Overt Scope Marking and Covert Verb-Second

*Samuel David Epstein*

In this article I investigate certain phenomena relating to superiority, the Empty Category Principle (ECP), and scope. I propose a chain-based scope-marking convention and a new analysis of adjunction, and hypothesize that English is a covert verb-second grammar. The analysis is couched within checking theory and ultimately within the bare theory of phrase structure. I propose category-neutral(ized) LF representations, displaying VP-recursion but lacking functional heads and their projections, and I suggest that this, in turn, allows significant simplification of index-sensitive head government conditions appearing in many contemporary formulations of the ECP.

*Keywords:* minimalism, checking theory, bare phrase structure, verb-second, ECP, superiority

In this article I reanalyze certain superiority phenomena and investigate consequences of the account I propose. In section 1 I briefly review two standard analyses of superiority effects and a problematic type of case (from Lasnik and Saito 1992) that neither analysis handles correctly. In section 2 I present the Operator Disjointness Condition (ODC), which Lasnik and Saito (L&S) (1992) propose as the central principle accounting for superiority effects (including the problematic cases they present). In section 3 I argue that the ODC can be reformulated as a more natural principle of scope marking: the Scope-Marking Condition (SMC). I then show that the success of this reformulation rests crucially on incorporating the theory of linked chains independently motivated by Chomsky and Lasnik (1993). In section 4 I present independent support for the SMC. In section 5 I tentatively propose a new theory of adjunction. In section 6 I investigate how subject position in English becomes properly governed by I-raising at LF (a process crucial to L&S’s analysis of superiority). In particular, I seek an explanation for the fact that LF I-to-C movement, but not syntactic I-to-C movement, renders subject position properly governed. I argue

This article is a revised version of a manuscript completed in July 1992, which appeared as Epstein 1993. The analysis presented in section 4.1 first appeared in Epstein 1989b. For very insightful discussion of the analyses presented here, I thank Maggie Browning, Noam Chomsky, Bob Freidin, Gunther Grewendorf, Howard Lasnik, Geoff Poole, Daiko Takahashi, and Hóskuldur Thraínsisson. I am especially indebted to Erich Groat, Madelyn Kissock, Hisatsugu Kitahara, Elaine McNulty, Robyne Tiedeman, and Esther Torrego for their extensive and invaluable comments. I am also grateful to all the Harvard University students in Linguistics 213r (Spring 1993) who listened patiently and thoughtfully to the material presented here. Portions of this article were presented at the University of Massachusetts at Amherst in February 1993, the University of Connecticut at Storrs in October 1993, and Princeton University in December 1993, and I thank the audiences at these presentations, too, for their helpful comments. Finally, I am indebted to an anonymous *LI* reviewer for carefully reading the manuscript and providing insightful commentary.
that a principled account can be provided within the framework of checking theory (Chomsky 1991, 1993). I propose, among other things, that (a) English is a ‘‘covert verb-second’’ grammar, (b) LF representations derived by V-raising are VP-recursion structures resulting from checking and deletion of functional heads and their projections, and (c) index-sensitive head government conditions may be at least partially eliminable and replaced by simpler and arguably more natural requirements on traces (chains), given the derived (VP-recursion) constituent structure I propose for checking-induced deletion. In section 7 I reformulate certain aspects of the proposed analysis within the bare theory of phrase structure presented in Chomsky 1994.

1 Superiority Effects: Previous Analyses and Counterexamples

The classic example of the superiority effect is given in (1).

(1) a. [who, [t, bought what]]

One account of this phenomenon is provided by Chomsky’s (1973:246) Superiority Condition.

(2) The Superiority Condition

a. No rule can involve X, Y in the structure . . . X . . . [ . . . Z . . . WYV . . . ] . . . , where the rule applies ambiguously to Z and Y, and Z is superior to Y.
b. ‘‘[T]he category A is ‘superior’ to the category B . . . if every major category dominating A dominates B as well but not conversely.’’

This constraint on rule application renders (1b) ungenerable. In the D-Structure representation, X = Comp, Z = who, and Y = what. Therefore, wh-movement, applying ambiguously to X(who) or to Y(what), is not allowed to ‘‘involve’’ X(Comp) and Y(what), since Z(who) is superior to Y(what). By contrast, (1a) is generable.

