

CHAPTER 3 LOCALITY ON MOVEMENT

3.0 Introduction

The previous chapter has proposed the Immediate Checking Principle (ICP), which requires that uninterpretable formal features (UFFs) should be checked immediately when they become accessible to a computation. It was shown that the ICP subsumes the properties of the two theory-internal linguistic levels in the EST. I have argued that the ICP approach is conceptually more desirable than Chomsky's (1993, 1995) approaches in that the former is local while the latter is global. I have also proposed the Earliness Principle (EP) on Select, which states that lexical items must be selected from a numeration (N) as early as possible. It was shown that the ICP coupled with the EP gives rise to the asymmetry regarding the composition of phrase structure. While arguments are required to be merged cyclically, adjuncts are required to be merged postcyclically. The chapters to follow explore empirical justification for our theory of the composition of phrase structure. I will argue that it receives strong empirical support from a wide range of facts pertaining to movement constraints, scrambling in Japanese, the distribution of *wh*-elements in-situ, and reconstruction effects. These empirical facts also lend support for the view that language is derivational in character rather than in the representational mode. This is because they can only be given a principled minimalist account by making use of information which is available at an intermediate stage of a derivation but later "wiped-out" by an operation before the output representation. Among these empirical facts, I will first consider locality

conditions on feature-driven A'-movement. Specifically, this chapter considers the "domain barrier" effects with feature-driven A'-movement, i.e., the Complex NP Constraint, the Adjunct Condition, the Subject Condition, and the non-bridge verb condition, and the ban against feature-driven extraction out of phrases which have undergone feature-driven A'-movement. I will argue that these locality conditions, which are left unexplained under the MP, straightforwardly follow from our theory of phrase structure. It is pointed out that our approach to these locality conditions diverges from all previous approaches in claiming that they should not follow from restrictions on movement but from restrictions on merger.

The organization of this chapter is as follows. Section 3.1 reviews previous generative analyses of locality restrictions on feature-driven A'-movement, taking overt wh-movement in English as an example. Section 3.2 deals with the "domain barriers," which have not been given any principled account under the MP. It is shown that if our theory of phrase structure is adopted, the "domain barrier" effects straightforwardly follow. Section 3.3 considers another locality condition on feature-driven A'-movement, i.e., no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. I will argue that this locality condition also follows from our theory of phrase structure. Section 3.4 discusses locality restrictions on adjunct feature-driven A'-movement. I will argue that they can also be accounted for by our theory of phrase structure. Section 3.5 makes concluding remarks.

3.1 Previous Analyses of Locality on Movement

3.1.1 Standard Theory Approaches

Locality conditions on movement have been one of the central issues for generative grammar. A lot of discussion has especially been around regarding locality conditions on so called feature-driven A'-movement like overt wh-movement and null operator movement. Feature-driven A'-movement is *prima facie* unbounded in the sense that it can extract an element out of deeply embedded phrases. In fact, however, A'-movement is not unbounded, but is restricted by some constraints on movement. It is Chomsky (1964) which marks the first substantial step toward establishing a general locality condition on movement. Chomsky (1964) proposes the A-over-A principle. The A-over-A principle prohibits movement operations from applying to phrases contained within those of the same category. Although the A-over-A principle is conceptually attractive in that it attempts to reduce all locality conditions to one explanatory principle, it faces serious empirical problems. It has been pointed by, among others, Chomsky (1964, 1968) and Ross (1967) that the A-over-A principle is too strong as well as too weak.¹

As an alternative to the A-over-A principle, Ross (1967) proposes island constraints on movement. Island constraints informally state that no elements can be extracted out of domains called "islands." Island

¹See Fukui (1997) for a minimalist account of island constraints based on the insight of the traditional A-over-A principle. As Fukui himself admits, however, the classical problems of the traditional version of the A-over-A principle are still left unexplained under his analysis.

constraints which are proposed in Ross (1967) and later works are shown below with their examples:²

- (1) Complex NP Constraint (CNPC)
 - a. Relative Clauses
*?**who** do you like [books that criticize *t*]
 - b. Non-relative Complex NPs
*?**what** did you study [the evidence that Harry stole *t*]
- (2) Subject Condition
 - a. Subject NPs (cf. Chomsky (1973))
*?**who** did [pictures of *t*] please you
 - b. Sentential Subjects
*?**what** did [that John saw *t*] surprised Mary
- (3) Adjunct Condition
 - a. Adjunct Clauses
*?**who** do you get jealous [because I spoke to *t*]
 - b. Adverbial PPs
*?**which class** did he fall asleep [during *t*]
 - c. PP Modifiers of Nouns
*?**which table** did you buy [the pictures [on *t*]]
 - d. Secondary Predicates
*?**what** did John arrive yesterday, [sad about *t*]
- (4) Non-bridge Verb Condition (cf. Erteschik-Shir (1973))
?**who** did John whisper [that he saw *t*]

²Ross (1967) also proposes the left-branching condition and the Coordinate Structure Constraint. We will put aside these constraints, since the nature of these phenomena is not entirely clear at this point.

(5) Wh-island Constraint (cf. Chomsky (1977))

*?what_i did you decide [who_j to persuade *t_j* to buy *t_i*]

Although these island constraints are empirically adequate, they are conceptually unattractive. This is because the island constraints are construction-specific and thus just observational generalizations rather than abstract explanatory principles.³

³There are some languages which prima facie do not fully obey these island constraints. It has been claimed by, among others, Rizzi (1982), Torrego (1986), and Sportiche (1981) that French does not exhibit any subject condition effects:

- (i) le diplomate **dont** [[la secrétaire *t*] t'a téléphoné]
'the diplomat of whom a secretary called you'

In the relative clause (i), extraction takes place from within the subject, but the result is acceptable. It has also been claimed by, among others, Allwood (1976), Erteschik (1973), Engdahl (1980b, 1982), Hellan and Christensen (1986), and Taraldsen (1978, 1982) that Scandinavian languages do not obey the Wh-island Constraint or the CNPC, which is exemplified by the following Norwegian examples:

- (ii) **denne boken** vet vi hvem som har skrevet *t*
this book know we who that has written
'this book that we know who has written'
- (iii) a. **denne boken** kjenner vi den mannen som har skrevet *t*
this book know we the man that has written
'this book that we know the man who has written'
- b. **denne boken** går det rykter om at du har lest *t*
this book go it rumors about that you have read
'this book that the rumor that you have read is around'
(Hellan and Christensen 1986:4)

Although extraction takes place from within the Wh-island in (ii) and the complex NP in (iii), these examples are all acceptable.

The nature of these phenomena, however, are not entirely clear at this point. Tellier (1990) discusses the Subject Condition in French, pointing out that the immunity from the Subject Condition is only observed with *dont*. The following examples show that French in fact obeys the Subject Condition:

- (iii) a. ?*le diplomate **de qui** [[la secrétaire *t*] t'a téléphoné]
'the diplomat of whom the secretary called you'
- b. ?***de qui** [[la secrétaire *t*]t'a-t-elle téléphoné]
'of whom did the secretary call you'

(Tellier 1990:307)

Turning to the lack of the Wh-island and the CNPC in Scandinavian languages, relativization in Scandinavian languages may not involve movement, as argued for Japanese relative clauses by some literatures (see, among others, Kuno (1973)). We will therefore put aside these phenomena in the discussion to follow.

3.1.2 EST Approaches

Since the advent of the island constraints, a considerable number of studies have been made on reduction of these island constraints to an explanatory principle within the framework of the EST (see, among others, Aoun (1985a, 1985b), Cattell (1976), Chomsky (1973, 1976, 1981, 1986a), Cinque (1990), Culicover and Wilkins (1984), Depréz (1989), Hegarty (1991), Huang (1982), Koster (1978, 1987), Kayne (1983), Lasnik and Saito (1992), Manzini (1992), and Pesetsky (1982)). All of these theories, however, crucially make use of notions which are no longer available in the MP where no structural relations are allowed to be invoked other than those required by BOCs and those induced in a natural way by the derivation itself like local relations to a head in its minimal domain. Hence, these EST analyses are incompatible with the MP. Let us quickly review Chomsky's (1986a) and Kayne's (1983) locality theories as representatives among locality theories in the EST.

Chomsky (1986a) proposes the notion of barrier, which is instrumental for the bounding theory as well as the government theory. The notion of barrier is defined based on the notion of Blocking Category (BC). Maximal projections which are not L-marked constitute BCs for their containing categories. Maximal projections which are L-marked as sisters to θ -role assigners, on the other hand, do not constitute BCs for their containing categories. Any category except IP that is a BC becomes a barrier. Barrierhood is also obtained through the inheritance mechanism. Any category that immediately dominates a BC inherits barrierhood. Movement is constrained by two locality restrictions, i.e., antecedent government and the subjacency condition, which are both defined based on the notion of barrier. According to the antecedent

government requirement, a trace left by movement may not be separated by its antecedent by any barriers. If an LF representation contains any trace which violates the antecedent government requirement, then the derivation is excluded by the Empty Category Principle (ECP). The subjacency condition also requires that a trace should not be separated by its antecedent by any barriers. Subjacency violations, however, are less severely deviant than ECP violations, which explains the argument/adjunct asymmetry with the degree of unacceptability of island violations. Although Chomsky's locality theory can subsume the island constraints, it crucially makes use of notions which are not available in the MP like government, antecedent government, L-marking, BC, and barrier. Hence, Chomsky's locality theory is incompatible with the MP.

Kayne (1983) proposes the connectedness condition (CC). The basic idea of the CC is that syntactic dependencies must be mediated by a path of nodes that link the related elements. Syntactic dependencies are only licit if dependents and their antecedents are members of a common well-formed subtree. If we restrict our discussion to a derivation where only one instance of movement is involved, the CC informally states that a trace must have a c-commanding antecedent that is contained within a g-projection of its structural governor. In other words, a trace and its antecedent must be members of a common subtree that is a g-projection of the structural governor of the trace. It then follows that a trace which does not have any structural governor may not be unbounded. Outside the first g-projection, which is determined by X'-principles and structural government, further g-projections are licensed by being in a canonical government configuration with a head. In languages like English where V governs NP to its right, [Y W Z] is a canonical government configuration,

with Z a maximal projection. Hence, in those languages, a trace may not be unbounded within a maximal projection that constitutes a left branch. Kayne's locality theory is conceptually attractive in that it subsumes the subjacency condition and the ECP under the CC. His theory, however, crucially makes use of notions which are no longer available in the MP like g-projection, government, canonical government, and structural government. Hence, it is incompatible with the MP.⁴

3.1.3 Minimalist Approaches

The previous subsection has reviewed the EST approaches to the island constraints. It was pointed out that although they are conceptually more attractive than Ross (1967), they crucially make use of notions which are no longer available in the MP. This subsection reviews previous minimalist approaches to the locality restrictions on feature-driven A'-movement.

Among the island constraints, the Wh-island Constraint has been the main object of study since the advent of the MP. One can cite Chomsky's (1993, 1994, 1995) approach to the Wh-island Constraint. Based on the insight provided by Rizzi's (1990) Relativized Minimality, Chomsky (1993, 1994, 1995) reformulates the Wh-island Constraint as a Minimal Link Condition (MLC), which is one of the economy conditions in Chomsky (1993, 1994) and later incorporated into the definition of

⁴Although Kayne's (1983) CC cannot be accommodated under the MP as it is, it can be translated into minimalist terms. One can safely say that the locality theory to be proposed here is a minimalist reinterpretation of Kayne's CC. Leaving the details to the discussion later in this chapter, I argue that structures belonging to different g-projections in the sense of Kayne have not been merged with each other when feature-driven A'-movement takes place. Hence, movement cannot take place across more than one g-projections.

Attract-F in Chomsky (1995).⁵ Only few attempts, however, have so far been made at providing a minimalist account of the other island constraints, i.e., the Complex NP Constraint, the Subject Condition, the Adjunct Condition, and the non-bridge verb condition. We will hereafter call these island constraints "domain barriers" as opposed to the relativized minimality effects including the Wh-island Constraint. In the subsections to follow, I will review two previous studies which attempt to give a minimalist account of the "domain barriers," pointing out that they are conceptually problematic.

3.1.3.1 Takahashi (1994)

Takahashi (1994) gives a minimalist account of the island constraints, claiming that the island constraints should be subsumed under the Shortest Movement Condition (SMC) and the Uniformity Corollary on Adjunction (UCA):

- (6) The Shortest Movement Condition (SMC)

Make the shortest movement.

(Takahashi 1994:8)

- (7) The Uniformity Corollary on Adjunction (UCA)

If $(\alpha_1, \dots, \alpha_n)$ is a chain ($1 \leq n$), then for any i ($1 \leq i \leq n$), $P(\alpha_i)$.

(Takahashi 1994:20)

⁵See Ishii (1997) for the different view that the Wh-island Constraint should not be subsumed under the MLC but under the "no extra baggage" condition. The MLC is instead responsible for the "crossing" constraint effects. It is shown that this analysis gives us a minimalist account of the contrast in acceptability between the "crossing" constraint and superraising effects, on the one hand, and the Wh-island effects, on the other.

