. Rather the maximal projection we create would have an abstract functional head. (28c) would be such a structure:

![Diagram](image)

We return to the role of functional heads in chapter 11. We do not pursue the problem of the restrictions on adjunction here. Obviously the choice between the different types of structures will have conceptual and empirical implications. (Cf. Kayne 1993 for a very restrictive theory in which adjunction structures are heavily constrained.)

### 3.6 Movement of Maximal Projections: A-movement vs. A'-movement

At this point it may be useful to summarize the discussion of movement so far and to compare the landing site of NP-movement, discussed in chapter 5, and that of wh-movement, discussed in this chapter.

29a Poirot, will be invited t_i.
29b Poirot, seems t_i to be the best detective.

30a Who/Whom, do you think Lord Emsworth will invite t_i?
30b Who, do you think t_i is the best detective?

29 illustrates NP-movement. An NP is moved to the subject position, an A-position. (30) illustrates wh-movement: a wh-element is moved to [Spec, CP], an A'-position. Based on the distinction in landing sites, we say that the chain created by NP-movement is an A-chain and that created by wh-movement is an A'-chain. We distinguish the antecedents of the two types of movement in terms of A-antecedent vs. A'-antecedent. The antecedent of NP-movement c-commands its coindexed trace: we say that it A-binds the trace. In the case of wh-movement the antecedent also c-commands the coindexed trace, but it occupies an A'-position: wh-antecedents A'-bind their traces (cf. section 7). Movement to an A-position is A-movement; movement to an A'-position is A'-movement.

### 4 Traces and Wh-Movement

As was the case for NP-movement, wh-movement leaves a trace coindexed with its A'-antecedent. We refer to traces of wh-movement as wh-traces and to traces of NP-movement as NP-traces. In this section we provide some arguments for postulating wh-traces. Our reasoning in the chapter is closely parallel to that used in chapter 6 to motivate NP-traces.

#### 4.1 Theta Theory and the Projection Principle

One motivation for wh-traces is analogy: having posited that NP-movement leaves traces it seems reasonable to also adopt this proposal for wh-movement. A second argument comes from theta theory. In chapter 6, section 4.2, we argued that the theta criterion applies to all syntactic levels. Internal theta roles are assigned by the lexical head under government. In an S-structure representation such as (31) invite will not be able to theta-mark whom, but it will assign the internal theta role to its trace.

31 Whom, do you t_i believe that Lord Emsworth will invite t_i?

#### 4.2 Agreement and Binding

In chapter 6, we provided some arguments for postulating NP-traces, based on locality constraints on syntactic processes. The same type of arguments can be advanced in favour of coindexed wh-traces:
4.3 Case

4.3.1 WH-PRONOUNS AND CASE

We have seen that abstract case is often not morphologically realized in English. For the *wh*-element what, for instance, there is no overt difference between the NOMINATIVE and the ACCUSATIVE, as the following echo questions demonstrate:

36a. I think that the castle will be destroyed.
36b. You think that WHAT will be destroyed?

37a. I think that Lord Emsworth will sell the pig.
37b. You think that Lord Emsworth will sell WHAT?

The situation with who is different:

38a. I think that Poirot will arrive first.
38b. You think that WHO/*WHOM will arrive first?

39a. I think that Lord Emsworth will invite Poirot.
39b. You think that Lord Emsworth will invite WHO/WHOM?

In (38b) the echo question contains a *wh*-constituent in the lower subject position: only who is admitted, whom is disallowed. In (39b) the *wh*-constituent occupies the object position: both who and whom are possible in spoken English, in writing whom is used. Putting aside many complications with respect to the use of whom here, we assume that the NOMINATIVE case is morphologically realized as who and that the ACCUSATIVE is realized either as who or as whom. Given this assumption let us turn to (40):

34a. Who/*Whom, do you think will arrive first?
34b. Who/*Whom, do you believe that Lord Emsworth will invite?

In (34b) the moved *wh*-phrase is NOMINATIVE; in (40b) it is ACCUSATIVE. In English NOMINATIVE and ACCUSATIVE are assigned at Spec, and under government (see chapter 3 for discussion). Neither in (34a) nor in (40b) do the case assigners, the finite I and the transitive verb
invoke respectively, govern the moved wh-phrase. But the traces of who and whom are governed by the relevant case assigners. We adopt the idea that the trace of wh-movement is case-marked. The case on the trace will make the theta position visible and will allow the verb to assign its theta role.

4.3.2 WH-TRACE VS. NP-TRACE: MORE CONTRASTS

Let us briefly compare wh-traces and NP-traces with respect to the assignment of case.

41a Poirot will be invited t_i.
41b Poirot, seems t_i to be the best detective.

42a Who/Whom, do you think Lord Emsworth will invite t_i?
42b Who, do you think t_i is the best detective?

NP-traces are not assigned case. In (41a) the passive verb fails to case-mark its complement NP and in (41b) unaccusative seem is unable to case-mark the subject of the lower non-finite clause. The wh-traces in (42) are case-marked: the verb invite in (42a) assigns ACCUSATIVE case to its complement; definite I in (42b) assigns NOMINATIVE to the wh-trace in the subject position. The situation of the antecedent of the trace is reversed: in the case of NP-movement the antecedent is assigned case. In the case of wh-movement the antecedent is not in a position to which case is assigned.

(43) summarizes the comparison between NP-movement which is A-movement and wh-movement, which in A’-movement.

43a A-movement; NP-movement
   A-chain
   Antecedent (head of the chain): +Case
   Trace (foot of the chain): −Case
43b A’-movement; wh-movement
   A’-chain
   Antecedent: −Case
   Trace: +Case

By simply inspecting the head or the foot of a chain we can identify the type of movement and its properties. Chapter 8 offers a more detailed comparison between the two kinds of movement.

4.4 Adjunct Traces

So far we have only discussed wh-traces of arguments. Adjunct phrases are also subject to wh-movement.

44 When did you tell her that Bill is coming?

In the sentence-initial position of (44) we have the temporal adjunct when. When can be related to, or construed with, the activity expressed in the matrix clause, i.e. ‘telling’, or with that in the subordinate clause, ‘coming’. The trace of the moved phrase will indicate which clause the time adjunct modifies. (45b) and (45d) suggest the type of answer for each interpretation:

45a When, did you tell her t_i [that Bill is coming]?
45b I told her yesterday that Bill is coming.
45c When, did you tell her [that Bill is coming t_i]?
45d I told her that Bill is coming tomorrow.

5 Subject Movement

In this section we turn to two problems related to wh-movement from subject position. Both phenomena will be described here and will be discussed in more detail in chapters 8, 9, 10 and 12.

5.1 Vacuous Movement

Let us look at sentences in which the subject is questioned.

46a Who, do you think [t_i will arrive first]?
46b Who will arrive first?