Another account of cases like (1b) is provided by Aoun, Hornstein, and Sportiche (1981:81), who attribute the account to Chomsky (see Chomsky 1981). Under this alternative, the Superiority Condition is reduced to the Empty Category Principle (ECP), given the following three assumptions:

(3) a. Comp-indexing: At S-Structure, [Comp XP₁ . . .] → [Comp₁ XP₁ . . .] iff Comp dominates only i-indexed elements.
b. At LF all wh-phrases occupy Comp.
c. Only an indexed Comp (Comp₁) coindexed with a subject trace (t_i) can properly govern t_i.

Aoun, Hornstein, and Sportiche consider the following ungrammatical French analogue of (1b):

(4) *Quel manteau qui a acheté?
what (sort of) overcoat who has bought

Given (3a), the S-Structure representation of (4) is (5).

(5) [S^[Comp₁ quel manteau,] [S qui, a acheté t_j]]
Given (3b), *qui moves to Comp in the LF component, yielding (6).

\[(6) \quad \left[ S' \right] \left[ \text{Comp qui} \right] \left[ s_1 \text{ a acheté t} \right] \]

Given (3c), the ECP is violated: \( t_i \) is not properly governed.

Thus, there are at least two general types of approaches to superiority phenomena: a constraint on rule application and an alternative hypothesis under which the constraint on rule application arguably reduces to the ECP, a constraint on representation.

A problem for both analyses is the grammaticality of examples like (7), from L&S 1992.

\[(7) \quad \begin{aligned} a. \text{Who wonders what who bought?} \\ b. \quad \left[ CP \text{ who_k C_i [IP t_i wonders [CP what_j C_j [who_k bought t_j]]]} \right] \end{aligned} \]

Incorrectly, the embedded CP in (7b) is ungenerable, given the Superiority Condition. Under the alternative account, the LF representation of (7b) violates the ECP, since the trace left by LF movement of who_k cannot be properly governed. This incorrect result comes about as follows: First, I assume (3a) can be reformulated in more current terms as (8).

\[(8) \quad \text{Specifier-head coindexing occurs at S-Structure.} \]

The effects of (8) are indicated by the indices in (7b). Second, I assume (3b) can be reformulated as (9).

\[(9) \quad \text{In LF representation, each wh-phrase either occupies [Spec, CP] or is adjoined to a wh-phrase occupying [Spec, CP].} \]

Consequently, who_k must move in the LF component, leaving a trace in subject position. Third, I assume the following descriptive generalization, discussed in detail in Epstein 1992:

\[(10) \quad \text{A wh-phrase occupying [Spec, CP] at S-Structure does not undergo LF movement.} \]

Under (10), both who_i and what_j remain in [Spec, CP] in the LF component. Consequently, the only way for who_k to satisfy (9) is to adjoin to a wh-phrase in [Spec, CP]. But in such an adjoined (nonspecifier) position, specifier-head indexing (in LF) is blocked. Therefore, no Comp (= C^0) will bear the index k. Consequently, proper government of the subject trace is precluded under the following definition of antecedent government (see L&S 1992), which dictates that a trace must be bound by an X^0 in order to be antecedent-governed by it:

\[(11) \quad \text{Antecedent government } \]

A antecedent-governs B iff

a. A = X^0, and
b. A binds B, and
c. B is subjacent to A.

---

In summary, if (1b) is excluded either by the Superiority Condition or under the reduction of the Superiority Condition to the ECP, then (7) will be wrongly excluded along with it.

The grammaticality of (7) is not the only fact to be explained: its interpretation must be accounted for as well. In particular, who_k, the wh-phrase in the embedded subject position at S-Structure, has no narrow scope interpretation; instead, it obligatorily takes wide scope (see L&S 1992). This is suggested by the fact that (12a) is, but (12b) is not, an appropriate answer (as indicated by ‘‘#’’).2

(12) a. Sue wonders what Bill bought.
   b. #Sue wonders what who bought.

Thus, in the LF representation of (7b), who_k obligatorily adjoins to who_i, occupying the matrix [Spec, CP], yielding an LF representation roughly as follows (see below for further details):

(13) [CP who_0 who,[C_i] [IP t wonders [CP [what,[C_j] [IP t bought t_j]]]]]

Notice that principle (9) alone fails to account for the scope of who_k.

Thus, (at least) two questions are raised by examples like (7): 3

(14) a. How can we at the same time exclude (1b) and allow (7b)?
   b. Why must who_k in (7b) take wide scope; that is, why must who_k adjoin to who_i in the LF component?