In (7), $P(\alpha)$ means that α has property P, which is the property of being adjoined to by X' or being not adjoined to. (7) virtually ensures that adjunction is impermissible to heads of nontrivial chains. This is because if something is adjoined to the head of a chain, then the head becomes distinct from the rest of the chain regarding property P. As for the Wh-island Constraint, Takahashi essentially follows the relativized minimality account advocated by Chomsky (1993, 1994, 1995), Chomsky and Lasnik (1993) and Rizzi (1990), claiming that it is subsumed under the SMC. Hence, not much has to be said about this. The discussion to follow therefore only concerns his analysis of the "domain barriers."

Let us look at how Takahashi's (1994) analysis works, considering the Subject Condition first:

(8) *?**who** did [pictures of t] please you

Since he adopts the VP-internal subject hypothesis, the subject *pictures of who* originates in the Spec of VP and then moves to the Spec of IP, as shown below:

(9) [C' C [IP [D' [NP **pictures of who**]]_j [t_j please you]]]

The *wh*-phrase *who* moves from within Spec of IP to the Spec of CP. The SMC requires that *who* should move in a successive cyclic fashion, landing at the nearest possible target at each step of movement, as shown below:

(10) [CP **who**_i did [IP t'''_i [IP [DP t''_i [D' [NP t'_i [NP pictures of t_i]]]]]_j [t_j please you]]]]

In (10), *who* first adjoins to NP and then moves to the Spec of DP. Since Takahashi assumes that specifiers are created by adjunction, the latter step actually involves adjunction to D'. It then adjoins to IP and finally moves to the Spec of CP, which counts as adjunction to C'. This

derivation, however, is excluded by the UCA, which applies at LF. This is because *who* is adjoined to D', the head of the non-trivial chain of the subject raising (*pictures of t_i, t_j*). The subject raising chain is therefore non-uniform; this violates the UCA. Hence, the Subject Condition can be accounted for by the SMC and the UCA.

Takahashi furthermore extends the UCA (7), claiming that it applies to coordinate structures as well as chains:

(11) The Uniformity Corollary on Adjunction (Revised)

If $(\alpha_1, \dots, \alpha_n)$ is a chain ($1 \leq n$) or $\alpha_1, \dots, \alpha_n$ are conjuncts of a coordination, then for any i ($1 \leq i \leq n$), $P(\alpha_i)$.

(adapted from Takahashi 1994:25)

He argues that the basic idea of the UCA is that elements belonging to a group should respect uniformity concerning the adjunction property. Elements of a chain count as belonging to a group and thus chains are subject to the UCA. He argues that it is intuitively correct to say that conjuncts of a coordination also count as belonging to a group. It is then plausible to assume that the UCA applies not only to chains but also to coordinate structures. If this conjecture is correct, then the Coordinate Structure Constraint also follows from the SMC and the UCA:

(12) *?which book was John reading *t* and driving cars

In the derivation of (12), before wh-movement takes place, the VP *reading which book* and the VP *driving cars* are conjoined. The SMC requires that *which book* adjoins to the former VP on its way to the Spec of CP. Adjunction to a conjunct, however, violates the UCA (11).

Extending this analysis of the Coordinate Structure Constraint, he argues that the Adjunct Condition is also subsumed under the SMC and the UCA. Following Higginbotham (1985), he claims that adjuncts

involve coordination. (13a), for example, is assigned semantic representation (13b):

- (13) a. John walks slowly
 b. $\exists e$ [Walk (John, *e*) & Slow (*e*)]

In (13b), the adjunct *slowly* is coordinated with the main predicate. It then follows that the Adjunct Condition can be accounted for in the same fashion as the Coordinate Structure Constraint:

- (14) *?who do you get jealous [because I spoke to *t*]

In the derivation of (14), the SMC requires that *who* adjoins to the adjunct on its way to the Spec of CP. Adjunction to the adjunct, however, is prohibited by the UCA (11), since the adjunct counts as a conjunct of the coordinate structure. Hence, the Adjunct Condition follows.

He furthermore argues that the CNPC can be accounted for in the same way as the Adjunct Condition. Since relative clauses are adjuncts in nature, the relative clause case of the CNPC is subsumed under the Adjunct Condition. Concerning non-relative complex NPs, Takahashi (1994) assumes following Grimshaw (1990) and Stowell (1981) that noun complement clauses are appositive and thus regarded as adjuncts. Then, the non-relative case of the CNPC is also subsumed under the Adjunct Condition.

Although Takahashi's analysis is attractive in reducing all the island constraints including the Coordinate Structure Constraint to the two general principles of grammar, i.e., the SMC and the UCA, his analysis is theoretically problematic. His analysis of the Adjunct Condition (and therefore the CNPC) crucially makes use of the idea that adjuncts involve coordination. As Takahashi himself admits, however, representations like (13b) where adjuncts involve coordination should not

count as LF representations but as semantic representations. The information that adjuncts involve coordination can never be read off in the syntactic component. It is only available in the semantic component. I argue that even if it is true that adjuncts semantically involve coordination, his UCA analysis of the Adjunct Condition is untenable. This is because the autonomous syntax thesis requires that the UCA, being syntactic in nature, should not refer to information only available in the semantic component.

3.1.3.2 Toyoshima (1997)

Toyoshima (1997) gives a minimalist account of the locality restrictions which were captured by the Condition of Extraction Domain (CED) within the EST. The CED was originally proposed by Huang (1982) and intended to subsume the Subject Condition and the Adjunct Condition. Toyoshima proposes the Derivation Condition on Extraction Domain (Derivational CED), claiming that it provides a minimalist account for the Adjunct Condition as well as the Subject Condition:

- (15) Derivational Condition on Extraction Domain
(Derivational CED)

A feature F is accessible for Attract triggered by another feature F' iff F and F' are both introduced as a part of the same process P.

(Toyoshima 1997:510)

The notion of process is defined as below:

(16) Process

A process P is a sequence of operations (OP), Merge or Attract, such that if OP $(\alpha, \beta) = \gamma$ is in P , then Merge (δ, γ) is also in P .

(Toyoshima 1997:511)

Let us look at how the Derivational CED works, taking the Subject Condition as an example:

(17) *?who did [pictures of t] please you

According to Toyoshima's analysis, the derivation of (17) would proceed in the following way:

- (18) ia. Merge (*please, you*)_{ia}. Merge (*of, who*)
 = *please you* = *of who*
- iib. Merge (*pictures, of who*)
 = *pictures of who*
- ib. Merge (*pictures of who, please you*)
 = *pictures of who please you*
- ic. Merge (T, *pictures of who please you*)
 = T *pictures of who please you*
- id. Attract (T, *pictures of who*)
 = [***pictures of who***] T *t please you*
- ie. Merge (C, [***pictures of who***] T *t please you*)
 = C [***pictures of who***] T *t please you*

The next step of this derivation should be to check the strong Q-feature of C through attracting the Q-feature of *who* in the subject phrase. The derivational CED, however, claims that this Attract is illegitimate. Derivation (18) involves two different processes, i.e., (i) and (ii). While C is introduced in process (i), *who* is introduced in process (ii). According to

the derivational CED, therefore, the Q-feature of *who* is not accessible for Attract triggered by the Q-feature of C. There is no way of checking the strong Q-feature of C; this derivation crashes. Hence, the Subject Condition follows. The Adjunct Condition can be accounted for in a similar fashion. Since a *wh*-phrase in an adjunct is introduced in a different process from C with a strong Q-feature, the former is inaccessible for Attract triggered by the latter. Hence, the strong Q-feature remains at the interface; the derivation crashes.

Toyoshima's derivational CED, however, has a conceptual problem. The derivational CED raises the problem of globality. Let us consider derivation (18) as an example. When we reach the stage of the derivation where the strong Q-feature of C is to be checked, it is necessary to know whether C and *who* are introduced in the same process or not. This information, however, can never be read off from the phrase structure at this stage of the derivation. In other words, although C and *who* are introduced in different processes at the earlier stages of the derivation, this information is no longer available at the stage where the Q-feature of C is to be checked. Under Toyoshima's analysis, therefore, we need global considerations, looking back to see whether the two elements are introduced in the same process or not. This is theoretically undesirable, since global considerations necessarily induce computational intractability.⁶

⁶Toyoshima's analysis also has empirical problems. Let us consider *wh*-movement out of a passive subject, taking (i) as an example:

(i) *?who_i were [pictures of t_i] stolen t_j

It has been assumed that (i) is excluded by the Subject Condition. According to Toyoshima, therefore, the deviancy of (i) should be accounted for by the Derivational CED, which is intended to subsume the Subject Condition. The Derivational CED, however, would wrongly predict that examples like (i) are acceptable. In (i), although

To summarize subsection 3.1.3, I have reviewed Takahashi's (1994) and Toyoshima's (1997) minimalist approaches to the island constraints. It was shown that they are confronted with conceptual problems. In the rest of this chapter, I will only consider the "domain barriers," which have not been given any principled account under the MP, putting aside the relativized minimality effects. The organization of the rest of this chapter is as follows. Section 3.2 argues that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure. In section 3.3, I will extend our analysis of the "domain barriers" to another locality phenomenon, i.e., no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. It is shown that this locality condition also follows from our theory of phrase structure. Section 3.4 considers locality restrictions on adjunct A'-movement. It is shown that they also follow from our theory of phrase structure.

pictures of who is in the Spec of TP when wh-movement applies, it is originally introduced in the complement position of the verb *steal*. Hence, C and *who* are introduced in the same process. The Derivational CED would claim that the Q-feature of *who* is accessible for Attract triggered by the Q-feature of C, which is undesirable. The same problem arises with wh-movement out of the subject of an unaccusative verb.

In order to solve this problem, Toyoshima claims that overt attraction of a complex category requires what he calls a subparallel process, which "constructs" a complex structure internally and in parallel to a given process. In (i), the subject of the passive verb, being a complex category, undergoes overt attraction. It therefore involves a subparallel process. The Q-feature of the *wh*-phrase within the subject is not accessible to the Q-feature of C due to the subparallel process. The subparallel process, however, is not independently motivated and only needed to solve this problem.

⁶Recall that Chomsky's (1995) definition of Attract/Move requires that what is raised should be c-commanded by its target. Under our analysis where Attract/Move is reinterpreted as consisting of Copy and Merge, Copy is subject to the c-command requirement. The c-command requirement on Copy, however, is not included in the definition of the last resort condition (33), since it is most likely to be derived from some general principles of UG.

3.2 An Account of the "Domain Barriers"

In this section, I will argue that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure. As argued in the previous chapter, the ICP coupled with the EP ensures that adjuncts, whose merger is not triggered by any UFF, are forced to be merged postcyclically. It is shown that the "domain barriers" are exactly those elements which are forced to be merged postcyclically. It then follows that when we come to a stage where the strong Q-feature of C becomes accessible to a computation, the "domain barrier" has not been merged with the main structure which contains the C. The main structure and the "domain barrier" each constitute an independent syntactic object at that stage. No Q-feature inside the "domain barrier" therefore can be attracted by the strong Q-feature of C. This violates the ICP; the derivation is canceled. In the following, we will look at in detail how this analysis works using concrete examples.

3.2.1 The Complex NP Constraint (CNPC)

We will begin by considering the relative clause case of the CNPC, taking (1a) (repeated here as (19)) as an example:

(19) *?**who** do you like [books that criticize *t*]

This structure can be divided into two parts: the main structure *who do you like books* and the relative clause *that criticize t*. The relative clause can be constructed by checking the UFFs of the selected items in accordance with the ICP and the EP. Given the copy theory of movement and the empty operator movement analysis of English relative clauses, the resultant structure is as follows. Among formal features, we will

only pay attention to Q-features here and in the relevant structures to follow, with irrelevant formal features being ignored:

$$(20) \quad [C^{\max} Op [C [T^{\max} Op [T [V^{\max} Op \quad [criticize \text{ who}_{[Q]}]]]]]]]$$

Turning to the main structure of (19), we first select the verb *like*. The ICP then requires that we should immediately check the internal thematic feature and Accusative Case feature of *like*, since these features, being uninterpretable, are accessible to the computation. In order to check these features, we select *books* and merge *like* with *books*. At this point, the external thematic feature of *like* becomes accessible to the computation. Since it is uninterpretable, it must be checked immediately by combining *you* and *like books*. The resultant structure is as below:

$$(21) \quad [V^{\max} \text{ you } [\text{like books}]]$$

The relative clause (20) is eventually merged with the D^{\max} *books* in order to be properly interpreted at LF. We assume following, among others, Browning (1987) that its interpretation is established through a predication relation between the head noun and the empty operator in the relative clause in the sense of Rothstein (1983) and Williams (1980). There is, however, no legitimate way of combining them before *books* is merged with *like*. This is because when *like* is selected, the ICP requires that the internal thematic feature and Accusative Case feature of *like* should be checked immediately by combining *like* with *books* without being interrupted by any other LF operation. Hence, merger of *books* and the relative clause (20) may not be applied at this point. Similarly, the ICP prohibits the relative clause (20) from being merged with *books* before merger of *like books* with *you*.