In (46a), an example of long subject movement, who has been extracted from the subject position of the lower clause, leaving a coindexed trace. In (46b) matters are not so obvious. In the literature two contrasting proposals have been put forward: it is usually assumed that, by analogy with object movement and long subject movement, the subject wh-phrase also moves in...
b Qui crois-tu qui viendra?
  who think-you 'who' will-come
  'Who do you think will come?'

In Dutch there is a lot of dialectal variation as indicated by the diacritic %; some dialects allow the sequence dat-trace, others do not.

54 Dutch
  %Wie denk je dat dat boek gekocht heeft?
  who think you that that book bought has
  'Who do you think has bought that book?'

Leaving aside the cross-linguistic variation and returning to English, we see that the that-trace filter in (51) can also account for the judgements in (55).

55a I would prefer Bill to come first.
55b Who would you prefer to come first?
55c I would prefer for Bill to come first.
55d *Who would you prefer to come first?

The ungrammaticality of (55d) can be related to the ban on a sequence complementizer-trace:

55e *Who, would you prefer [CP for [IP t, to come first]]?

6 Bounding Theory

6.1 Island Constraints: The Data

Consider (56), an example of wh-movement:

56 [CP How did [IP you [VP say [CP that [IP Jeeves thinks [CP that [IP Lord Emmerich will solve this problem t,]]]]]]].

The wh-phrase how is moved from inside the lowest clause and ends up in the matrix [Spec, CP]. Research initiated by Ross in the 1960s (Ross, 1967) has shown that wh-movement is not unconstrained. In this section we shall consider the central data. In later sections we shall provide an analysis of the data.

6.1.1 THE COMPLEX NP CONSTRAINT

Consider the data in (57). (57a) and (57b) are grammatical, (57c), which is closely similar in structure and in interpretation to (57b), is not.

57a [CP Who, did [IP he see t, last week]]?
57b [CP Who, did [IP Poirot claim [CP that [IP he saw t, last week]]]]?
57c *[CP Who, did [IP Poirot make [NP the claim [CP that [IP he saw t, last week]]]]]

In the ungrammatical (57c) the wh-phrase, who, is extracted from inside a complex NP, an NP whose head N (claim) takes a sentential complement. Ross proposes that movement out of a complex NP is blocked. Complex NPs are islands for movement. The constraint which bans movement out of a complex NP is often referred to as the complex NP constraint, abbreviated as CNPC.

6.1.2 WH-ISLANDS

Now consider the data in (58):

58a [CP1 How, do [IP you [VP think [CP2 that [IP John could [VP solve this problem t,]]]]]]?
58b I wonder [CP which problem, [IP John could solve t, this way]].
58c *[CP1 How, do [IP you [VP wonder [CP2 which problem, [CP2 which problem, [CP1 John could [VP solve t, t,]]]]]]]]?
58d *[CP1 Which problem, do [IP you [VP wonder [CP2 how, [CP2 John could [VP solve t, t,]]]]]]?

In (58a) we extract how from the lower clause and move it to the matrix [Spec, CP1] analogously to the movement of how in (56) or who in (57b). This movement is unproblematic. Equally unproblematic is the movement of which problem in (58b): which problem is extracted from the VP-internal position and moves to the embedded [Spec, CP]. In (58c) we combine the two types of movement: how moves to the matrix [Spec, CP1] and which problem moves to the lower [Spec, CP2]. Though each of these movements is in itself
Let us first consider the application of the subjacency condition in a grammatical sentence. Consider (56), repeated here as (60a):

60a  \[ cp_1 \text{How, did [IP] you say [cp_2 that [ip_2] Jeeyes thinks [cp_3 that [ip_3 Lord Emsworth will solve this problem t_1]]]}? \]

At first sight the reader might be tempted conclude that the subjacency condition (59) is violated in (60a): how is extracted from the lower CP and ends up in the higher [Spec, CP1], crossing, it would appear (i) the lower IP3, (ii) the intermediate IP2, and (iii) the matrix IP1. Condition (59) rules out this derivation. However, observe that there is an alternative representation possible: we do not have to move how in one single step from the lowest clause to the matrix domain; the wh-element could also move stepwise: first it moves to the lowest [Spec, CP3], then it moves to the intermediate [Spec, CP2] and finally it moves to the matrix [Spec, CP1]. We assume that at each intermediate step the movement leaves a trace. We shall refer to the traces in between the foot of the chain and its head as intermediate traces.

60b  \[ cp_1 \text{How, did [IP] you say [cp_2 [t_1] that [ip_2] Jeeyes thinks [cp_3 [t_1] that [ip_3 Lord Emsworth will solve this problem t_1]]]]? \]

Each of the intermediate steps is an application of wh-movement. Each clause (CP) defines a domain of application for wh-movement, a syntactic domain in which wh-movement can apply, or a cycle. We say that the movement is successive cyclic: it applies in successive cycles, from bottom to top. We have seen in chapter 6 that NP movement is also cyclic and also leaves intermediate traces. In general, it is assumed that transformations are subject to cyclicity: all transformations that are restricted in application to a lower cycle will apply prior to those that involve higher cycles. As the discussion shows, a vacat [Spec, CP] serves as a sort of passway for movement: thanks to the availability of this vacant position, movement out of the lower clause is possible. [Spec, CP] is like an escape hatch. In section 6.2.3 we shall see that when the intermediate [Spec, CP] is filled this gives rise to subjacency violations, as movement would then have to cross two consecutive IPs.
6.2.2 THE COMPLEX NP CONSTRAINT

Let us first turn to the data introduced in section 6.1.1. We have seen that extraction out of a complex NP leads to an ungrammatical result; (61a) violates Ross’ complex NP constraint:

61a  *[CP1 What, did [IP2 Poirot make [NP the claim [CP2 that [IP3 he saw t1 last week]]]]]

The example violates the subjacency condition. Consider the representation of (61b) where we have made maximal use of landing sites available for movement:

61b  *[CP1 What, did [IP1 Poirot make [NP the claim [CP1 t1 that [IP2 he saw t1 last week]]]]]

Step 1 is legitimate: only one bounding node is crossed, IP2. Step 2 is illegitimate: two bounding nodes, NP and IP1, are crossed. Violations of the complex NP constraint are violations of subjacency.

In our discussion of (61b) we represent each step of the derivation of the sentence by means of an arrow. This is done for expository reasons, we do not need to rely on the derivational history of this example in order to check the subjacency condition. The S-structure representation of the sentence in itself records effects of movement: traces indicate the vacated positions and includes both the D-structure position of the moved phrase, and the intermediate landing sites. By inspecting the distance between the tracings terms of bounding nodes we can detect subjacency violations.

6.2.3 WH-ISLANDS

Now let us turn to the examples of extraction out of wh-islands, illustrated in (58c) and (58d). We repeat them here as (62a) and (62b).