2 The Operator Disjointness Condition

In this section I present L&S’s (1992) answers to the questions posed in (14). First, with respect to (14a), the exclusion of the central case of superiority (1b), they reject an account in terms of the Superiority Condition, since, as noted, this condition wrongly excludes (7b) as well. They also reject the ECP account, which rests on the assumption that, in English, a trace left by a subject undergoing LF movement cannot be properly governed. As L&S note, this assumption cannot be maintained. The well-formedness of the LF representation (13) indicates that a wh-phrase in subject position at S-Structure can indeed undergo LF movement (in fact, long-distance LF movement) and leave a trace that is somehow (see below) properly governed.

Thus, it appears that subject position somehow becomes properly governed in LF, and therefore the ECP analysis of (1b) cannot be maintained. As L&S discuss, further evidence for rejecting this analysis is that it is unable to account for ‘‘pure’’ superiority effects such as the following (see Hendrick and Rochemont 1982):

(15) *I wonder [CP who, C_i [you told whom [CP (t_i) [IP PRO to see t_i]]]]

---

2 An LI reviewer correctly notes that (12b) is not just infelicitous, but also ungrammatical. This, too, will be predicted by the analysis presented below.

3 For other current analyses of superiority phenomena concerning data not discussed here, see, among others, Kitahara 1993a, Williams 1994: sec. 5.2.7, and Freidin 1995.
That is, LF movement of whom satisfies the ECP, given the following disjunctive formulation (from L&S 1992:28, 52):

(16) a. **ECP**
   A trace must be properly governed.
   b. A trace is properly governed iff it is antecedent-governed (see (11)) or lexically properly governed.
   c. **Lexical proper government**
      A trace is lexically properly governed iff it is assigned Case or a 0-role by a lexical head.

Under (16c), even though the LF trace of whom cannot be antecedent-governed (since who1 has indexed C0 at S-Structure—see (8)—and remains in this position in LF—see (10)), it satisfies the ECP by being lexically properly governed.

Therefore, L&S argue, one can appeal neither to the ECP nor to the Superiority Condition to explain superiority phenomena, in particular, the contrast between the ungrammaticality of (1b) and the grammaticality of (7). In order to describe the contrast, L&S propose the following condition (pp. 120–121):

(17) **The Operator Disjointness Condition (ODC)**
   a. A wh-phrase X in [Spec, CP] is O-disjoint (operator-disjoint) from a wh-phrase Y if the assignment of the index of X to Y would result in the local A-<binding of Y by X (at S-Structure).
   b. If two wh-phrase X and Y are O-disjoint, then they cannot undergo Absorption.

(Regarding Absorption, see Higginbotham and May 1981.)

First, consider how (17) excludes (1b). The S-Structure representation of (1b) is (18).

(18) \[ CP \text{what}_j \text{did \_IP who}_i \text{buy \_t}_j \]

The first step is to determine whether whatj is O-disjoint from whoi (17a). This is done by hypothetically assigning the index j to who. This results in the hypothetical representation (19).

(19) \[ CP \text{what}_j \text{did \_IP who}_i \text{buy \_t}_j \]


(20) X locally A-binds Y iff
   a. X binds Y, and
   b. X occupies an A-position, and
   c. if Z binds Y, then Z binds X.

Since the hypothetical assignment of the index of what to who in (19) does indeed yield local A-binding of who by what in the hypothetical representation (19), these two operators are marked as O-disjoint in the (nonhypothetical) S-Structure representation (21).
Now, in the LF component, (9) applies, dictating that who must either occupy [Spec, CP] or be adjoined to a wh-phrase in [Spec, CP] in the LF representation. As discussed above, who cannot occupy [Spec, CP] since this position is occupied by what. Consequently, who has only one option: it must adjoin to what. However, in the resulting configuration, in which one wh-phrase is adjoined to another, the structural description of Absorption is met. Assuming (with L&S) that Absorption is obligatory, the two wh-phrases must undergo this process—but they cannot, by (17b), since they are (were) O-disjoint. Hence, (1b) has no well-formed LF representation and is therefore excluded.4

Now that (1b) is excluded, how is (7) allowed? Consider again its S-Structure representation.

\[
(22) \begin{array}{c}
[CP \text{ who}_1 [IP \text{ ti wonders } [CP \text{ what}_j [IP \text{ who}_k \text{ bought } t_j]]]]
\end{array}
\]

What_j and who_k are O-disjoint in (22). This is because hypothetical assignment of j to who_k does indeed yield hypothetical local A*-binding of the embedded subject who by what.5 Since who_k and what, are O-disjoint, who_k cannot be adjoined to what_j in LF since obligatory Absorption

4 An LI reviewer asks whether there is any independent motivation for L&S’s assumption that Absorption is obligatory. L&S (1992:189n. 22) attribute the obligatory wide scope reading of what in (i) to the impossibility of Absorption (“. . . in approximately the sense of Higginbotham and May (1981) . . .”) between what and whether:

(i) who, t wonders whether John bought what

If Absorption in this sense were merely optional, the representation of the unavailable reading would presumably be generable.