The next step is to select T. Note that the EP prohibits the relative clause (20) from being merged with the main structure before we select T. The ICP then requires that the V-feature of T, being uninterpretable, should be checked immediately by combining T with the V^{\max} (21). Then, the Nominative Case feature and D-feature of T become accessible to the computation. Since they are uninterpretable, they must be checked immediately by the raising of *you* to the Spec of T^{\max} . When T^{\max} is constructed, the next step must be to select C. Note again that due to the EP, we cannot merge the relative clause (20) with the main structure before we select C. The ICP requires that the T-feature of C, being uninterpretable, should be checked immediately by merging C with T^{\max} :

$$(22) \quad [C^{\max} C_{[Q]} [T^{\max} \text{you} [T [V^{\max} \text{you} [\text{like books}]]]]]$$

At this point, the strong Q-feature of C becomes accessible to the computation. Since it is uninterpretable, the ICP requires that it should be checked immediately. There is, however, no way of checking the strong Q-feature of C immediately. This is because the relative clause (20), which contains the *wh*-phrase *who*, could not have been merged with the main structure until this point of the derivation due to the ICP and the EP. In other words, the relative clause and the main structure each constitute an independent syntactic object at this stage of the derivation. Since C and *who* belong to different phrase markers at this stage, the former does not c-command the latter. The strong Q-feature of C cannot be checked by the raising of *who* to the Spec of C^{\max} . Since the strong Q-feature cannot be checked immediately, it violates the ICP; the derivation is canceled. To put it another way, our analysis claims that the raising of *who* to the Spec of C^{\max} would be an instance of "movement across

phrase structures," which is prohibited due to the c-command requirement on movement. Therefore, the deviancy of (19), an example of the relative clause case of the CNPC, straightforwardly follows.

Let us next consider the non-relative case of the CNPC, taking (1b) (repeated here as (23)) as an example:

(23) *?what did you study [the evidence that Harry stole *t*]

We follow Grimshaw (1990), Stowell (1981), and Takahashi (1994) in claiming that the head nouns of non-relative complex NPs like *evidence* in (23) do not assign any θ -roles to the following clauses.

It first appears that the relation between the noun and the following clause in (25a-c) is parallel to the one between the corresponding verb and its complement in (24a-c):

- (24) a. Andrea guesses [that Bill was lying]
 b. John claimed [that he would win]
 c. Paul explained [that he was temporarily insane]
- (25) a. Andrea's guess [that Bill was lying]
 b. John's claim [that he would win]
 c. Paul's explanation [that he was temporarily insane]

(Stowell 1981:199)

Stowell (1981) argues, however, that these two relations are not parallel to each other. While the verbs *guess*, *claim*, and *explain* refer to the actions of guessing, claiming, and explaining something, the nouns *guess*, *claim*, and *explanation* do not. Rather, the nouns *guess*, *claim*, and *explanation* refer to a thing which is guessed, a thing which is claimed, and a thing which is explained, respectively. In other words, the nouns refer to the same things as the following clauses. Hence, the relation

between the nouns and the following clauses in examples like (25a-c) is an appositive one rather than one of a θ -role assignment.

Stowell (1981) presents evidence in support of his appositive analysis. As shown in (26), the identity relation holds between the noun and the following clause:

- (26) a. [Adrea's guess] was [that Bill was lying]
 b. [John's claim] was [that he would win]
 c. [Paul's explanation] was [that he was temporarily insane]

(Stowell 1981:200)

This fact further supports the claim that the head nouns of the non-relative complex NPs do not assign any θ -roles to the following clauses.

Let us return to (23). Since merger of the appositive clause in the non-relative complex NP is not triggered by any UFF, the ICP coupled with the EP ensures that it can only be merged with the main structure postcyclically. It then follows that when we come to the stage of the derivation where the strong Q-feature of C in the main structure is to be checked, the appositive clause which contains *who* has not been merged with the main structure yet. The appositive clause and the main structure each constitute an independent syntactic object at that stage. Since C and *who* belong to different phrase markers, the former does not c-command the latter. The strong Q-feature of C cannot be checked immediately by the raising of *who*; this violates the ICP. This derivation is therefore canceled. The non-relative clause case of the CNPC follows.

3.2.2 The Adjunct Condition

The Adjunct Condition can be analyzed essentially along the same line as the CNPC. Let us consider (3a-d) (repeated here as (27a-d)) as examples:

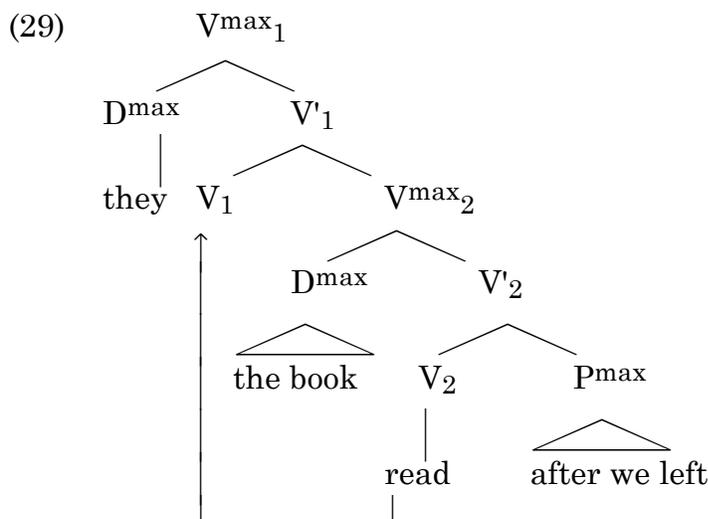
- (27) a. Adjunct Clauses
 *?**who** do you get jealous [because I spoke to *t*]
- b. Adverbial PPs
 *?**which class** did he fall asleep [during *t*]
- c. PP Modifiers of Nouns
 *?**which table** did you buy [the pictures [on *t*]]
- d. Secondary Predicates
 ***what** did John arrive yesterday, [sad about *t*]

Note that merger of adjuncts like those in (27) and main structures is not triggered by any UFFs. It then follows from the ICP and the EP that adjuncts are required to be merged postcyclically. Hence, when we come to the stage when the strong Q-feature of C in the main structure is to be checked, the adjunct has not been merged with the main structure yet. The adjunct and the main structure each constitute an independent syntactic object at that stage. Hence, C does not c-command the *wh*-phrase in the adjunct clause. The strong Q-feature of C cannot be checked immediately by the raising of the *wh*-phrase in the adjunct clause. The derivation is canceled due to a violation of the ICP. The Adjunct Condition straightforwardly follows.

Recall that according to the MP where the bare phrase structure is assumed, adjuncts may appear in so called "specifier or complement positions," as argued by the Larsonian analysis of adjuncts advocated by, among others, Chomsky (1995), Larson (1988), McConnell-Ginet (1982),

and Stroik (1990). For example, under the Larsonian analysis of adjuncts, the structure underlying (28) is (29):

(28) they read the book after we left



In (29), the adjunct clause *after we left* appears as the sister of V_2 in the underlying structure. The verb *read* raises to V_1 . Under this analysis, since adjuncts as well as arguments may appear as the sisters of V , there is no way of making a distinction between the two by means of structural terms. Recall that all the previous analyses of the Adjunct Condition but Takahashi (1994) crucially rely on the assumption that arguments and adjuncts appear in structurally distinct positions. They assume that adjuncts appear in adjoined positions whereas arguments appear in either specifier or complement positions. Hence, those analyses are incompatible with the Larsonian analysis of adjuncts and thus the bare phrase structure.

Our analysis of the Adjunct Condition, on the other hand, is compatible with the Larsonian analysis of adjuncts and thus the bare phrase structure. Recall that our analysis defines the notions of adjunct and complement derivationally. Under our analysis, the difference

between adjuncts and complements resides in the fact that the former is required to be merged postcyclically whereas the latter is required to be merged cyclically. Crucially, when the strong Q-feature of C is to be checked, adjuncts have not been merged with main structures; the Adjunct Condition follows. Hence, our analysis of the Adjunct Condition works no matter where adjuncts may appear structurally.

3.2.3 The Non-bridge Verb Condition

Let us next consider the non-bridge verb condition, taking (4) (repeated here as (30)) as an example:

(30) ?**who** did John whisper that he saw *t*

As first observed by Erteschik-Shir (1973), extraction out of the "complements" of non-bridge verbs like *whisper*, *whine*, and *murmur* is illegitimate. Kayne (1983) and Stowell (1981) claim that non-bridge verbs do not assign any θ -roles to their "complements." In our terms, this observation can be interpreted as follows. Non-bridge verbs do not have any thematic features which are to be checked by their merger with their internal arguments. In other words, the clauses following non-bridge verbs should count as clausal adjuncts rather than as complements.

If this conjecture is correct, the ICP coupled with the EP enables us to account for the non-bridge verb condition. Note that merger of the clausal adjunct with the main structure is not triggered by any UFF. It then follows that the clausal adjunct is forced to be merged postcyclically. Hence, when we come to the stage when the strong Q-feature of C in the main structure is to be checked, the clausal adjunct has not been merged with the main structure yet. Hence, C does not c-command the *wh*-phrases in the clausal adjunct. The strong Q-feature of C cannot be

checked immediately. This derivation is canceled due to a violation of the ICP. The non-bridge verb condition straightforwardly follows.

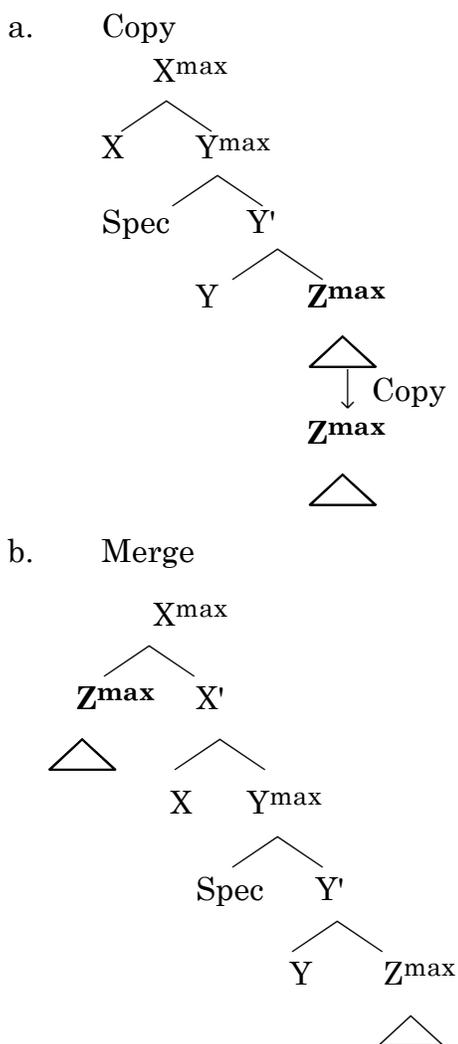
3.2.4 The Subject Condition

This subsection discusses the Subject Condition, arguing that it also follows from the ICP coupled with the EP. Before turning to the Subject Condition, it is necessary to explicate the notion of Attract/Move. I will argue that Attract/Move is not a primitive operation but a complex operation consisting of two primitive operations, i.e., Copy and Merge. It is shown that unlike the standard Attract/Move approach, our Copy + Merge approach only makes use of the operations of deletion and erasure, entirely dispensing with the operation of checking. The effects of the operation of checking under Chomsky's (1993, 1994, 1995) analysis are brought about by the operation of deletion or erasure which makes UFFs invisible at LF under our analysis. I will argue that this constitutes theoretical support for our Copy + Merge approach, since the operation of checking is untenable within the most recent development of the MP. Since no operation of checking is posited under our Copy + Merge approach, the ICP should be reformulated accordingly. I will argue that the ICP should be reformulated as a condition on LF visibility, which requires that UFFs should be made invisible at LF immediately by deletion or erasure when they become accessible to a computation. Finally, I will propose that erasure should be interpreted as conversion of UFFs to phonological features. It is shown that this conversion ensures that all UFFs must be erased before Spell-Out.

3.2.4.1 Attract/Move = Copy + Merge

Following Bobaljik (1995) and Collins (1997), I claim that there is no primitive operation Attract/Move. Attract/Move in fact is a complex operation, consisting of two primitive operations, i.e., Copy and Merge. The operation Copy creates a new, independent syntactic object by copying a term within a syntactic object. This copied term is available to be merged with another syntactic object. For example, the raising of Z^{\max} from the complement position of Y to the Spec of X^{\max} proceeds as depicted below:

(31) Move = Copy + Merge



First, Z^{\max} undergoes the copy operation, creating another Z^{\max} . This newly created Z^{\max} is then merged with the X^{\max} in (31a), yielding the larger X^{\max} as in (31b).

Let us consider the operation Copy in detail. Suppose that H is a head dominating a feature F_1 . I argue that the copy operation is subject to the following last resort condition (though the latter is subsumed under the revised ICP to be proposed below):⁷

- (32) α undergoes the copy operation only if F_1 of H enters into a deletion relation with F_2 of α .