62a  *[CP1 How, do [IP1 you [VP wonder [CP2 which problem C0 [IP2 John could [VP solve t1 t1]]]]]]

62b  *[CP1 Which problem, do [IP1 you [VP wonder [CP2 how, C0 [IP2 John could [VP solve t1 t1]]]]]

Now which problem would have to move to the intermediate [Spec, CP2]. This movement raises several problems. In (62a) the movement of how involves the lower cycle (CP2) and the higher cycle (CP1) which contains the landing site. The subsequent movement of which problem would be restricted to the lower cycle, CP2. This means that the movement of which problem to the intermediate [Spec, CP2] violates the cyclicity condition. Second, consider the intended movement of which problem. It has to be moved to the intermediate [Spec, CP2]. If we allow this movement to take place, we will have to obliterate the intermediate trace of how in [Spec, CP2] and the chain between the antecedent how and the trace in the base position will be destroyed. Again the S-structure representation of (62a) will also register the problem: the antecedent how in [Spec, CP1] will not be locally related to its trace, since the wh-phrase in [Spec, CP2], which problem, will interfere and prevent the formation of the chain <how, t1>.

The problem which we return to in later chapters is posed by the contrast between (62a) and (62b). Though (62b) is not fully acceptable, it is not as degraded as (62a). Anticipating the later discussion, consider that both sentences violate subjacency. In (62a) we extract an adjunct, how, from a wh-island; in (62b) we extract an argument, the direct object which problem. The subjacency violations would account for the degradation in both sentences. We return to the adjunct-argument asymmetry in chapters 9, 10 and 12, where we shall see that (62a) violates an additional principle of the grammar.
6.3 Subjacency as a Diagnostic for Movement

Wh-movement is subject to the subjacency condition. Whenever the possible linking of an empty position and its antecedent can be seen to be subject to this condition we can conclude that *wh*-movement is involved (cf. Riemsdijk, 1978b). In this section we consider two constructions for which a movement analysis has been proposed: left dislocation (6.3.1) and relative clause formation (6.3.2 and 6.3.3). Using the subjacency condition as a diagnostic we show that a movement analysis is not appropriate for left dislocation but that relative clauses are derived by means of *wh*-movement. In section 6.3.4 we see that NP-movement is also subject to the subjacency condition.

6.3.1 LEFT DISLOCATION; MOVEMENT AND COPYING?

(63) illustrates left dislocation:

63a Simonon, I don't like him.
63b Simonon, I always wonder when I discovered him.

In (63a) the NP *Simonon* is in a sentence-initial position. Let us assume it is adjoined to IP (for adjunction see section 3.4). The pronoun *him* is coindexed with the NP *Simonon*. One might propose that the NP *Simonon* is the subject object of *like* and that it has been moved to the sentence-initial position. The pronoun *him* would then be interpreted as a pronominal complement inserted at the vacated site.

64a [in *Simonon, [np I don't like him]]]

A closer look at the examples shows that a movement analysis is not appropriate. The distance between *Simonon, and him* is not subject to the subjacency condition: in (64b) two bounding nodes (IPs) intervene, indicated by #:

64b Simonon, [np I always wonder [cp when [np I discovered him]]]]

This leads us to the conclusion that left dislocation is not the result of movement. The sentence-initial NP *Simonon* has not been moved from inside IP. We assume that the NP is present in the adjoined position at D-structure.

This conclusion is important. So far we have suggested that adjunction structures are created by *wh*-movement. The examples above lead us to the conclusion that adjunction structures can be base-generated, i.e. that they also occur at D-structure. At this point then, we must reconsider the discussion of phrase structure in chapter 2 and include adjunction structures. The *X*-schema has to be completed with the phrase structure rule in (65a), where the semicolon means that order is irrelevant, allowing both right adjunction (65b) and left adjunction (65c):

65a XP ———> XP; YP
65b XP ———> XP—YP
65c XP ———> YP—XP

The *X*-format can then be summarized as in (66):

66a XP ———> XP; YP
66b XP ———> Spec; X'
66c X' ———> X; YP
66d X' ———> X; YP

6.3.2 RELATIVE CLAUSES AND WH-MOVEMENT

Consider the following example, with partial bracketing:

67 I know [np the man [cp whom [np Emsworth will invite]].]

(67) contains a complex NP with a relative clause: the head noun *man* is modified by a clause (CP). We focus on the internal structure of the relative clause here. Based on the presence of a tensed auxiliary (*will*) and a subject NP, we propose that the relative clause is an IP preceded by the relative pronoun *whom*. This is strikingly similar to the structure of indirect questions. Let us assume that relative clauses are sentences, i.e. CPs, and that the relative pronoun *whom* occupies [Spec, CP]. Note that as a maximal projection, *whom* could only occupy [Spec, CP], C being reserved for heads. The predicate of the relative clause, the verb *invite*, needs an internal argument. There is no overt element present but by analogy with our analysis of *wh*-questions we propose that *invite* is followed by a trace whose antecedent is *whom*. The complete S-structure of the relative clause in (67) is (68a) and its D-structure (68b):

---

(68)
68a [\textsc{cp} Whom, \textsc{ip} Emsworth will [\textsc{vp} invite t]]])?
68b [\textsc{cp} \textsc{ip} Emsworth will [\textsc{vp} invite whom]]?)

The hypothesis is that relative clause formation involves \textsc{wh}-movement. If this is correct then the relative clause construction should be subject to an adjacency condition on movement (59).

69a This is the man whom Emsworth claims that he will invite.
69b *This is the man whom Emsworth made the claim that he will invite.

In (69a) long movement is allowed in relative clauses. (69b) is a violation of the complex NP constraint. The S-structure representations of the relevant NP in the sentences in (69) are given below.

70a [\textsc{np} the man [\textsc{cp} whom, \textsc{ip} Emsworth claims [\textsc{cp} t', that \textsc{ip} he will invite t]]]])

The violation of adjacency in (69b) will be signalled by means of the ditritic # on the brackets.

70b *\textsc{np} the man [whom, \textsc{ip} Emsworth made \textsc{np} the claim [t', that # #\textsc{ip} he will invite t]]]])

Using the adjacency condition as a diagnostic we find confirmation that relative clause formation is indeed a result of movement. In chapter 10 we reformulate the adjacency condition in terms of the notion barrier.

As it stands the S-structure representations in (70) are not sufficient to allow us to interpret the relative clause. (70a), for example, does not indicate that \textit{whom}, is to be linked to \textit{the man}. We assume that the interpretation of the relative pronoun is achieved through a rule of coindexation where \textit{the man} and \textit{whom} end up having the same index. This coindexation is used to represent the fact that the relative clause modifies or is 'predicated of' \textit{the man} as a \textbf{predication rule}.