From L&S 1992:189n. 29, we can infer that L&S assume that, in the absence of Absorption, the LF representation of, say, (ii)

(ii) Who bought what?

would contain a free variable, namely, the trace of what. Thus, Absorption would be, in effect, obligatory in that failure to apply it invariably yields ill-formedness.

5 For ease of exposition, I have suppressed C^0 indexing in (22). Under specifier-head indexing, the embedded CP in (22) appears as in (i):

(i) . . . [CP what [C^0_j [who_k bought t_j]]]

Hypothetical indexing changes who_k to who_j. Even though C^0_j “intervenes,” what_j locally A*-binds who provided that the second occurrence of bind in (20c) is defined as in (ii),

(ii) X binds Y iff
   a. X and Y are coindexed, and
   b. X m-commands Y.

where m-command (see Aoun and Sportiche 1983) may be defined as in (iii).

(iii) X m-commands Y iff the minimal maximal projection dominating X dominates Y.

Another possibility is that heads are simply ignored in calculating local binding relations among maximal projections (contra Epstein 1991). To an LI reviewer, the latter possibility appears preferable, since the former incorporates an m-command definition of binding (internal to the definition of local binding) that, as we shall see momentarily, would coexist with a c-command definition of binding (internal to the definition of antecedent government), an arguably unattractive state of affairs.
will be blocked by (17b). So far this is exactly what happened in (1b). Unlike in (1b), however, in (22) who_k has another possible LF landing site: it can be adjoined to who_i occupying the matrix [Spec, CP]. In fact, this is a licit landing site: who_i and who_k are not O-disjoint (i.e., hypothetical assignment of i to who_k does not yield local Â-binding of who_k by who_i because of the “intervening” matrix subject t_i). Thus, provided who_k adjoins to who_i in LF, the derivation is well formed: the obligatory rule of Absorption can apply. Thus, only one LF representation can be derived from the S-Structure representation (22), roughly like (13).

\[ (23) \{_{CP} who_k, who_i, [_{IP} t_i, I_i wonders \{_{CP} what_j, C_j, [_{IP} t_k, I_k bought t_j]\}] \}\]

In summary, the ODC (17) has answered the question posed in (14a): (1b) is excluded and (7) is generable. Question (14b) has been answered too: who_k must take wide scope, as in (23), since, at S-Structure, what_j and who_k are O-disjoint.

There is, however, at least one outstanding question. The LF movement of who_k constitutes (wh-island hopping) long-distance subject movement, the result of which produces a “classic Comp-trace” configuration. Why then does the LF representation (23) satisfy the ECP, when the S-Structure representation (24), exhibiting similar long-distance subject movement, violates it?

\[ (24) *_{IP} who_i, [_{C} do]_{IP} the men wonder \{_{CP} what_j, C_j, [_{IP} t_i, I_i bought t_j]\}\]

The violation in (24) is explicable given the definition of the ECP in (16). The trace t_i, created by long-distance subject movement of who_i, is not lexically properly governed since it receives neither a Case feature nor a θ-role from a lexical head (Case is assigned by I; θ-role is assigned by VP compositionally, or, under the VP-Internal Subject Hypothesis, the subject position lacks a θ-role (see Koopman and Sportiche 1986 and references cited there)). Furthermore, t_i is not antecedent-governed according to (11) since it is not subjacent to a head that binds it. The only head coindexed with t_i is I_i, t_i is indeed subjacent to I_i; moreover, they are coindexed. To obtain the desired result that I_i does not in fact antecedent-govern t_i, it must be the case that I_i (= head of IP) does not bind t_i (= specifier of IP). This result can be obtained if the definition of binds in (11b) incorporates the branching node definition of c-command given in Reinhart 1979 (a definition I seek to deduce in Epstein 1994, to appear, and which is discussed both in further detail and in a broader theoretical context in Epstein et al., to appear).