In (32), α is the minimal element including F_2 that allows for convergence, as required by the "no extra baggage" condition. Recall that under the standard Attract/Move approach, the raising of α to target K only takes place if F of α enters into a checking relation with a sublabel of K (where a sublabel of K is a feature of the zero-level projection of the head $H(K)$ of K). Under Copy + Merge approach, on the other hand, Copy, the first half of the complex operation Attract/Move, only takes place if a deletion relation, but not a checking relation, is established.

The definition of a deletion relation is shown below:

- (33) F_1 enters into a deletion relation with F_2 iff F_1 is deleted.
(F_2 may also be deleted.)

Given that deletion only takes place under identity, F_1 deletes only if identical with F_2 . Let us assume following Chomsky (1995) that deletion

⁷Recall that Chomsky's (1995) definition of Attract/Move requires that what is raised should be c-commanded by its target. Under our analysis where Attract/Move is reinterpreted as consisting of Copy and Merge, Copy is subject to the c-command requirement. The c-command requirement on Copy, however, is not included in the definition of the last resort condition (33), since it is most likely to be derived from general principles of UG.

makes formal features invisible at LF. Formal features may be made invisible at LF by deletion unless it would contradict the principle of recoverability, which states that unrecoverable items may not be deleted. Interpretable features receive interpretations at LF. Hence, they may not be made invisible at LF by deletion due to the principle of recoverability. Uninterpretable features, on the other hand, may be made invisible at LF by deletion. This is because uninterpretable features do not contribute to any content at LF and thus their deletion does not violate the principle of recoverability. Hence, in the definition of a deletion relation (33), F_1 must be a UFF. When F_2 happens to be a UFF, it is also deleted. It then follows from (32) and (33) that Copy, the first half of the complex operation Attract/Move, only takes place if a UFF enters into a deletion relation.

The UFFs which undergo deletion through Copy are further subject to erasure. Erasure is another operation which makes formal features invisible at LF. Erasure, a stronger form of deletion, not only makes formal features invisible at LF but also inaccessible to a computation (see Chomsky (1995)).⁸ I argue that UFFs, whether deleted or undeleted, are required to be erased and thus made inaccessible to a computation before Spell-Out. Formal features can only be erased when they enter into an erasure relation. Suppose that H is a head dominating a feature F_1 . The definition of an erasure relation is shown below:

⁸Our interpretation of erasure is different from Chomsky's (1995). The latter claims that only deleted formal features are subject to erasure. Our analysis claims, on the other hand, that erasure may apply not only to deleted formal features but also to undeleted formal features.

- (34) F_1 enters into an erasure relation with F_2 iff F_2 is in the erasure domain of H and F_1 is erased. (F_2 may also be erased.)

Given that erasure only takes place under identity, F_1 erases only if identical with F_2 . Note in passing that exactly like deletion, erasure is subject to the principle of recoverability, since it makes formal features inaccessible to a computation and thus invisible at LF. Only UFFs therefore may enter into erasure relations. F_1 in the definition of an erasure relation (34) must be a UFF.

The erasure domain of H consists of:

- (35) a. X adjoined to H and any features dominated by X
 b. Any X^{\max} in the Spec of H and any features dominated by X
 c. Any X^{\max} in the complement of H and any features dominated by X

Note that our notion of erasure domain corresponds to the notion of checking domain in the standard Attract/Move approach. Recall that according to the standard Attract/Move approach, the notion of checking domain is responsible for the fact that a raised term (or a copied term under the Copy theory of movement) with a feature F_2 is merged into the local domain of H with F_1 , where F_1 is identical with F_2 . Since our Copy + Merge approach does not assume the checking operation, however, the notion of checking domain is no longer available. We therefore have to look for an alternative way of expressing the local relation which holds between F_1 and F_2 . The notion of erasure domain plays that role, ensuring that F_1 of H is only erased when F_2 is in the local domain of H

through application of Merge. Under our Copy + Merge approach, therefore, the local requirement which is assumed to be imposed on the checking operation in terms of the notion of checking domain in the standard Attract/Move approach is imposed on the erasure operation in terms of the notion of erasure domain. I will later argue that erasure captures the effects which have traditionally been accounted for by agreement and assignment. It has been observed by, among others, Chomsky (1981) that agreement and assignment only take place under a local relation. Hence, it is plausible to claim that erasure is subject to the local requirement.

Let us illustrate how our analysis works, taking (31) as an example. Suppose that X dominates F_1 , a UFF. Suppose also that Z^{\max} dominates F_2 , which is identical with F_1 . Then, since F_1 enters into a deletion relation with F_2 , Z^{\max} undergoes the copy operation given that Z^{\max} is the minimal element including F_2 that allows for convergence. After the copy operation, F_1 deletes and thus becomes invisible at LF. If F_2 is uninterpretable, it is also deleted. If F_2 is interpretable, it remains intact. This yields (31a), where there are two independent syntactic objects, X^{\max} and Z^{\max} . Recall that deleted features, though not visible at LF, are still accessible to a computation. Since UFFs, even if deleted, are required to be erased before Spell-Out, F_1 must enter into an erasure relation with F_2 through merging Z^{\max} into the Spec of X , an erasure domain of X . F_1 is erased and made inaccessible to the computation. If F_2 is uninterpretable, it is also erased. If F_2 is interpretable, it remains intact. This yields (31b). Hence, Copy Z^{\max} is followed by Merge Z^{\max} into the Spec of X , which correctly brings about the effects of the traditional notion of Attract/Move.

This gives validity to our claim that Copy takes place under a deletion relation. As shown above, a UFF serves as a trigger for operations which make an element with the matching feature appear in its local domain. When it discharges its triggering function, it is eliminated (made inaccessible to a computation) through erasure. Suppose that we come across a UFF during a derivation. If there is an element with the matching feature in the N, the UFF requires that element to be merged in its local domain. Since its triggering function is discharged by this merge operation, it is eliminated through erasure. If there is an element with the matching feature not in the N but in a syntactic object already formed, on the other hand, the UFF must first trigger Copy of that element. This is because unless that element undergoes Copy, the UFF can never be eliminated through merger of that element in its local domain. The application of Copy only partially discharges the triggering function of the UFF. It is then natural to assume that the UFF does not get eliminated through erasure by the application of Copy, since it does not completely discharge its triggering function. Rather, it only becomes invisible at LF through deletion, a weaker form of erasure. Hence, it is plausible to claim that Copy takes place under a deletion relation

The salient property of our Copy + Merge approach is that it only makes use of the operations of deletion and erasure, entirely dispensing with the operation of checking.⁹ The checking operation was originally proposed by Chomsky (1993), which first set up the MP. Recall that Chomsky (1993) assumes that lexical items are fully inflected in the

⁹See Collins (1997) for a different way of dispensing with the operation of checking.

lexicon. The addition of morphological features to a lexical item in the lexicon involves the simultaneous addition of those features, which play no role at PF or LF. Those features, which are visible at the interface levels, must be eliminated by checking operations prior to PF and LF for convergence. For example, the verb *V* is taken to be a sequence $V = (\alpha, \text{Infl}_1, \dots, \text{Infl}_n)$ where α is the morphological complex $[\text{R-Infl}_1-\dots-\text{Infl}_n]$, *R* a root and Infl_i an inflectional feature. These inflectional features must be eliminated by the raising of *V* to the functional categories with the matching features for convergence.

Chomsky (1995) argues, however, that the operation of checking has odd features and thus should be dispensed with. First, the operation of checking is redundant in that the relevant properties can be determined by algorithms at LF. The LF representation which contains an inflected item, say $[\text{V-AGR-TENSE}]$, tells us that checking operations have applied to this item, eliminating the inflectional features *AGR* and *TENSE*. Such redundant operations should be abandoned because of the notion of "simplicity." Second, the operation of checking is incompatible with the most recent development of the MP where formal features are classified based on $[\pm \text{Interpretable}]$. Recall that among formal features, only uninterpretable features but not interpretable ones may be made invisible at LF due to the principle of recoverability. According to the operation of checking proposed by Chomsky (1993), however, checked features must be eliminated irrespectively of whether they are interpretable or not. Even interpretable features would be eliminated by checking operations. This would violate the principle of recoverability. The operation of checking is therefore incompatible with the most recent development of the MP unless we change its original formulation. Hence,

our Copy + Merge approach, where no operation of checking is posited, is theoretically desirable.¹⁰

Since there is no operation of checking under our Copy + Merge approach, the ICP, which presupposes the operation of checking, must be reformulated accordingly. In the next subsection, I will argue that the ICP should be reformulated as a condition on LF visibility, which requires that UFFs should be made invisible at LF immediately when they become accessible to a computation.

3.2.4.2 Elimination of the Operation of Checking

Since no checking operation is available any longer, the ICP, which presupposes the existence of the operation of checking, should be reformulated accordingly. I argue that the ICP should be reformulated as a condition on LF visibility:

¹⁰Chomsky (1995) claims that his system dispenses with the operation of checking. Contrary to his claim, however, the operation of checking is not entirely dispensed with under his system. Let us consider his definition of the last resort condition, which is to be incorporated into the definition of Attract/Move:

- (i) Last Resort
Attract/Move F raises F to target K only if F enters into a checking relation with a sublabel of K.

(adapted from Chomsky 1995:280)

Checked features are subject to deletion operations and deleted features are further subject to erasure operations:

- (ii) a. A checked feature is deleted when possible.
b. A deleted feature is erased when possible

(Chomsky 1995:280)

Note that his system crucially makes use of the distinction between checked and unchecked features, claiming that only checked features are subject to deletion. Hence, his system still assumes the operational content of checking, though none of his definitions directly make reference to the operation of checking.

(36) The Immediate Checking Principle (Revised)

Uninterpretable formal feature (UFFs) must be made invisible at LF immediately when they become accessible to a computation.

(36) requires that when we come across a UFF during a derivation, we must immediately make it invisible at LF either by deletion or erasure. Through the application of Copy, UFFs are made invisible at LF by deletion. They are made invisible at LF by erasure through being merged into the erasure domain of H, a head, with the matching features. Note that (36) subsumes the last resort condition on Copy (32) and thus makes it possible to maintain the simplest definition of Copy, which only states that Copy creates a new independent syntactic object by copying a term within a syntactic object.

3.2.4.3 Erasure as Conversion of UFFs to Phonological Properties

The above discussion has shown that Copy only takes place when a deletion relation is established. Deleted UFFs are further required to enter into erasure relations before Spell-Out. Since erasure only takes place within an erasure domain, copied terms are required to be merged into that domain. This correctly brings about the effects of Attract/Move. A question now arises what requires deleted UFFs to enter into erasure relations before Spell-Out. I propose that erasure should be interpreted as conversion of UFFs to phonological properties, arguing that erasure of UFFs is required for PF-convergence.¹¹

¹¹See Chomsky (1995) for a similar interpretation of UFFs.

It is well known that there are cases where UFFs have their phonetic reflexes. The ϕ -features of predicates, which are UFFs, often have their phonetic reflexes. For example, in rich subject-agreement languages like Italian and Spanish, the subject ϕ -features of predicates have their phonetic reflexes. Presented below is the conjugation of the present tense indicative of the Spanish verb *hablar* 'to say':

(37)	a.	habl-o	'I speak'	1s
	b.	habl-as	'you speak'	2s
	c.	habl-a	'he/she speaks'	3s
	d.	habl-amos	'we speak'	1pl
	e.	habl-áis	'you (pl.) speak'	2pl
	f.	habl-an	'they speak'	3pl

The phonetic realizations of the subject ϕ -features of predicates are also observed in languages like English, French, German, and Irish though they have relatively impoverished agreement morphology. Similarly, the object ϕ -features of predicates are phonetically realized in, among others, the participial verbs with wh-movement in Romance languages (Kayne (1989)) and the perfect aspect sentences in Pashto (Huang (1984, 1989)).

The Case features of nominals, being UFFs, also have their phonetic reflexes. The phonetic realizations of the Case features of nominals are clearly observed in languages like Latin. Presented below is the conjugation of the singular form of the Latin masculine noun *puer* 'boy':

- (38) a. Nominative: puer
 b. Accusative: puerum
 c. Genitive: pueri
 d. Dative: puero#
 e. Ablative: puero#

The phonetic realizations of the Case features of nominals are also observed in languages like English and French though they have relatively impoverished Case morphology.

In order to account for the phonetic realizations of these UFFs, I argue that erasure should be interpreted as conversion of UFFs to their corresponding phonological properties, which are to be stripped away at Spell-Out. I argue that this conversion takes place before Spell-Out and thus the phonetic reflexes of UFFs appear at PF. The uninterpretable ϕ -features of predicates are converted to their corresponding phonological properties through entering into erasure relations with the ϕ -features of nominals. The Case features of nominals are converted into their corresponding phonological properties through entering into erasure relations with the Case features of verbs. It is well known that languages vary as to whether the ϕ -features of predicates and the Case features of nominals are phonetically realized. Even in languages where they are not phonetically realized, I argue that these UFFs are converted to their corresponding phonological properties, presumably phonetically null, through entering into erasure relations.