\footnotesize 17 For a discussion of predication the reader is referred to Williams (1980). It is proposed in the literature that the co-indexation rule does not apply at S-structure but at the level of logical form, LF, which is discussed in chapter 9. Further discussion of the predication rule is found in Chomsky (1982: 92–3) and Safrir (1986).

6.3.3 RELATIVE CLAUSES AND RESUMPTIVE PRONOUNS

In the section above we have seen that relative clauses in English are derived by means of \textsc{wh}-movement. However, relative clauses need not be formed by means of movement. Consider the following examples from French, taken from Zribi-Hertz (1984).

71a Voici l’homme, à qui Marie a parlé t.
here-is the man to whom Marie has talked
71b Voici l’homme, que Marie lui a parlé.
here-is the man that Marie to-him has talked

72a Voici la maison, à laquelle Marie pense encore t.
here-is the house about which Marie thinks still
72b Voici la maison, que Marie y pense encore.
here-is the house that Marie of-it thinks still

73a Voici le courrier, qui t, est arrivé ce soir.
here-is the mail which is arrived tonight
73b Voici le courrier, qu’il est arrivé ce soir.
here-is the mail that-it is arrived tonight

The (a) examples above illustrate ‘standard’ French, the (b) examples illustrate ‘popular’ French. The (a) examples are straightforward illustrations of \textsc{wh}-movement: in (71a), for instance, the PP \textit{à qui} is moved to [Spec, CP] and finds a trace in its extraction site. The (b) example illustrates an alternative strategy for the formation of relative clauses. The \textit{relative} clause is headed by the complementizer \textit{que} and it contains a pronoun \textit{lui} which is coindexed (by the rule of predication) with the relativized NP \textit{l’homme}. The pronoun which is related to the relativized NP is called a \textbf{resumptive pronoun}.

English too has a substandard resumptive pronoun strategy for the formation of relative clauses. Zribi-Hertz (1984: 27) gives the following example (from Chomsky, 1982: 11, his (8b)).

74 the man who, John saw him,

74 differs from the French examples in that the resumptive pronoun \textit{him} is associated with a \textsc{wh}-element in [Spec, CP]. Given that the pronoun occupies its base-position, we must conclude that the \textsc{wh}-element must be
base-generated in [Spec, CP], i.e. it does not move to that position. As an important consequence of this analysis is that because no \emph{wh}-movement is involved, the subjacency condition should not come into play:

75 the man who, [\emph{wh} they think [\emph{cp} that [\emph{cp} when [\emph{cp} Mary marries him]]] then everyone will be happy]]

The resumptive pronoun \emph{him} is inside a clause introduced by \emph{when}. As (76) shows, such clauses are \emph{wh}-islands; we assume \emph{when} is in [Spec, CP]:

76 *the man who they think that [\emph{cp} when [\emph{cp} Mary marries]] then everyone will be happy

For many speakers of English, the resumptive pronoun strategy is a way of overcoming subjacency violations. For French we shall assume that the resumptive pronoun is related to a non-overt element in [Spec, CP]. Such non-overt elements will be discussed in Chapter 4.

77 I am looking for those documents which I can never remember where I put them.

6.3.4 NP-MOVEMENT

If the subjacency condition is a constraint on movement, then we expect it will also apply to NP-movement. In (78) the subjacency condition is respected: in (78a) no bounding nodes are crossed; in (78b) one bounding node is crossed.

78a John, was invited \emph{t} at Mary’s house.
78b John, seems [\emph{cp} \emph{t} to have lost].

Now consider the ungrammatical (79a):

79a [\emph{cp} \emph{t} John, seems [\emph{cp} that [\emph{cp} it is likely [\emph{cp} \emph{t} to [\emph{vp} \emph{t} win]]]]].

\emph{John} originates as the subject of the lowest clause, IP3, and moves to [Spec, IP1] in the matrix domain. For completeness’ sake we add the trace in (79b):

79b *\emph{wh} John, seems [\emph{cp} that [\emph{cp} it is believed \emph{t} by everyone]]].

The movement of \emph{John} in (79b) does not violate subjacency and the sentence remains ungrammatical. In Chapter 6, section 4.5.2. we saw that (79b) is ungrammatical because the trace of \emph{John} is not bound in its governing category, IP2. In Chapter 6 we argued that NP-traces are subject to Principle A of the binding theory. In (79a): the governing category of \emph{t}, is IP2: IP2 contains the trace itself, its governor \emph{likely}, and a subject, the NP \emph{it} in [Spec, IP2]; \emph{t} is not bound in its governing category. According to our definition of chains (Chapter 6 (105a)) the chain <\emph{John}, \emph{t}> cannot be formed in (79a): \emph{John} does not locally bind \emph{t}.

6.4 The Subjacency Parameter

Consider the following Italian NPs (Rizzi, 1982b: 50):

80a tuo fratello, a cui mi domando che storie your brother, to whom myself I ask which stories they have told
80b il solo incarico che non sapevi a chi the only charge that not you knew to whom they would have entrusted
80c la nuova idea di Giorgio, di cui immagino che cosa pensi the new idea of Giorgio, of which I imagine what you think
80d Giorgio’s new idea, of which I imagine what you think
The English equivalents of (80) are far less acceptable:

81a  *your brother, to whom I wonder which stories they told
81b  *the only task which you ignore to whom they'd entrust
81c  *George's new idea, of which I can imagine what you think

The reader will probably be able to identify the English examples in (81) as violations of the subjacency condition. In (81a), for instance, to whom has been extracted out of an embedded question introduced by which stories, crossing IP₂ and IP₁:

82  your brother [CP to whom [IP I wonder [CP which stories [IP₂ they told [t₁ ⋄ ⋄]]]]]

Apparently this type of extraction is allowed in Italian: omitting irrelevant details the S-structure of (80a) is (83). A cui has crossed IP₂ and IP₁, without any harm.

83  tuo fratello [CP a cui [IP mi domando [CP che storie [IP₂ abbiano raccontato [t₁ t₂ ⋄ ⋄]]]]]

One possibility would be to claim that the subjacency condition is language-specific, like the doubly filled COMP filter, and does not apply in Italian. If this were true one would equally expect that extraction out of any type of indirect question and out of complex NPs is freely possible, contrary to fact (example from Rizzi, 1982b: 51):

84  * tuo fratello [CP a cui [IP temo [NP la possibilità [CP t₁ your brother to whom I fear the possibility che [IP₂ abbiano raccontato [t₁ tutto]]]]] that they have told everything

Rizzi's proposal to account for the example given here is NOT that subjacency is irrelevant for Italian. Rather he proposes that the bounding nodes are parametrized, i.e. that different languages may have different bounding nodes. While we assume that in English NP and IP are the relevant bounding nodes, for Italian bounding nodes would be NP and CP. On the basis of this proposal the grammaticality of (80) follows: in each of the

85  *tuo fratello [CP a cui [IP mi temo [NP la possibilità [CP t₁ che [IP₂ abbiano raccontato [t₁ tutto]]]]]

In (85) the PP a cui (to whom) originates in the lowest clause: it is the complement of raccontare (tell). It is first moved to the lowest [Spec, CP], and then it has to move across CP and across NP, crossing two bounding nodes.