\[ (25) Binds \text{ in (11b)} = X \text{ binds } Y \iff \]
\[ a. \ X \text{ and } Y \text{ are coindexed, and } \]
\[ b. \ X \text{ c-commands } Y. \]

\[ (26) X \text{ c-commands } Y \iff \text{ the first branching node dominating } X \text{ dominates } Y \text{ and neither } \]
\[ X \text{ nor } Y \text{ dominates the other. } \]

Given (25), an (unmoved) head does not bind, and therefore cannot antecedent-govern, its specifier. Thus, in (24) t_i is neither antecedent-governed nor lexically properly governed; it therefore violates the ECP, as desired.

But notice now that the LF representation (23) will be wrongly excluded, along with the S-Structure representation (24), since the subject trace similarly fails to be properly governed. We
are thus led to believe that (23) is not in fact the LF representation of (7). Rather, there would seem to be some LF operation that renders the subject position properly governed at this level. L&S (1992:117) propose:

(27) In the LF component, I can substitute into C or can adjoin to IP.

Under either option, the subject trace in (23) will be bound (under definition (25)) and hence antecedent-governed, and the ECP will therefore be satisfied. Thus, L&S account for all the relevant properties of (7).

3 Reformulating the ODC

To summarize up to this point: L&S account for the grammaticality of (7) by rejecting both Chomsky’s (1973) account and an ECP account of superiority phenomena and by postulating (27), whereby the subject position can become properly governed at LF. The unique wide scope interpretation of (7) and the ungrammaticality (or uninterpretability) of (1b) (the classic case of the superiority effect) are accounted for by the ODC (17). However, the ODC is descriptive and, as L&S note, the question arises whether it can be deduced from, or replaced by, natural principles. I will show that this is indeed possible, by reformulating the ODC as a principle of scope marking.

To begin with, consider again the ODC (17) and the central contrast (1a) versus (1b) (repeated here).

(1) a. [who, [ti bought what]]
   b. *[what, did [who buy ti]]

In a nutshell, (1a) is allowed by the ODC because the trace of the wh-phrase who (who in [Spec, CP]) c-commands the wh-in-situ at S-Structure; (1b) is excluded because this condition is not met. Suppose then that we simply restate the ODC as the following (preliminary) condition:

(28) In the LF component, a wh-in-situ Y can adjoin to a wh-phrase X occupying [Spec, CP] only if the trace of X c-commands Y at S-Structure. (preliminary formulation)

---

6 Regarding the disjunction in (27), L&S note that LF I-substitution into C is not sufficient to generate (i), for example, given Quantifier Raising.

(i) Everyone left.

LF application of QR (May 1977, 1985) and I-to-C movement yield the following representation:

(ii) *[CP[1, I]]IP everyone, [IP ti, ti left]]]

The ECP is violated: two IPs ‘intervene’ between the ti subject and I. Therefore, ti is not subjacent to I, (see L&S 1992 and below), with the result that ti is not antecedent-governed by I, given (11). In order to generate a well-formed LF representation of (i), L&S assume that, as one option, in addition to adjoining to C, I can adjoin to IP, yielding the LF representation in (iii), which satisfies the ECP.

(iii) [CP[1, I]]IP everyone, [IP I, [IP ti, ti left]]]

I will return to this disjunction in condition (27) below, in particular the non-structure-preserving adjunction of I to IP, expressed in the second disjunct.

7 Regarding (27), the question arises why syntactic I-to-C movement does not render subject position properly governed at S-Structure. This difference between the syntax and LF is analyzed in section 6.
With this in mind, let us consider the following example:

(29) *[CP [IP John likes who]]

Apart from the echo-question interpretation, (29) is ungrammatical (see, e.g., Chomsky 1986:53). Given the LF requirement (9), the derivation in which who moves to [Spec, CP] in LF must be excluded. Neither the ODC (17) nor the alternative (28) applies to this type of derivation; hence, neither can exclude it. Following Chomsky (1986:53), suppose we assume that a grammar like that of English incorporates the positive value of the following parameter (whereas the grammar of, say, Chinese incorporates the negative value):8

(30) In the LF component, wh-in-situ must adjoin to a wh-phrase in [Spec, CP] (LF substitution into empty [Spec, CP] is prohibited).