Such a conversion is required to take place before Spell-Out by the fact that lexical items with these features have morphological slots to be filled by their phonetic reflexes. Predicates have the morphological slots for ϕ -features which are to be filled by the phonetic reflexes of their

uninterpretable ϕ -features. Nominals have the morphological slots for Case features which are to be filled by the phonetic reflexes of their uninterpretable Case features. The information that those lexical items have such morphological slots belongs to their intrinsic property, though what kinds of features are to fill the slots does not. If the slots are not filled, it leads a derivation to crash at PF (or possibly it is ruled out in the morphological component, which stays between Spell-Out and the PF interface).¹²

There are, however, UFFs which never have their phonetic reflexes like the uninterpretable categorial features of functional heads including the Q-feature of C, the thematic features of predicates, and the Case features of verbs. The difference between the UFFs with phonetic reflexes and those without them resides in the fact that while the former is an optional feature, the latter is an intrinsic feature. I argue that UFFs, whether intrinsic or optional, have their corresponding phonological properties, though *prima facie* intrinsic features are never phonetically realized.

Recall that intrinsic features are part of lexical items within the lexicon. It is then reasonable to claim that the phonological properties of intrinsic UFFs constitute part of those of lexical items themselves. If they are not properly converted to their corresponding phonological properties through conversion, lexical items end up having insufficient phonological information to be properly realized at PF. Such improperly

¹²The same can be said about the tense features of verbs, which are also optional UFFs. According to our conversion analysis, verbs are not inflectionally represented but rather bare in the lexicon. They acquire inflectional morphology through syntactic operations. See Lasnik (1994) and Rohrbacher (1993, 1995) for detailed discussion of this subject.

realized lexical items are illegitimate at PF; the derivation would crash. For example, let us consider the verb *see*. It has Accusative Case and Theme features as its intrinsic UFFs. Although the verb *see* is assigned phonological properties in the lexicon, those properties alone are not sufficient to make this verb properly realized at PF. Its intrinsic UFFs must be converted to their corresponding phonological properties through erasure before Spell-Out. The converted phonological properties together with the lexically assigned phonological properties make this verb properly realized at PF. Otherwise, an illegitimate object appears at PF; the derivation crashes at that level. Hence, intrinsic UFFs must also be converted to their corresponding phonological properties through erasure before Spell-Out, though they do not have their own phonetic reflexes. Optional features, on the other hand, are not assigned to lexical items within the lexicon but rather added to them arbitrarily at any point of derivations. It is then reasonable to claim that the phonological properties of optional UFFs do not constitute part of those of lexical items within the lexicon. They are rather added to the lexical items during derivations through conversion. In other words, they constitute inflectional morphemes. Hence, optional UFFs have their phonetic reflexes of their own, which are to be added to the lexical items as their inflectional morphemes, though they may be phonetically null. We can therefore conclude that UFFs, whether intrinsic or optional, must be converted to their corresponding phonological properties through erasure for PF-convergence.¹³

¹³This interpretation of erasure differs from Chomsky's (1995) interpretation in that the latter, but not the former, claims that erasure is prohibited if it would create an illegitimate object. Crucial is the case where a term only consists of a strong feature

With the above discussion in mind, let us reconsider how our Copy + Merge approach works. Suppose that we come across a UFF of H, a head, during a derivation. The UFF is required to be immediately made invisible at LF by the ICP. If it is made invisible at LF by erasure through merger of a term with the matching feature into the erasure domain of H, it undergoes conversion to its corresponding phonological property and thus becomes inaccessible to the computation anymore. If it is made invisible at LF by deletion through the application of Copy, on the other hand, it is further subject to erasure. The copied term, which has the matching feature, must be merged into the erasure domain of H. This is because the UFF of H must be converted to its corresponding phonological property for PF-convergence. This ensures that the copied term is merged into the erasure domain of H before Spell-Out.^{14, 15}

and thus the erasure of the strong feature leads to erasure of the entire term. Chomsky's interpretation of erasure depends on the validity of his analysis of pure expletives. He claims that pure expletives like English *it* only consist of strong D-features. Hence, their strong feature may never be erased. It is not clear at this point, however, whether his analysis of pure expletives is on the right track. The present discussion therefore claims that erasure of a UFF is always mandatory, leaving an analysis of pure expletives for further study,

¹⁴One might wonder what ensures that subjects are "docked" in the right positions in complex sentences like (i):

- (i) John said that Bill read that book

Note that *John* and *Bill* have the same agreement and Case features. Before merger of *John* and *Bill* with the main structure, we have the following:

- (ii) a. $[T^{\max} T [V^{\max} \text{John said } [C^{\max} \text{that } [T^{\max} T [V^{\max} \text{Bill read that book }]]]]]$
 b. John
 c. Bill

If we "docked" *John* and *Mary* to the embedded and matrix Spec's of T^{\max} , respectively, we would get the following:

- (iv) $[T^{\max} \text{Bill } [T [V^{\max} \text{John said } [C^{\max} \text{that } [T^{\max} \text{John } [T [V^{\max} \text{Bill read that book }]]]]]]]$

Chain formation applies in the LF-component, yielding the following two chains:

- (v) a. CH = (*Bill*, *Bill*)
 b. CH = (*John*, *John*)

These chains, however, are not legitimate at LF. Specifically, these chains do not count as A-chains, since C^{\max} intervenes between the head and tail positions. The resultant representation would be ruled out. Note in passing that (vb) also violates the condition on an A-chain which states that the head of an A-chain is in a Case-marked position while its tail is in a θ -position. Hence, we can ensure that the subjects *John* and *Bill* are "docked" into the right positions. The cases where there are two instances of object raising within a sentence can be analyzed in a similar way.

One might also wonder what ensures that *wh*-phrases are "docked" into the right positions in examples like (i):

- (i) who wonders what John bought

Before merger of *who* and *what* with the main structure, we have the structures in (ii). Here, we assume just for expository purposes that the subjects have already been merged with the main structure:

- (ii) a. $[C^{\max} C [T^{\max} \text{who} [T [V^{\max} \text{who wonders} [C^{\max} C [T^{\max} \text{John} [T [V^{\max} \text{John bought what}]]]]]]]]]$
 b. who
 c. what

If we "dock" *who* and *what* to the embedded and matrix Spec's of C^{\max} , respectively, we would get the following:

- (iii) $[C^{\max} \text{what} [C [T^{\max} \text{who} [T [V^{\max} \text{who wonders} [C^{\max} \text{who} [C [T^{\max} \text{John} [T [V^{\max} \text{John bought what}]]]]]]]]]]]$

Among the chains created in this derivation, the following chains are relevant here:

- (iv) a. CH = (*what*, *what*)
 b. CH = (*who*, *who*)

Let us assume following Chomsky (1993) that the non-head positions of chains delete. This yields the following structure:

- (v) $[C^{\max} \mathbf{what}_i [C [T^{\max} \mathbf{t}'_j [T [V^{\max} \mathbf{t}_j \text{wonders} [C^{\max} \mathbf{who}_j [C [T^{\max} \text{John} [T [V^{\max} \text{John bought } \mathbf{t}_i]]]]]]]]]]$

In (v), the operator *who* does not have any variable to bind in its c-command domain. This violates the ban against vacuous quantification. Furthermore, the variable t'_j , more precisely the A-chain (t'_j , t_j), does not have any c-commanding operator. Hence, this violates the ban against free variables. Hence, we can ensure that the *wh*-phrases *who* and *what* are "docked" into the right positions.

¹⁵One could argue that this analysis is conceptually undesirable since it necessarily needs global considerations. Merger of a copied term into an erasure domain of a head H is required for PF-convergence. We have to look at the PF interface level to determine whether to apply Merge to a copied term. The reference to the PF interface level has the "look-ahead" property, which should be avoided on conceptual grounds.

It might be possible to claim, however, that the problem of globality does not arise in this analysis. Recall that the notion of globality is defined as below:

- (i) A condition C is global if it cannot determine whether to apply an OP or not only on the basis of information available in Σ .

Note that according to the definition of the notion of globality (i), the problem of globality only arises when we have an option of applying an OP or not during a derivation. If we follow Chomsky (1995) in assuming that Merge counts as costless and thus does not need any motivation for its application, there is no option of not applying Merge to a copied term. Insufficient application of Merge simply generates no derivation. Hence, concerning Merge, the problem of globality never arises if Chomsky's interpretation of

Through the rest of this thesis, unless the distinction between the standard Attract/Move approach and our Copy + Merge approach is relevant, we will pretend for expository purposes that we are assuming the former. Given the established usage, the discussion to follow also sticks to the phrase "checking operation," which is meant to represent a process of making UFFs invisible at LF through either a deletion or erasure operation unless any complication arises. Accordingly, we will stick to the appellation "the Immediate Checking Principle (ICP)," though, to be precise, it does not refer to the operation of checking.

3.2.4.4 An Account of the Subject Condition

Returning to the Subject Condition, let us consider (2) (repeated here as (39)) as an example:

(39) ***who** did [pictures of *t*] please you

Let us first consider how to construct the subject phrase *pictures of who*. We first select *of*. The selectional restriction feature of *of*, being uninterpretable, must be made invisible at LF immediately by erasure through merger of *of* with *who*. Then, the noun *pictures* is selected. The ICP requires that its selectional restriction feature, being uninterpretable, should be made invisible at LF immediately by erasure through merger of *pictures* with *of who*. Finally, D is selected. Its selectional restriction feature is made invisible immediately at LF by erasure through combining D with *pictures of who*. The resultant structure is as follows:

Merge is correct. If this conjecture is correct, we can say that the inspection of PF only tells us where a copied term should be merged. It does not tell us whether we should apply merger of a copied term or not.

(40) $[D^{\max} [N^{\max} \text{ pictures } [P^{\max} \text{ of } [D^{\max} \text{ who}_{[Q]}]]]]]$

Let us now turn to consider how to construct the main structure.

We first select the verb *please*. The ICP requires that we should immediately make the internal thematic feature and Accusative Case feature of *please* invisible at LF, since these features are uninterpretable. Hence, we merge *please* with *you*, erasing these features. At this point, the external thematic feature of *please* becomes accessible to the computation. Since it is uninterpretable, it must be made invisible at LF immediately by erasure through combining *pictures of who* with *please you*. The resultant structure is as below:

(41) $[V^{\max} [D^{\max} \text{ pictures of who}_{[Q]}] [\text{please you}]]]$

The next step is to select T. The ICP requires that the selectional restriction feature of T, which states that T takes V^{\max} as its complement, should be made invisible at LF immediately by erasure through combining T with the V^{\max} *pictures of who please you*, as shown below:

(42) $[T^{\max} T [V^{\max} [D^{\max} \text{ pictures of who}_{[Q]}] [\text{please you}]]]]]$

Since the strong D-feature and Nominative Case feature of T become accessible to the computation at this stage, the ICP requires that they should be made invisible at LF immediately. They are made invisible at LF by deletion through application of Copy to the D^{\max} *pictures of who* given that the D^{\max} is the minimal element including these features that allows for convergence:¹⁶

¹⁶I will later explicate what counts as a minimal element including a feature F that allows for convergence.

- (43) a. $[T^{\max} T [V^{\max} [D^{\max} \text{ pictures of who}_{[Q]}] [\text{please you}]]]$
 b. $[D^{\max} \text{ pictures of who}_{[Q]}]$

The newly created D^{\max} (43b) is eventually merged into the Spec of T^{\max} , an erasure domain of T, for PF-convergence. It is important to note, however, that merger of the newly created D^{\max} (43b) with the main structure (43a) cannot take place at this point of the derivation. Recall that this merger is not required by the ICP, since the UFFs of T have already been made invisible at LF by deletion through application of Copy. We can say that the newly created D^{\max} functions as an adjunct in the sense that it is required to be merged postcyclically. The EP then requires that we should select C rather than merge the newly created D^{\max} with the main structure at this stage. When C is selected, the ICP requires that its selectional restriction feature, which states that it takes T^{\max} as its complement, should be made invisible at LF immediately by erasure through combining C with the T^{\max} *pictures of who please you*:

- (44) a. $[C^{\max} C_{[Q]} [T^{\max} T [V^{\max} [D^{\max} \text{ pictures of who}_{[Q]}] [\text{please you}]]]]]$
 b. $[D^{\max} \text{ pictures of who}_{[Q]}]$

At this stage of the derivation, the strong Q-feature of C, which is a UFF, becomes accessible to the computation. The ICP requires that it should be made invisible at LF immediately. It should be noted that the strong Q-feature of C cannot be made invisible at LF by deletion through application of Copy to *who* in the D^{\max} (44b). The *wh*-phrase *who* is the potential candidate for being subject to the copy operation triggered by the strong Q-feature of C. The *wh*-phrase *who* in the D^{\max} (44b),

however, is not c-commanded by C. This is because the D^{\max} (44b) and the main structure (44a) each constitute an independent syntactic object at this stage of the derivation. Since C does not c-command *who* in the D^{\max} (44b), there is no way for the strong Q-feature of C to trigger Copy of that *wh*-phrase. The strong Q-feature of C cannot be made invisible at LF through application of Copy to *who* in the D^{\max} (44b), which is eventually merged into the Spec of T^{\max} . Hence, we can derive the fact that no extraction is possible out of the subject in the Spec of T^{\max} .