The subjacency parameter is one of the earliest formulated in the present theory.²⁰

7  Binding Theory and Traces of Wh-Movement

7.1  Typology of NPs

In the discussion of wh-movement it has become clear that some of the wh-traces (86) have the status of NPs:

86a  Whom, will Lord Emsworth invite t₁?
86b  Which detectives, do you expect [t₁ to admire themselves most]?

The trace in (86a) occupies a position normally taken by an NP, it is case-marked by the verb and it is assigned a theta role. In (86b) the trace binds a reflexive with which it shares features of person, number and gender. If these wh-traces are NPs, the question is how they behave with respect to the BT. Or to put it differently: what type of NPs are those traces? In chapter 4 we identified four NP types based on the features [± Anaphor] and [± Pronominal]:

²⁰ For a discussion of the subjacency parameter in French the reader is referred to Spottiswoode (1981). Further discussion of the Italian data is found in Rizzi (1982b). For various modifications of the parameter see also Chomsky (1986b) and chapters 9 and 10.

²¹ Clearly, traces of PPs, for example, will not have the status of NPs but rather that of PPs.
Typology of NPs

<table>
<thead>
<tr>
<th>Type</th>
<th>OVERT</th>
<th>NON-OVERT</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+Anaphor, –Pronominal]</td>
<td>anaphors</td>
<td>NP-trace</td>
</tr>
<tr>
<td>[–Anaphor, +Pronominal]</td>
<td>pronouns</td>
<td>?</td>
</tr>
<tr>
<td>[–Anaphor, –Pronominal]</td>
<td>R-expressions</td>
<td>?</td>
</tr>
<tr>
<td>[+Anaphor, +Pronominal]</td>
<td>PRO (chapter 5)</td>
<td></td>
</tr>
</tbody>
</table>

We have assimilated NP-traces with anaphors. Could we do the same for wh-traces? At first glance one might wish to say yes. After all, wh-traces need a c-commanding antecedent. But a more careful analysis shows that the answer is ‘No’. The moved wh-constituent is coindexed with its trace and c-commands it. Since the wh-constituent is in an A’-position, it does not A-bind its trace. We have said (section 3.5) that the moved wh-constituent A-binds its trace. The binding theory developed in chapter 4 is about A-binding, and not about A’. The reader can for himself that the wh-trace is not A-bound by anything in its GC in the examples above.

We do not dwell too long on the question whether the wh-trace is like a null element PRO, discussed in chapter 5. It must be clear that the wh-trace is governed and PRO must not be governed.

Is the wh-trace then like a pronoun? Principle B of the binding theory says that pronominal elements must be free in their GC. In other words a pronoun may be bound by something outside the GC. If wh-traces were like pronouns they should have the same distribution. Let us try to construct an example.

88a The detective thinks that [IP he, likes Bill best].
88b *Who does the detective think [IP t, [IP t, likes Bill best]]?

In (88a) the pronoun he is allowed to be coreferential with the NP the detective, since the latter is outside its GC. In (88b) the lowest trace of who occupies the position filled by the pronoun he in (88a). But in (88b) the natural answer is not that ‘the detective thinks that he himself likes Bill best’, i.e., that the NP the detective can be coindexed with who and consequently the trace of who; t, and the NP the detective must not be coreferential in (88b). If the trace of wh-movement were like a pronoun then the facts would rather hard to explain.

Last but not least we turn to the final option: what if wh-traces were like R-expressions? Following Principle C of the binding theory they would have to be free everywhere. A brief look at the data above confirms that this is indeed the right answer. The fact that the trace of who in (88b) cannot be bound by the detective follows directly. The example is structurally parallel to (88c).

36c *He, thinks [IP John, likes him best].

We can now identify one more null element in the table above: wh-traces are like R-expressions:

Typology of NPs

<table>
<thead>
<tr>
<th>Type</th>
<th>OVERT</th>
<th>NON-OVERT</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Anaphor, –Pronominal]</td>
<td>anaphors</td>
<td>NP-trace</td>
</tr>
<tr>
<td>Anaphor, +Pronominal</td>
<td>pronouns</td>
<td>?</td>
</tr>
<tr>
<td>Anaphor, –Pronominal</td>
<td>R-expressions</td>
<td>wh-trace</td>
</tr>
<tr>
<td>Anaphor, +Pronominal</td>
<td>PRO</td>
<td></td>
</tr>
</tbody>
</table>

In the discussion we have distinguished the concepts A’-binding from A-binding, and A’-bound from A-bound. If we wish to refer to ‘any’ binding we will use the terms X-binding or X-bound. Traces must be X-bound.

By way of summary, let us look at some examples of movement in order to see how the binding theory applies.

34a Who do you think is the best detective?
34b *Poirot seems is the best detective.

We invite the reader to provide the S-structure and the D-structure for these examples before continuing to read. The sentences in (90) have S-structures (91a) and (91b) respectively:

1a Who do [IP you think [IP t, [IP t, is believed [IP t to be the best detective]]]]?
1b *Poirot, seems [IP t, [IP t, is the best detective]].

1a contains a combination of NP-movement and wh-movement. Who, introduces as the subject of the lower infinitival clause. Being caseless — believed is passive — who moves to the subject position of the higher clause.
where it is assigned NOMINATIVE case. (In passing we draw the reader's attention to the fact that it is quite possible, as in (91a), that a *wh-constituent undergoes NP-movement.) From this subject position *who* is then *wh*-moved to the matrix [Spec, CP], via the intermediate [Spec, CP]. The trace in the subject position of the infinitival clause has all the properties of an NP-trace: it is caseless, it is A-bound and like anaphors it is bound in its governing category. The trace in the subject position of *is believed* is a trace of *wh*-movement: it has case, it is A'-bound and so, like R-expressions, it is not A-bound.

Let us turn to the ungrammatical (90b), whose S-structure is given in (91b). The idea is here that the NP Poirot moves from the lower subject position of a finite clause to the subject position of the matrix clause via the intervening specifier of CP where it leaves an intermediate trace. The lower trace in the subject position of *is* has case and thus to all appearances is a trace of *wh*-movement. The lower trace is A'-bound from the intermediate [Spec, CP]: it is the foot of an A'-chain. The analysis implies that NPs can undergo *wh*-movement, a possibility independently allowed as we discuss in section 8.1. Admitting this possibility for the moment without further discussion, the representation (91b) is still problematic: the lower trace, being like an R-expression, must be free. In the example it is bound by the NP Poirot, in an A-position, the subject position of *seems*. (91b) violates the binding theory. Movement which goes from an A-position to an A'-position and back to an A-position is often referred to as improper movement.22

Let us briefly consider another derivation for (90b): take the S-structure representation (92), where we assume that the NP Poirot, moves directly from the lowest subject position to the matrix position, leaving a coindexed trace.