There is, however, an alternative candidate for making the strong Q-feature of C invisible at LF. Since we are assuming that Attract/Move consists of Copy and Merge, there exist two instances of the D^{\max} *pictures of who* after application of Copy. The one is in the Spec of V^{\max} while the other is the D^{\max} (44b). Although the strong Q-feature of C cannot trigger Copy of *who* in the D^{\max} (44b), it can trigger Copy of *who* within the Spec of V^{\max} . This is because C c-commands *who* within the Spec of V^{\max} . Then, the resulting structure would be as follows after the strong Q-feature of C is made invisible at LF by deletion through application of Copy:

- (45) a. $[C^{\max} C [T^{\max} T [V^{\max} [D^{\max} \text{pictures of } \mathbf{who}_{[Q]}]$
 [please you]]]]
- b. $[D^{\max} \text{pictures of } \mathbf{who}_{[Q]}]$
- c. $\mathbf{who}_{[Q]}$

(45c) is the newly created *who*. Note that (45c) is the copy of *who* in (45a) but not that of *who* in (45b). The newly created *who* (45c) is not merged with the main structure (45a) at this point. This stage of the

derivation, therefore, can be characterized as consisting of the three independent syntactic objects (45a-c).

At the later stage of the derivation, (45b-c) are merged into the right positions in (45a) for PF-convergence. This yields the following structure, with all formal features including Q-features being ignored:

$$(46) \quad [C^{\max} \text{ who } [C [T^{\max} [D^{\max} \text{ pictures of who } [T [V^{\max} [D^{\max} \text{ pictures of who } [\text{please you}]]]]]]]$$

Although this derivation reaches LF without being canceled, I argue that there is a violation of the ban against vacuous quantification, a condition of FI at LF.

Recall that the operation that copies α introduces α a second time into a syntactic object. The element α appears twice in the syntactic object, in its initial and raised positions. This is the only case in which two terms can be identical, since we distinguish among distinct selections of a single lexical item from the lexicon. This ensures that such pairs that consist of identical terms only arise through Copy. Although the two terms are identical in constitution, they are positionally distinct. Recall also that such terms that are identical in constitution but positionally distinct from each other form a chain in the LF-component. In (46), *pictures of who* is introduced a second time in the syntactic object through Copy. The two occurrences of *pictures of who* are therefore identical in constitution but positionally distinct from each other. Furthermore, *who* in the Spec of C^{\max} is introduced by copying *who* within the Spec of V^{\max} . These two instances of *who* are also identical in constitution but positionally distinct. Hence, the following two chains are formed during the derivation:

- (47) a. CH = (*pictures of who*, *pictures of who*)
 b. CH = (*who*, *who*)

Let me stress again that chain (47b) consists of *who* in the Spec of C^{\max} and *who* within the Spec of V^{\max} , not *who* within the Spec of T^{\max} .

Following Chomsky (1993), let us assume that the non-head positions of chains delete within the LF component. Hence, *who* within the Spec of V^{\max} and *pictures of who* in the Spec of V^{\max} delete. (46) therefore yields the following LF-representation:^{17, 18}

- (48) [C^{\max} **who** [C [T^{\max} [D^{\max} pictures of who]_i [T [V^{\max} t_i [please you]]]]]]]

In (48), the operator *who* does not bind any variable and thus violates the ban against vacuous quantification, a condition of the FI. Hence, this derivation crashes at LF.¹⁹

It is possible to claim that (48) also violates the condition of inclusiveness. Recall that the condition of inclusiveness requires that any structure constructed by a computation from N to LF only consists of elements which are already present in the lexical items selected for the N. This condition holds with the single exception of the terms introduced by Copy. Copy introduces elements which are not present in an N during a derivation. Those introduced by Copy, however, form chains in the LF-component. The non-head positions of the chain are deleted and only the

¹⁷See chapter 6 for detailed discussion of the construction of an operator-variable pair.

¹⁸Just for expository purposes, we use indices to identify elements of a chain here and in the relevant representations to follow. Recall that the condition of inclusiveness does not allow indices to be assigned to any structure constructed by a computation from N to LF.

¹⁹See the discussion of the previous chapter for the alternative view that the ban against vacuous quantification does not lead a derivation to crash but only makes an interpretation anomalous. Note that the arguments to follow are valid under either interpretation of the ban against vacuous quantification.

head-position remains. At LF, therefore, the condition of inclusiveness should hold without any exceptions. No new objects are allowed to appear at LF. In LF-representation (48), however, there are two occurrences of *who*, although the N of this derivation only includes one occurrence of *who*. Hence, this derivation adds the new object which is not present in the N; this violates the condition of inclusiveness.

The above discussion showed that the ICP and the EP together with the Copy + Merge approach correctly derives the subject condition effects. Before closing this subsection, let us look at how the present analysis deals with the case where a subject itself is a *wh*-phrase:

(49) **who** *t* bought that book

During the derivation, we come to the following stage:

(50) [_T^{max} T [_V^{max} **who**_[Q] [buy [_D^{max} that book]]]]

At this stage, the ICP requires that the strong D-feature and Nominative Case feature of T should be made invisible at LF immediately by deletion through application of Copy to *who*. The resultant structure is as follows:

(51) a. [_T^{max} T [_V^{max} **who**_[Q] [buy [_D^{max} that book]]]]
 b. **who**_[Q]

Note that, as argued above, we cannot merge the newly created *who* (51b) with the main structure (51a) at this point of derivation, since the EP requires that C should be selected.

When C is selected, the ICP requires that its selectional restriction feature should be made invisible at LF immediately by erasure through merger of C and the T^{max} *who buy that book*:

- (52) a. [C^{max} C_[Q] [T^{max} T [V^{max} who_[Q] [buy [D^{max} that
book]]]]]]
b. who_[Q]

At this stage of the derivation, the strong Q-feature of C becomes accessible to the computation. Since it is uninterpretable, it must be made invisible at LF immediately by deletion through copying *who* in the Spec of V^{max}. Note that the strong Q-feature of C cannot trigger Copy of the newly created *who* (52b), since the former does not c-command the latter. The resultant structure is as follows:

- (53) a. [C^{max} C [T^{max} T [V^{max} **who**_[Q] [buy [D^{max} that book]]]]]]
b. who_[Q]
c. **who**_[Q]

Finally, the two instances of *who* are merged with the main structure as the Spec of C^{max}, an erasure domain of C, and the Spec of T^{max}, an erasure domain of T, respectively. The resultant structure is as follows, with all formal features including Q-features being ignored:

- (54) [C^{max} **who** [C [T^{max} **who** [T [V^{max} who [buy [D^{max} that
book]]]]]]]]

In (54), since *who* in the Spec of T^{max} is introduced by copying *who* in the Spec of V^{max}, these two occurrences of *who*, though positionally distinct, are identical in constitution. Furthermore, *who* in the Spec of C^{max} is also introduced by copying *who* in the Spec of V^{max}. These two occurrences of *who* are also identical in constitution but positionally distinct. On the assumption that the identity relation is transitive, the three occurrences of *who* in (54) are all identical in constitution, though they are positionally distinct. Hence, *who* in the Spec of C^{max} is

identical in constitution with *who* in the Spec of T^{\max} , which in turn is identical in constitution with *who* in the Spec of V^{\max} . Chain formation applies to (54), yielding the following two chains:^{20, 21}

- (55) a. CH = (*who*, *who*)
 b. CH = (*who*, *who*)

(55) is an operator-variable construction which consists of *who* in the Spec of C^{\max} and *who* in the Spec of T^{\max} . (55), on the other hand, is an A-chain which consists of *who* in the Spec of T^{\max} and *who* in the Spec of V^{\max} . Then, the non-head members of the chains, i.e., *who* in the Spec of T^{\max} and *who* in the Spec of V^{\max} , are deleted. This yields the following LF-representation:

- (56) [C^{\max} **who** [C [T^{\max} *t'* [T [V^{\max} *t* [buy [D^{\max} that book]]]]]]]]

LF-representation (56) only consists of legitimate objects. Specifically, there is no violation of the ban against vacuous quantification, since *who* properly binds *t'* in the Spec of T^{\max} (more precisely, the A-chain (*t'*, *t*)) as its variable.

²⁰Note that this is different from successive cyclic A'-movement from an A-position, which is discussed in Chomsky (1991). The latter yields a chain that is not legitimate, a "heterogeneous chain" which consists of an adjunct chain and an A'-A pair. Chomsky argues that this "heterogeneous chain" must be made a legitimate object, an operator-variable construction, by eliminating the intermediate A'-trace(s) in accordance with the economy condition. In the present case, however, we can construct legitimate objects, an operator-variable chain and an A-chain, without eliminating any intermediate trace. Since the intermediate trace in the Spec of T^{\max} need not be deleted, it may not be deleted by the "least effort" principle.

²¹Alternatively, as argued by Chomsky and Lasnik (1993), chain formation applies to (54), yielding the following three-membered chain:

- (i) CH = (*who*, *who*, *who*)

Even under this analysis, we get LF-representation (56) after the deletion of the non-head members of the chain.

3.2.4.5 Lack of the Subject Condition Effects in Japanese

It has been observed by, among others, Fukui (1995), Kayne (1984), Kuno (1973), Lasnik and Saito (1992), Ross (1967), Saito (1985, 1992), and Takahashi (1994) that Japanese does not obey the Subject Condition.

Let us look at the following examples:

(57) a. The Cleft Construction

[**Op**_i [John-ga [[Mary-ga **t**_i katta no]-ga
 -Nom -Nom bought fact-Nom
 mondai da to] omotte iru] no]-wa
 problem is Comp think Comp-Top
 sono hon_i-o da
 that book-Acc be

Lit. 'it is that book_i that John thinks that the fact
 that Mary bought *e*_i is a problem'

b. The *Tough* Construction

[[zibun-no ootoo-kara]_i-ga [(John_j-nitotte) [**Op**_i [[[*pro*_j
 self's brother-from-Nom -for
t_i okane-o takusan karite iru no]-ga mondai da]
 money-Acc a lot borro fact-Nom problem be
 to] mitome] nikui]]
 Comp admit hard

Lit. '[from self's brother]_i is hard (for John_j) to admit
 that the fact that he_j has borrowed a lot of money *e*_i is
 a problem'

c. The Comparative Deletion Construction

[*Op_i* [John-ga [Bill-ga *t_i* katta no]-ga mondai
 -Nom -Nom bought fact-Nom problem
 da to] itte iru] yorimo] Mary-wa takusan hon-o
 be Comp say than -Top many book-Acc
 katta
 bought

Lit. 'Mary bought more books than John says that
 the fact that Bill bought is a problem'

As will be discussed extensively in the next chapter, the cleft construction with an NP-Case or PP focus, the *tough* construction with a PP subject, and the comparative deletion construction all involve movement of an empty operator (see, among others, Hoji (1990), Ishii (1991), Kikuchi (1987), and Takezawa (1987)). In (57), although empty operators are extracted out of the subject phrases, the results are acceptable. This indicates that the Subject Condition does not hold in Japanese.

There are some speakers who find the examples in (57) awkward. It is possible to claim, however, that their awkwardness is due to the fact an empty operator is extracted out of a *no* clause. For those speakers, extraction out of an object *no* clause is also awkward, as shown below:

(58) a. The Cleft Construction

[*Op_i* [John-ga [Mary-ga *t_i* katta no]-o
 -Nom -Nom bought fact-Acc
 mondai-ni site iru] no]-wa sono hon_i-o da
 problem-Dat make Comp-Top that book-Acc be

Lit. 'it is that book_i that John is calling the fact that
 Mary bought *e_i* into question'

b. The *Tough* Construction[[zibun-no ootoo-kara]_i-ga [(John_j-nitotte) [**Op**_i [*pro*_j

self's brother-from-Nom -for

[Mary-ga *t_i* okane-o takusan karite iru no]-o

-Nom money-Acc a lot borrow fact-Acc

mondai-ni si]] nikui]]

problem-Dat make hard

Lit. '[from self's brother]_i is hard (for John_j) to call the fact that Mary has borrowed a lot of money *e_i* into question'

c. The Comparative Deletion Construction

[**Op**_i [John-ga [Bill-ga *t_i* katta no]-o

-Nom -Nom bought fact-Acc

mondai-ni site iru] yorimo] Mary-wa takusan

problem-Dat make than -Top many

hon-o katta

book-Acc bought

Lit. 'Mary bought more books than John calls the fact that Bill bought into question'

Hence, even for those speakers who find the examples in (57) awkward, there is no contrast in acceptability between the examples in (57) and those in (58). This suggests that there are no subject condition effects in Japanese.

I argue that the lack of the subject condition effects in Japanese straightforwardly follows from our analysis if we assume following, among others, Fukui (1986), Kuroda (1988), Lasnik and Saito (1992), and Takahashi (1994) that subjects in Japanese stay in-situ at least in the

overt component. Recall that subjects in English overtly move from within V^{\max} to the Spec of T^{\max} . Under our analysis, this overt raising yields the adjuncthood of a subject in the sense that the raised subject is required to be merged with the main structure postcyclically. The subject in English therefore constitutes "domain barriers." Given that the subject in Japanese never undergoes movement, on the other hand, our analysis correctly predicts that the Subject Condition does not hold in Japanese. This is because unmoved subjects count as arguments throughout derivations, being required to be merged cyclically.²²

Our account of the lack of the subject condition effects in Japanese can be extended to an analysis of extraction from some postverbal subjects in Italian. It has been claimed by, among others, Belletti and Rizzi (1981), Jaeggli (1981), and Rizzi (1982) that some postverbal subjects behave like objects with respect to *ne*-cliticization:

- (59) a. ho letto tre libri
'I have read three books'
b. **ne** ho letto tre
'I have read three'
- (60) a. sono arrivati tre ragazzi
'three kids have arrived'
b. **ne** sono arrivati tre
'three have arrived'

(Jaeggli 1981:139-140)

²²It should be pointed out that our analysis can also accommodate the lack of the Subject Condition in Japanese under the overt subject raising analysis if subject raising in Japanese is a non-feature-driven movement like scrambling. As will be discussed in detail in chapter 4, our analysis claims that phrases which have undergone non-feature-driven movement do not constitute "barriers" for further extraction.

Ne-cliticization takes place from within the direct object in (59b) and the postverbal subject in (60b). They are both acceptable. This indicates that the postverbal subject in (60) does not constitute a "domain barrier." I assume following, among others, Jaeggli (1981) that such a postverbal subject originates there and thus never undergoes movement. It then follows from our analysis that such a postverbal subject that does not undergo movement never constitutes a "domain barrier." Hence, we can correctly predict that examples like (60b) are acceptable.

3.2.5 Summary

To summarize section 3.2, I have argued that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure, taking English overt *wh*-movement as an example. It was shown that the "domain barriers" are exactly those elements which are forced to be merged postcyclically. When we come to the stage of a derivation where the strong *Q*-feature of *C* is to be made invisible at LF, the "domain barrier" has not been merged with the main structure. The "domain barrier" and the main structure each constitute an independent syntactic object at this stage. No element inside the "domain barrier" may undergo Copy to make the strong *Q*-feature of *C* invisible at LF. Since the strong *Q*-feature cannot be made invisible at LF immediately, this violates the ICP; the derivation is canceled. This analysis gives us a minimalist account of the "domain barrier" effects.

3.3 No Extraction out of Moved Phrases

The analysis of the "domain barriers" proposed in the previous section can be extended to another locality condition on feature-driven *A'*-

movement, i.e., no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. I will argue that this locality condition can be accounted for along the line of the above mentioned analysis of the Subject Condition.

Let us consider the following examples (cf. Lasnik & Saito (1992)):²³

- (61) Topicalization out of Topic
 *?**vowel harmony**_i, I think that [[articles about **t_i**]_j [you read **t_j**]]
- (62) Wh-movement out of moved *wh*-phrases
 *?**who**_i do you wonder [[which pictures of **t_i**]_j [Mary bought **t_j**]]
- (63) Wh-movement out of Topic
 *?**who**_i do you think that [pictures of **t_i**]_j John wanted **t_j**

In (61), topicalization applies to *articles about vowel harmony* in the embedded clause, out of which *vowel harmony* is further extracted to the initial position of the matrix clause by topicalization. In (62), wh-movement first applies to *which pictures of who* in the embedded clause. At the matrix clause level, *who* within the embedded Spec of C^{max} is further extracted by wh-movement. In (63), topicalization applies to *pictures of who* in the embedded clause, out of which *who* is further extracted by wh-movement. These examples are all deviant. The observational generalization can be stated as follows. Phrases which have undergone feature-driven A'-movement like topicalization and wh-

²³Lasnik and Saito (1992) claims that although these examples are deviant, they are less severely deviant than the "domain barrier" violation cases. Base on such judgments, they propose that A'-binder do not function as "barriers." I depart from them in claiming that there is no substantial contrast in acceptability between these examples and the "domain barrier" violation cases.

movement constitute "barriers" for further feature-driven movement. I will argue that this generalization straightforwardly follows from the ICP coupled with the EP.^{24, 25}

Let us consider (61) as an example. Its derivation proceeds as follows. Let us consider the stage where we construct the embedded T^{\max} :

²⁴Based on the observation made by Torrego, Chomsky (1986) claims that extraction out of the *wh*-phrase in the Spec of C^{\max} is acceptable:

- (i) [de que autora]_i no sabes [[qué traducciones t_i] han ganado premios internacionales]
'by what author don't you know what translations have won international awards'

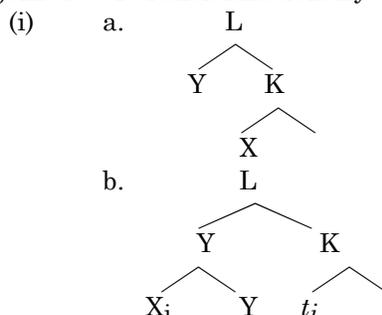
(Chomsky 1986:26)

If this observation is correct, then this counts as evidence against this locality condition. Jim Huang (personal communication) suggested to me that the contrast between (i) and (61-63) might be due to the difference in the categorial status of the extracted phrases. While the extracted phrase is a PP in the Spanish example, those in the English examples are NPs. There are cases where PP-movement but not NP-movement is allowed for some unknown reason:

- (ii) a. *These students, two of are crazy
b. Of these students, two are crazy

I leave further discussion of this subject for future research.

²⁵Note that this locality condition differs from the freezing principle advocated by, among others, Culicover (1976), Culicover and Wexler (1973a, 1973b, 1977), Culicover and Wilkins (1984), and Wexler and Culicover (1980). The freezing principle states that no transformation may apply to any node under a frozen node, where a frozen node is one which does not immediately dominate a base structure. Suppose, for instance, that X is adjoined to Y transformationally:



Suppose further that such an adjunction structure cannot be base-generated. The freezing principle claims that the node L is frozen. No transformation may apply to any node which is dominated by L.

- (64) a. [T^{max} T [V^{max} you [read [D^{max} articles about vowel
harmony]]]]
b. you

Recall that the ICP together with the EP ensures that *you*, which undergoes Copy for making the Nominative Case feature and D-feature of T invisible at LF, may not be merged with the main structure (64a) at this point of the derivation.

Let us assume that topicalization is triggered by the strong [TOPIC] feature under a functional head F (see, among others, Authier (1992) and Watanabe (1993)). Let us assume that [TOPIC] features are uninterpretable wherever they may appear, since unlike a *wh*-element, no phrase can be interpreted as a topic in-situ. F appears between C and T. In other words, F selects T^{max} and its maximal projection is selected by C. F, which has the following hierarchical structure of features, is selected at this stage:

- (65) F_{[[TOPIC] T]}

The ICP requires that the T-feature of F should be made invisible at LF immediately by erasure through combining F with T^{max}. After the T-feature of F is checked, its strong [TOPIC] feature becomes accessible to the computation. It must be made invisible at LF immediately in accordance with the ICP. This feature is made invisible at LF by deletion through copying *articles about vowel harmony*. The resultant structure is as follows:

3.4 Adjunct A'-movement and the "Domain Barrier" Effects

In the preceding sections, I have argued that the "domain barrier" effects straightforwardly follow from our theory of the composition of phrase structure. It was also shown that our theory of phrase structure accounts for the fact that no feature-driven extraction is possible out of phrases which have undergone feature-driven A'-movement. The discussion in the preceding sections, however, has confined its attention to argument A'-movement. A question now arises how to deal with locality conditions on adjunct A'-movement. In this section, I will argue that the "domain barrier" effects with adjunct A'-movement also follow from our theory of phrase structure.

3.4.1 Adjunct A'-movement

Before turning to the "domain barrier" effects with adjunct A'-movement, let us first consider how to construct simplex *wh*-adjunct questions, taking (72) as an example:

(72) why did he leave

During the derivation of (72), we construct (73), conforming to the ICP and the EP:

(73) a. [C_[Q] [T^{max} T [V^{max} he leave]]]
 b. he

Since the strong Q-feature of C is uninterpretable, the ICP requires that it should be checked immediately. Note that *why*, whose merger has not been triggered by any UFF, has not been introduced into the derivation but is still in the N. Following Bromberger (1986), Longobardi (1984), Reinhart (1982), and Rizzi (1990), I claim that *why* is the *wh*-version of a sentential adverb. Exactly like other sentential adverbs, *why* is simply

required to have the clause it modifies in its immediate c-command domain in order to be properly interpreted at LF. Hence, *why*, being a sentential adverb, is base-generated in the Spec of C^{\max} which it modifies. I therefore argue that it is direct merger of *why* into the Spec of C^{\max} that checks the strong Q-feature of C:

- (74) a. $[C^{\max} \mathbf{why}_{[Q]} [C [T^{\max} T [V^{\max} \text{he leave}]]]]$
 b. he

The strong Q-feature of C, being uninterpretable, is erased by this merger. On the other hand, the Q-feature of *why*, being interpretable, remains intact. Finally, *he* is merged into the Spec of T^{\max} , yielding (72). Under this analysis, unlike *wh*-arguments like *who* and *what*, there is no trace of *why* within the clause it modifies. In other words, *why* does not have any variable to bind. This is supported by the fact that unlike *wh*-arguments like *who* and *what*, *why* does not range over individuals but over higher order entities, yielding a function as its value (see, among others, Reinhart (1992, 1993) and Szabolsci and Zwarts (1992)).

Let us next consider how our analysis constructs complex *wh*-adjunct questions, taking as an example (75) where *why* modifies the embedded clause:

- (75) why do you think that John left

I argue that *why* is first merged into the embedded Spec of C^{\max} , where it gets an interpretation at LF, and then raised to the matrix Spec of C^{\max} .

Let us look at the derivation of (75) in detail. During the derivation, we come to the following structure:

- (76) a. $[C_{[Q]} [T^{\max} T [V^{\max} \text{John left}]]]$
 b. John

Recall that we are assuming that C is divided into two types, [+Q]-C and [-Q]-C. Between these two types of C, while [-Q]-C never has a strong Q-feature, [+Q]-C has a strong Q-feature as its intrinsic feature, though, for simplicity, we have stuck to the expression that C has a strong Q-feature as its intrinsic feature. Since the Q-feature of C is uninterpretable, it never contributes any content to LF. The interrogative/noninterrogative interpretation of a clause depends not on the existence of a Q-feature in C but on the existence of a *wh*-element in the Spec of C^{\max} at LF. A clause is interpreted as interrogative when the Spec of C^{\max} is occupied by a *wh*-element at LF. Otherwise, it is interpreted as noninterrogative. I claim that the embedded C in (76) is a [+Q]-C, which has a strong Q-feature as its intrinsic property. The strong Q-feature of the embedded C, being uninterpretable, is checked by merger of *why* with (76):

- (77) a. [C^{\max} **why**_[Q] [C [T^{\max} T [V^{\max} John left]]]]
 b. John

The strong Q-feature of C, being uninterpretable, is erased by this merger. On the other hand, the Q-feature of *why*, being interpretable, remains intact.

As the derivation proceeds, we come to the stage where the matrix C is selected and merged with the main structure. The matrix C is also a [+Q]-C and thus assigned a strong Q-feature as its intrinsic property. The strong Q-feature must be checked immediately in accordance with the ICP:

- (78) a. [C_[Q] [T^{\max} T [V^{\max} you think [C $^{\max}$ **why**_[Q] [C
 [T^{\max} T [V^{\max} John left]]]]]]]]
 b. John
 c. you

One might claim that this analysis would allow any intermediate clause of a *wh*-question to be headed by a [+Q]-C with a strong Q-feature. The proliferation of a strong Q-feature, however, is blocked by the following condition proposed by Chomsky (1995):

- (80) α enters into a numeration (N) only if it has an effect on output.

(Chomsky 1995:294)

Let us consider (81) as an example:

- (81) **what** do you think that John read *t*

In constructing the N of (81), condition (80) prevents the embedded clause from being headed by a [+Q]-C with a strong Q-feature. This is because the strong Q-feature would not have any effect on PF or LF. If the embedded C had a strong Q-feature, the *wh*-phrase *what* would undergo successive cyclic movement, first moving to the Spec of the embedded C^{\max} and then to the Spec of the matrix C^{\max} . It would not have any

Spec of C^{\max} is occupied by the scope marker *was*. The scope marker *was* also appears in every intermediate Spec of C^{\max} between the matrix Spec of C^{\max} and the *wh*-phrase.