92  *[Poirot, seems [CP [P t, is the best detective]]].

The trace in (92) is assigned NOMINATIVE by the finite INFL of *is*, so we conclude that it is a *wh*-trace. There are two problems with this derivation. First, a *wh*-trace should have an antecedent in an A'-position, which is not the case in (92), and second, it should not be A-bound, which it is. On the other hand, suppose we were boldly to ignore the case diagnostic and assume *t* is an NP-trace. We only do this to our detriment. An NP-trace is subject to Principle A of the BT: it must be bound in its GC. The GC is the lower finite clause (see chapter 4 for the definition of the GC and for the role of AGR in particular) and clearly the trace is not bound there. Whatever syntactic representation we imagine for (90b) it will violate some principle of our grammar. (90b) has no legitimate syntactic representation and is ungrammatical.

7.2 Crossover

In the literature23 the following examples have received a lot of attention (cf. (88b)):

93a *Who, does he think t left?
93b *Who, does he think you saw t?
93c *Who, does he see t?

The ungrammaticality was at one time attributed to the fact that *wh*-movement moves a constituent across a coindexed pronoun. It was proposed that these examples are ruled out by the so-called leftness condition, re-formulated here for expository reasons (Koopman and Sportiche, 1982: 140) and illustrated in (94b):

94a Leftness condition

A *wh*-trace cannot be coindexed with a pronoun to its left.

94b *Who, does he think t left?

The ungrammatical examples in (93) are usually described as illustrating strong crossover (SCO).24 It is clear that such examples can also be explained in terms of the discussion in section 7.1: in all instances the *wh*-trace will be A-bound.

Now consider the contrast in (95). An example such as (95b) is referred to as weak crossover (WCO). The term is chosen because the ungrammaticality is less strongly felt than that illustrated in (93).

95a Who, loves his, mother?
95b *Who, does his, mother love t?
The contrast in grammaticality between the sentences in (95) can also be explained by the leftness condition (cf. Koopman and Sportiche, 1982: 140). The contrast in (95), however, does not follow from the BT. In (95b) the trace is not $A$-bound; the pronoun $bis$ does not $c$-command, hence does not bind, the $wh$-trace. We conclude that SCO follows from the binding theory, but WCO does not. One proposal is to maintain the leftness condition. This will rule out both the examples in (93) and (95b). However, there is some redundancy in our theory since (93) is ruled out both by the binding theory and by the leftness condition. The leftness condition and the BT do the same job in (93), this suggests that one of the two is superfluous. Restricting the leftness condition for WCO seems an ad hoc solution.\footnote{For an alternative analysis of weak crossover, the reader is referred to Koopman and Sportiche (1982). Further discussion of crossover and relative clauses is found in Safr (1986). See also Lasnik and Stowell (1991).}

8 Movement to the Right in English

So far we have discussed only leftward movement of constituents in English. NP-movement takes a constituent to a $c$-commanding [Spec, IP] and $wh$-movement takes an element to a $c$-commanding [Spec, CP]. In this section we illustrate two instances of rightward movement, known as heavy NP-shift and PP-extrapolation from NP. We shall see that these two are instances of $wh$-movement. Our discussion will entail that the term $wh$-movement is to be interpreted as movement to an $A$-position.

8.1 Heavy NP-shift

In chapter 3 we discussed case assignment in English. On the basis of examples like those in (96) we postulated an adjacency constraint on ACCUSATIVE assignment.

96a Poirot speaks English badly.
96b *Poirot speaks badly English.
96c Bertie drinks whisky every night.
96d *Bertie drinks every night whisky.
96e Jeeves introduced him to the guests.
96f *Jeeves introduced to the guests him.

Certain examples seem to provide counterevidence for the adjacency requirement on case assignment in English:

97a Jeeves introduced to the guests [\text{NP} the famous detective from Belgium].
97b My doctor told me to drink every night [\text{NP} two glasses of mineral water with a slice of lemon].

The bracketed NPs in (97) are internal arguments of the verbs \textit{introduce} and \textit{drink} respectively; they are directly theta-marked. Theta theory specifies that direct theta-marking is achieved under government. At S-structure the relevant NPs should be made visible by case. In order to account for the thematic relations between the verbs and their complements we assume that the sentence-final positions of the NPs in (97) are derived positions and that the D-structure of these sentences is as in (98):

98a Jeeves [\text{VP} introduced \text{NP} the famous detective from Belgium] to the guests.
98b My doctor told me to [\text{VP} drink \text{NP} two glasses of mineral water with a slice of lemon] every night.

The S-structure of these examples will be (99):

99a Jeeves [\text{VP} [\text{VP} introduced \text{NP} to the guests \text{NP} the famous detective from Belgium]].
99b My doctor told me to [\text{VP} [\text{VP} drank \text{NP}] every night] [\text{NP} two glasses of mineral water with a slice of lemon]].

The question that arises is how to characterize the rightward movement of the object NP: is it $A$-movement or is it $A$'-movement? Recall that $A$-movement is movement to an $A$-position, and is exemplified by NP-movement in raising and passive sentences; $A'$-movement is movement to an $A'$-position and is exemplified by $wh$-movement in interrogative and relative clauses.

As is standard by now, we assume that movement leaves traces. At first sight it might appear that the rightward movement of the NP in (99) is NP-movement, i.e. $A$-movement, and leaves an NP-trace: after all, the moved constituent is an NP. This would mean that the chain created by the movement of the NP in (99) is an $A$-chain, i.e. a chain whose head is an $A$-position. Two problems arise. First, the traces which constitute the foot of the chain in (99) are case-marked by the verb. NP traces are ordinarily not
assigned case. In addition, if the movement is A-movement, then the landing site of the movement must be an A-position. A-positions are positions in which arguments occur, they are positions which are assigned grammatical functions, i.e. the object position and the subject position. On the basis of this definition it is hard to see how the landing site of the rightwardly moved NP could be an A-position. Another property of A-movement as instantiated by NP-movement is that it substitutes an NP for an empty position. NP movement in passive sentences, for instance, moves an object NP into a subject position, raising moves a subject NP of a non-finite clause into the subject position of a higher clause. In both cases we assume that the subject position [Spec, IP] is generated in the base and then filled by the moved constituent. But it is hard to see how to motivate a base-generated sentence-final position which could become the landing site for the rightward movement of the NP in (99). There is no motivation to postulate a sentence-final position which can host the moved NP, whether this position be categorically specified as an NP position, or categorially unspecified. We conclude that the NP is moved to a position created for it. In other words the moved NP is adjoined. (See the discussion of adjunction in connection with structure preservation in chapter 6; and also section 3.4 above.) Let us assume that the moved NP is adjoined to VP producing a structure like in (100):

\[ \text{VP} \]

\[ \text{VP} \]

\[ \text{V'} \]

\[ \text{NP} \]

\[ \text{V} \]

\[ \text{t} \]

\[ \text{every} \]

\[ \text{night} \]

\[ \text{two glasses of mineral water with a slice of lemon} \]

The moved NP, c-commands its trace, a desirable result since we have seen that both NP-traces and wh-traces are c-commanded by their antecedents.