Cheng (1997) argues that in partial *wh*-movement like (ib-c), the Q-feature of a *wh*-phrase (to be precise, a bundle of feature including the Q-feature) undergoes overt movement, leaving a copy. The Q-feature of a *wh*-phrase (strictly speaking, the bundle of formal features) and its copy are spelled out as the scope marker *was* at PF unless it is followed by category movement. In the latter case, the repair strategy in the sense of Chomsky (1995) takes place at PF between the bundle of formal feature and the category, yielding overt *wh*-movement. Hence, in (ib-c), the chain consisting of the *wh*-feature (strictly speaking, the bundle of formal features) is formed in the LF component. Given that the non-head positions of a chain delete, the *wh*-feature only remains in the head position of the chain, i.e., the matrix Spec of C^{\max} . Hence, we can correctly predict that the matrix clause, whose Spec is occupied by the *wh*-feature at LF, is interpreted as interrogative while the other clauses, whose Spec is not occupied by any *wh*-feature at LF, are interpreted as noninterrogative. According to our analysis, the intermediate clauses in (ib-c) are all headed by [+Q]-C, which ensures successive cyclic movement of a *wh*-feature. This does not violate (80), since the strong Q-feature of those [+Q]-C's has a PF effect. A question still remains, however, what forces successive cyclic movement of a *wh*-feature. I am indebted to Lisa Cheng (personal communication) for bringing my attention to this subject.

effect on PF, however, since it would not change the phonetic form. It would not have any effect on LF either, since it would not change the LF-representation. Note that the intermediate trace in the Spec of the embedded C^{\max} would be eliminated to make the heterogeneous chain created by successive cyclic *wh*-movement a legitimate object, an operator-variable pair. Hence, the embedded C cannot be a [+Q]-C with a strong Q-feature in (81).

In complex adjunct *wh*-questions like (75), however, the strong Q-feature of the embedded [+Q]-C has an effect on LF. Let us consider (75) again. If the embedded clause is headed by a [+Q]-C with a strong Q-feature, *why* is interpreted as modifying the embedded clause. If the embedded C does not have any strong Q-feature, *why* is not interpreted as modifying the embedded clause but rather as modifying the matrix clause. Hence, the embedded clause in (75) is headed by a [+Q]-C with a strong Q-feature without violating condition (80).^{27, 28}

There is empirical evidence to support the base-generation of *why* in the Spec of C^{\max} . It has been observed by, among others, Huang (1982) and Lasnik and Saito (1992) that *why* never stays in-situ:

²⁷One might argue against this analysis, claiming that (80) necessarily calls for global considerations. When constructing an N, we must look ahead to see whether the inclusion of a lexical item into the N has any effect on the interface levels. Recall, however, that the construction of an N is not part of a computation. Hence, even if (80) is global in nature, we can still maintain the view that a computation only needs local information.

²⁸As discussed in detail in chapter 1, in languages like Irish where [+Q]-C and [-Q]-C have distinct phonetic forms, the intermediate clause in a *wh*-question may be headed by a [+Q]-C without violating condition (80). This is because the existence of a strong Q-feature has an effect on PF. Morphological realizations of successive cyclic *wh*-movement are also observed in Chamorro (Chung 1982, 1994), Ewe (Collins 1993), Kikuyu (Clements 1984), Moore (Haik 1990), and Palauan (Georgopoulos 1985, 1991a, 1991b). Condition (80) also allows a strong Q-feature to be assigned to C in the intermediate clause of a *wh*-question in these languages, since the existence of a strong Q-feature has a PF effect.

- (82) a. *who left **why**
 b. *who thinks Mary left **why**

In the EST, the unacceptability of examples like (82) has been attributed to the ECP. Since the ECP is not available in the MP anymore, we have to look for an alternative account of the distribution of *why*. Let us consider how our analysis accounts for its distribution, taking (82a) as an example. During the derivation, we construct the following structure:

- (83) a. [C_[Q] [T^{max} T [V^{max} who_[Q] left]]]
 b. who_[Q]

When we come to the stage of the derivation where the strong Q-feature of the matrix C is to be checked, there are two possible continuations. We either copy *who* or select and merge *why*. Among these options, the EP requires that we should choose the latter. The strong Q-feature of C can only be checked by merger of *why* but not by copy of *who*. Hence, there is no legitimate way of deriving (82a). The unacceptability of (82b) can be accounted for in the same way. Our analysis therefore offers a minimalist account of the lack of *why* in-situ.

With the above discussion in mind, the next subsection considers the "domain barrier" effects with adjunct A'-movement.

3.4.2 The "Domain Barrier" Effects with Adjunct A'-movement

Let us consider the "domain barrier" effects with adjunct A'-movement, taking adjunct wh-movement as an example:

- (84) Complex NP Constraint (CNPC)
 a. Relative Clauses
 ***why** do you like [books that criticize John]

b. Non-relative Complex NPs

***why** did you deny [the evidence that Harry stole money]

(85) Subject Condition

***why** was [that John bought the book] predicted by you

(86) Adjunct Condition

***why** do you get jealous [because I spoke to Mary]

(87) Non-bridge Verb Condition

***why** did John whisper [that Bill criticized Mary]

(84)-(87) are all deviant under the readings where *why* modifies the clauses within the islands.

Let us first consider the non-relative clause case of the CNPC, taking (84b) as an example. Considering the appositive clause in (84b), i.e., *that Harry stole money*, we can construct it through checking the UFFs of the selected items in accordance with the ICP and the EP. Note that *why*, whose merger has not been triggered by any UFF, has not been selected or merged yet. The resultant structure is as follows:

- (88) a. [C^{max} that [T^{max} T [V^{max} Harry stole money]]]
 b. Harry

Note in passing that at this point of the derivation, since merger of *Harry* is not triggered by any UFF, *Harry* has not been merged with (88a) yet. In other words, there are two independent syntactic objects at this stage of the derivation as far as the appositive clause is concerned.

Turning to the main structure, let us consider the stage where we construct the following structure:

- (89) a. [C^{max} C_[Q] [T^{max} T [V^{max} you [deny the evidence]]]]
 b. you

headed by a [+Q]-C with a strong Q-feature, there is no way of generating (84b).

Alternatively, we can think of a derivation based on the N where both the matrix and embedded clauses are headed by [+Q]-C's with strong Q-features. In the derivation, the C of the appositive clause also has a strong Q-feature:

- (93) a. [C^{max} that_[Q] [T^{max} T [V^{max} Harry stole money]]]
 b. Harry

The strong Q-feature is checked if we select *why* and merge it into the Spec of C^{max}:

- (94) a. [C^{max} **why**_[Q] [that [T^{max} T [V^{max} Harry stole money]]]]
 b. Harry

After the checking operation takes place, the strong Q-feature of C is erased while the Q-feature of *why* remains. Note that this derivation would yield a representation where *why* modifies the appositive clause, since *why* is merged into the position where it takes the appositive clause in its immediate c-command domain.

Turning to the main structure, let us consider the stage where the strong Q-feature of the matrix C is to be checked:

- (95) a. [C^{max} C_[Q] [T^{max} T [V^{max} you [deny the evidence]]]]
 b. you

Note that the appositive clause (94a) has not been merged with the main structure (95a). The matrix C does not c-command *why*, which stays within a different syntactic object. The former cannot be checked by copying the latter. There is no way of checking the strong Q-feature of the matrix C. The derivation is canceled due to a violation of the ICP.

We cannot generate (84b) either from the N where both the matrix and embedded clauses are headed by [+Q]-C's with strong Q-features or from the one where only the matrix clause is headed by a [+Q]-C with a strong Q-feature. Hence, there is no legitimate way of deriving (84b). The non-relative case of the CNPC with adjunct wh-movement follows. The relative clause case of the CNPC, the Adjunct Condition, and the non-bridge verb condition can be accounted for in similar fashions.²⁹

Let us next consider the subject condition effects with adjunct wh-movement, taking (85) (repeated here as (96)) as an example:

(96) ***why** was [that John bought the book] predicted by you

If we construct the N where only the matrix clause is headed by a [+Q]-C with a strong Q-feature, *why* is directly merged into the Spec of the matrix C^{\max} . Although this derivation converges, it only derives an LF-representation where *why* modifies the matrix clause but not the embedded clause. Hence, we cannot generate (96) based on that N.

Let us next consider the N where not only the matrix clause but also the sentential subject is headed by a [+Q]-C with a strong Q-feature. Let us first consider the construction of the sentential subject *that John bought the book*. During the derivation, we come to the stage where the strong Q-feature of C of the sentential subject is to be checked:

²⁹The relative case of the CNPC is worth a mention in passing. One of its logically possible derivations is based on the N where the embedded clause as well as the matrix clause is headed by a [+Q]-C with a strong Q-feature. In that derivation, multiple specifier constructions proposed by Chomsky (1995) come into a play if we conform to the assumption that checking only takes place in specifier positions. Since the relative clause involves empty operator movement to the Spec of the relative clause, *why* must be merged into another Spec. If we assume contra Chomsky that checking may take place in adjoined positions, on the other hand, then *why* is adjoined to the relative clause to check the strong Q-feature. The same is also true of the clausal adjunct case of the Adjunct Condition if Larson's (1987) empty operator movement analysis of a clausal adjunct is correct.

- (97) a. [that_[Q] [T^{max} T [V^{max} John [bought the book]]]]
 b. John

The strong Q-feature of the C is checked immediately by merger of *why* into the Spec of the sentential subject, resulting in the following structure:

- (98) a. [C^{max} **why**_[Q] [that [T^{max} T [V^{max} John [bought the book]]]]]
 b. John

The strong Q-feature of C, being uninterpretable, is erased by this merger. The Q-feature of *why*, being interpretable, remains intact.

Turning to the main structure, we come to the stage where the strong feature of the matrix T is to be checked:³⁰

- (99) a. [T [V^{max} [C^{max} **why**_[Q] [that [T^{max} T [V^{max} John [bought the book]]]]] [was predicted by you]]]
 b. John

The strong feature of T is checked by copying the sentential subject:

- (100) a. [T [V^{max} [C^{max} **why**_[Q] **that John bought the book**] [was predicted by you]]]
 b. John
 c. [C^{max} **why**_[Q] **that John bought the book**]

As the derivation proceeds, we come to the stage where the strong Q-feature of the matrix C is to be checked

³⁰I assume following, among others, Boskovic (1995) and Delahunty (1983), that exactly like D^{max} subjects, sentential subjects originate in the Spec of V^{max} and then move to the Spec of T^{max}. It is not clear at this point, however, what strong feature triggers the raising of a sentential subject to the Spec of T^{max}. See Authier (1991), Koster (1978), and Safir (1985) for a different view.

- (101) a. [C_[Q] [T^{max} T [V^{max} [C^{max} why_[Q] that John bought the book] [was predicted by you]]]]
 b. John
 c. [C^{max} why_[Q] that John bought the book]

The strong Q-feature of the matrix C can only be checked by copying *why* within the Spec of V^{max} but not *why* in (101a). This is because the matrix C c-commands the former but not the latter. This yields the following structure:

- (102) a. [C [T^{max} T [V^{max} [C^{max} **why**_[Q] that John bought the book] [was predicted by you]]]]
 b. John
 c. [C^{max} why_[Q] that John bought the book]
 d. **why**_[Q]

After combining these syntactic objects, we get the following structure, with all formal features including Q-features being ignored:

- (103) [C^{max} why_i [C [[C^{max} why that Johnbought the book]_j [T [[C^{max} why_i that bought the book]_j [was predicted by you]]]]]]

Chain formation applies in the LF-component. Recall that *why that John bought the book* in the Spec of the matrix T^{max} is introduced by copying *why that John bought the book* in the Spec of V^{max}. These two occurrences are therefore identical in constitution but positionally distinct. *Why* in the Spec of the matrix C^{max} is introduced by copying *why* within the Spec of V^{max}. Hence, these two occurrences of *why* are identical in constitution but positionally distinct. These two copy operations create the following two chains:

- (104) a. CH = (*why that John bought that book, why that John bought that book*)
 b. CH = (*why, why*)

If we delete the non-head positions of the chains, we get the following LF-representation:

- (105) [_{C^{max}} **why** [C [[_{C^{max}} **why that John bought the book**]_j [T [_{t_j} [was predicted by you]]]]]]]

This apparently violates the condition of inclusiveness. This is because there are two occurrences of *why* in (105) while the N for this derivation only includes one occurrence of *why*. This representation is also excluded by the selectional property of the verb. Recall that we are assuming that the Q-feature of a *wh*-phrase is interpretable while the Q-feature of C is uninterpretable. The clause whose specifier position is occupied by a *wh*-phrase at LF is interpreted as an interrogative. Since the sentential subject has *why* in its specifier position at LF, it is interpreted as an interrogative. It is well known, however, that no verb selects an interrogative in its subject position, not to mention the verb *be*. Note in passing that since *why* does not need any variable to bind, there is no violation of the ban against vacuous quantification in (105).

To summarize this section, I have first argued that an adjunct *wh*-phrase should be directly merged into the Spec of C^{max} to check the strong Q-feature of C. I have then argued that the "domain barrier" effects with adjunct *wh*-movement straightforwardly follows from our theory of phrase structure.

3.5 Concluding Remarks

This chapter has considered the "domain barrier" effects with feature-driven A'-movement and shown that they straightforwardly follow from our theory of the composition of phrase structure. It was also shown that our analysis of the "domain barriers" can be extended to account for another locality condition which constraints extraction out of moved phrases. I have argued that the pre-minimalist locality theories are incompatible with the MP, since they crucially rely on the notions not available under the MP. I have also argued that our theory has conceptual advantages over the previous minimalist approaches to locality. This is because our locality theory only makes use of the ICP and the EP, which are local "heuristic algorithms" ("computational tricks") for globality induced by the interface conditions on UFFs and Ns. It was shown that our theory of locality diverges from all the previous approaches in claiming that locality facts should not follow from restrictions on movement but from restrictions on merger.