Obviously, the adjunction analysis is incompatible with the hypothesis that the chain \(<\text{NP}_1, t>\) is an A-chain. The adjoined position is an A'-position. We conclude that the chain created by the moved NP in our example is like the chain created by wh-movement. The trace is A'-bound. That the trace is assigned case is as expected.

If the empty category in the object position in (100) is indeed a trace created by movement — our hypothesis — then the link with its antecedent should be subject to the subjacency condition on movement. This is easy to check. In (101) we see that rightward movement of NPs must not cross more than one bounding node: NP, crosses its containing IP and in addition an NP-node and the result is ungrammatical:

101 *\[ \text{The man \{CP who, \{t does t, every night\}\} bothers me \{NP, two glasses of mineral water with a slice of lemon\}.} \]

We may wonder why the object NPs in (96) cannot be adjoined to the VP. What distinguishes the grammatical examples from the ungrammatical ones is that the moved NP in the grammatical examples is rather heavy. Apparently the adjunction of object NPs to the VP is only admitted with heavy NPs. A precise definition of the concept of heaviness has not been formulated but the intuitive idea is clear. Rightward movement of NPs as exemplified in (97) is called heavy NP-shift.26

26 Not every reordering of complements is necessarily due to A'-movement of the object. For some examples the reader is referred to Belletti and Shlonsk (forthcoming). These authors show that the variation in word-order V - PP - NP in the Italian examples in exercise 9 in chapter 6 is obtained via two distinct derivations. The pattern in (ib) is derived by rightward movement of the object; while that in (iib) is derived by leftward movement of the PP.

(i) \[ \text{Ho dato \{NP un libro che mi avevano consigliato la settimana scorsa\} \{NP a Gianni\}} \]

\[ \text{have given a book that me they had suggested last week to Gianni} \]

(ii) \[ \text{Ho dato \{NP a Gianni\ \{NP un libro che mi avevano consigliato la settimana scorsa\}} \]

\[ \text{have give to Gianni a book that they me had advised last week} \]

\[ 'I gave a book to Gianni which they had suggested to me last week.' \]

(iia) \[ \text{Gianni ha messo \{NP tre libri\ \{NP sulla tavola\}} \]

\[ \text{Gianno has put three books on the table} \]

\[ 'Gianni has put three books on the table.' \]

(iib) \[ \text{Gianni ha messo \{NP sulla tavola\ \{NP tre libri\}} \]

\[ \text{Gianni has put on the table three books} \]

\[ 'Gianni has put on the table three books.' \]
8.2 PP-extrapolation from NP

In the examples below we concentrate on the bracketed NPs:

102a I read [NP a description [PP of Hockney’s latest picture]] yesterday.
102b I read [NP a description] yesterday [PP of Hockney’s latest picture].

The sentences in (102) are paraphrases. The bracketed PP is the complement of the head N description. We shall assume that the D-structure of the examples corresponds to (102a) and that the surface order of (102b) is derived. The PP of Hockney’s latest picture, has been moved out of the NP and is adjoined to the right. Movement of constituents out of NPs is referred to as extrapolation; PP is extraposed from the object NP.

Let us assume that extrapolation leaves a coindexed trace and that the extrapolated constituent must c-command its trace. Analogously to the discussion of heavy NP-shift above it would not be reasonable to argue that the moved PP is inserted in an unfilled PP position. As before we would have hard time motivating that such a position is projected at D-structure. We conclude that the PP must be in an adjoined position.

Assuming that the extrapolated PP in (102b) is in a derived position as a result of movement, we expect the subjacency effects illustrated in (105).

```
103 *[NP A translation [PP of [NP a description t]]] has appeared [NP of Hockney’s latest picture].
```

If PP is to be construed with the N description then it would have been extraposed out of NP and subsequently out of NP. This is not a possible construal.

8.3 Conclusion

In this section we illustrate two types of rightward movement: heavy NP shift and extrapolation from NP. Both movements are assumed to involve adjunction of the moved constituent to a maximal projection; these examples illustrate A’-movement. The moved element moves to an A’-position, its trace.

---

27 For discussion of extrapolation from NPs see, among others, Coopman and Roovers (1986), Guéron (1980) and Rochemont (1978).

9. Summary

This chapter discusses the properties of wh-movement, an instantiation of A’-movement. A’-movement moves a constituent to an A’-position, leaving a coindexed trace in the base position. Typically wh-movement is involved in the derivation of interrogative sentences where a wh-constituent moves to [Spec, CP]. It is also involved in the formation of relative clauses. We have seen that A’-movement may also affect NPs, as is illustrated in heavy NP-shift, or postnominal modifiers as is seen in extrapolation from NP. A’-movement is either done by substitution (to [Spec, CP]) or by adjunction (heavy NP-shift, extrapolation from NP).

Wh-movement is subject to the subadjacency condition on movement:

1. **Subjacency condition**
   Movement cannot cross more than one bounding node.

The bounding nodes are subject to parametric variation. In English IP and NP are bounding nodes; in Italian CP and NP are bounding nodes.

In addition we discuss two filters which both involve the content of C:

2. **Doubly filled COMP filter**
   When an overt wh-phrase occupies the Spec of some CP the head of that CP must not dominate an overt complementizer.

3. **That-trace filter**
   The sequence of an overt complementizer followed by a trace is ungrammatical.

These filters are language-specific.

Traces of NPs which are wh-moved are characterized as [-anaphor, nonnominal] and are subject to Principle C of the binding theory.

In our discussion we have also paid attention to adjunction structures, illustrated in (4):
In (4) YP is adjoined to XP₁. Adjunction may arise from movement as in heavy NP-shift. Left-dislocation structures suggest that adjunction structures may also be base-generated. We have now extended our phrase structure rules to allow for adjunction:

5a XP → XP₁ YP
5b XP → Spec; X’
5c X’ → X’₁ YP
5b X’ → X; YP.

The relation between the adjoined element and the phrase to which it is adjoined has led us to redefine the notion dominance and also to introduce the concept exclusion:

6 Dominance
A is dominated by B only if A is dominated by every segment of B.

7 Exclusion
B excludes A if no segment of B dominates A.

10 Exercises

Exercise 1
Discuss the motivation for the intermediate traces in the following representations:

1a Who does the detective think \([c_p t]_1 [t_1 \text{ he likes t_1 best}]\)?
1b Who does the detective think \([c_p t]_1 [t_1 \text{ he likes him best}]\)?

Exercise 2
Consider the following sentences, some of which we have already given in the exercises to the Introduction of this book.

1 ?Which man do you wonder when they will appoint?
2 *Who do you wonder which present will give?
3 ??Which present do you wonder who will give?
4 ??Which man do you wonder whether John will invite?
5 *Which man do you wonder whether will invite John?
6 *Which man do you wonder what will give to John?
7 *Which man do you wonder when will invite John?

Although none of the sentences above is entirely acceptable to all native speakers, the degree of unacceptability varies. A sentence with an asterisk is worse than one with a question mark. Try to account for the relative unacceptability of these sentences using the concepts developed in chapter 7. When discussing these sentences you should first of all determine their syntactic representations, D-structure and S-structure. Then you should try to identify which principle or principles are violated.

In your analysis you will no doubt discover that the extraction of a subject wh-constituent is consistently more difficult than that of an object. This type of asymmetry was discussed in terms of the that-trace filter. From the analysis of the examples above try to check whether the filter as formulated in the chapter is adequate and if not, try to reformulate it. In subsequent chapters we shall return to examples such as those above.

If you are a native speaker of a language other than English then check how translations of examples like those above fare in this language.

Exercise 3
In section 6.3.1 we have seen that an analysis of left dislocation in terms of movement and pronoun insertion is not consistent with our present version of the grammar. A structure that closely resembles left dislocation is topicalization:
1a Detective stories, I have never liked them.
1b Detective stories, I have never liked.

(1a) illustrates left dislocation, (1b) topicalization. On the basis of the example above and of the examples given below try to decide whether a movement analysis would be adequate to account for topicalization.

2a ?Detective stories, I wonder if he likes.
2b ?Detective stories, I wonder who reads.
2c *Detective stories, I don't believe the rumour that they will ban.
2d *Detective stories, I don't like linguists who read.
2e Detective stories, I expect will be quite successful.
2f *Detective stories, I expect that will be successful.
2g *Detective stories, I wonder if will be successful.
2h *Detective stories, I wonder when will be successful.

**Exercise 4**

Discuss the derivation (D-structure, S-structure and the various principles that determine them) of the following sentences:

1 Which detective do you think will invite Miss Marple?
2 This is the author whom I like best.
3 Which detective will be invited next week?
4 These are stories which are believed everywhere.
5 Which detective do you think seems to be nicest?
6 Which ships will the enemy sink first?
7 Which ships do you think will sink first?
8 Which sailors do you think will arrive first?

**Exercise 5**

Compare the properties of NP-movement and WH-movement on the basis of section 2.3 in chapter 6, where we discuss the properties of NP-movement. Make a list of similarities and differences between A-chains and A'-chains. This exercise prepares you for the next chapter.

**Exercise 6**

Consider the application of the binding theory in the following examples:

1 Which pictures of himself will John sell?
2 Which pictures of each other do you think that your parents prefer?
3 Those are the pictures of himself which John likes best.
4 Every picture of him, John likes.

Do examples like these produce evidence for Belletti and Rizzi’s (1988) proposal that Principle A can be satisfied either at D-structure or at S-structure? We return to these data in chapter 9.

**Exercise 7**

Consider the application of the binding theory in the following examples:

1 Which pictures of himself, will John sell?
2 Which pictures of each other, do you think that your parents, prefer?
3 Which pictures of himself, does John, think that Jane will sell?
4 Which pictures of himself, does Jane think that John, will sell?
5 Criticize himself, John never will.
6 Criticize himself, Mary never thought that John, would.
7 *Criticize himself John, never thought that Mary would.
8 John, wondered which pictures of himself, Mary liked.

Do the examples above provide support for Belletti and Rizzi’s (1988) claim that Principle A can be satisfied either at D-structure or at S-structure? We return to examples like the ones above in chapter 9.

**Exercise 8**

How can the grammaticality judgements of the following English examples be accounted for?
Introduction to Government and Binding Theory

1. "Which book did you wonder when would be published?"
2. "Who did Poirot tell you why he had interviewed?"

How could we account for the fact that (1) is worse than (2)?
Consider the following data from Italian (Rizzi, 1982b: 54-6). Can the subjacency parameter discussed in section 6.4 account for them?

3. Non so proprio chi possa avere indovinato a chi not know really who could have guessed to whom affiderò questo incarico I'll-entrust this task 'I really don't know who might have guessed to whom I will entrust this task.'

4. *Questo incarico, che non so proprio chi possa avere this task, which not know really who could have indovinato a chi affiderò, mi sta creando un sacco guessed to whom I'll-entrust, me is giving a lot di grattacapi of trouble 'This task, which I really don't know who might have guessed to whom I'll entrust, is giving me a lot of trouble.'

5. Il mio primo libro, che credo che tu sappia a chi my first book, which I believe that you know to whom ho dedicato, mi è sempre stato molto caro. have dedicated, to-me is always been very precious 'My first book, which I know that you know to whom I have dedicated, has always been very dear to me.'

6. *Il mio primo libro, che so a chi credi che abbia my first book, which I know that you-think that I have dedicato mi è stato sempre molto caro. dedicated, to-me is been always very dear.

Jaeggli (1981: 170) gives the following Spanish data:

7. *el único encargo que no sabías a quién iban the only task which you didn't-know to whom they-would a dar give

8. *A quién no sabías qué le regalaron? to whom didn't-you-know what they-had-given

9. *tú hermano, a quién me preguntó que historias le habrán your brother, to whom I-wonder what stories they have contado told.

Assuming that the bounding nodes for subjacency may be parametrized (along the lines suggested by Rizzi 1982b), would the data above suggest that the bounding nodes in Spanish are like in English or like in Italian?

Exercise 9

In section 5.1 we discussed the vacuous movement hypothesis for subject extraction. Consider the following examples from Chomsky (1986b: 50, example (109)). According to Chomsky (1) is more acceptable than (2). Would this contrast in grammaticality be relevant to the discussion in section 5.1?

1. He is the man to whom I wonder who knew which book to give.
2. He is the man to whom I wonder who John told which book to give.

In chapter 5 (section 3.3) we discussed the position of whether as being possibly in [Spec, CP]. According to Chomsky (1986b: 50), example (3) (his (110)) is more acceptable than (2) above.

3. He is the man to whom I wonder whether John told us which book to give.

Does this throw any light on the discussion in chapter 5?