This parameter is in turn arguably derivable from parameterizing the morphological features of C⁰ (as proposed in Chomsky 1993); that is, in English [+ wh] C⁰ has strong N-features. Thus, when a [+ wh] C⁰ is generated, a wh-phrase must move to [Spec, CP] prior to Spell-Out, so that the strong [+ wh] feature (an illegitimate PF object) can be checked and hence deleted. Thus, in (29) (assuming the matrix C⁰ is [+ wh]) the postponement of wh-movement until LF entails that the strong N-feature of C⁰ is not checked prior to Spell-Out and therefore remains at PF, at which level it is illegitimate. This derivation thus does not converge (i.e., it “crashes”) at PF. If, in (29), the matrix C⁰ is [− wh] (possible if matrix (hence lexically unselected) C⁰ is freely assigned [+ / − wh], as in L&S 1992), then there is simply no position in which who can have its [+ wh] feature checked, since only [+ wh] C⁰ can check a [+ wh] feature. This unchecked feature therefore remains throughout the derivation. Such unchecked [wh] features borne by wh-in-situ must be legitimate at PF (if they were not, what in (1a)—and, more generally, all instances of wh-in-situ—would induce nonconvergence at PF). I assume then that these features are illegitimate LF objects; hence, wh-in-situ must move to the checking domain of a [+ wh] C⁰ in LF.⁹

Summarizing this analysis, I assume:

(31) a. Universal: Wh-phrases bear a [+ wh] feature that can be checked only in the checking domain of a [+ wh] C⁰.
b. Universal: An unchecked [+ wh] feature borne by a wh-phrase is an illegitimate LF object.
c. Parameter: [+ wh] C⁰ = strong vs. weak (value in English: strong)

Under this analysis, there is no convergent derivation that includes the S-Structure representation (29). By contrast, (31) alone fails to block superiority violations such as (1b), where C⁰ checks the [+ wh] feature on what prior to Spell-Out and the [+ wh] feature on who in LF. I therefore incorporate the movement constraint (28), which restricts the class of configurations in which a

---

8 Chomsky writes, “In English... a wh-phrase in situ must move to specifier of CP [at LF].... and it may do so... only if this position is already occupied by a wh-phrase; wh-in-situ is ‘attracted’ by a wh-phrase in ‘scopal’ position.”

9 This “minimalistic” analysis thus derives (9). “Checking domain” is defined below.
wh-in-situ can adjoin to a wh-element in [Spec, CP] (required for convergence at LF). In the configuration displayed by (1b), for instance, (28) prevents LF movement of wh-in-situ and therefore precludes convergence at LF; that is, the LF-illegitimate [+ wh] feature borne by who-in-situ remains unchecked at LF.

Regarding (28), notice also that if the trace of X c-commands Y at S-Structure (or Spell-Out), then X itself (= the wh-phrase in [Spec, CP]) will also c-command Y at S-Structure, given that the Proper Binding Condition on traces applies at this level. This then suggests (see below) that (28) can be redefined as (32a), with the definition of chain c-command given in (32b).

32a. In the LF component, a wh-in-situ can adjoin to a wh-phrase X occupying [Spec, CP] only if the chain of X c-commands Y at S-Structure. (preliminary formulation)
32b. A chain C c-commands a position P iff every member of C c-commands P.

The constraint (32a) amounts to the arguably natural condition that a c-commanding wh-chain is a scope marker for wh-in-situ. However, with respect to (32a), one might ask why adjunction of a wh-phrase in situ Y to the wh-phrase X should depend on the configurational relation of the chain of X (as opposed to the wh-phrase X itself) to Y. The relevance of the configurational relation between the chain of X, on the one hand, and Y itself, on the other (and the irrelevance of the X/Y relation itself) is expected under the following provisional proposal:

33 In the LF component, wh-in-situ is not adjoining to a wh-phrase (= part of a chain) occupying [Spec, CP]; rather, wh-in-situ is adjoining to the wh-chain headed by the wh-phrase in [Spec, CP].

It just happens that adjunction to a chain head and adjunction to a chain yield identical configurations.

As intended, (33) allows us to restate (32a) as the more natural Scope-Marking Condition.

34 Scope-Marking Condition (SMC)
In the LF component, a wh-in-situ Y can adjoin to a wh-chain X only if X c-commanded Y at S-Structure.

The SMC amounts to the (simple) assumptions stated in (35).

35a. Only a c-commanding wh-chain is a scope marker for wh-in-situ.
35b. Scope marking is determined at S-Structure.
(Regarding (35b), see footnote 17.) From here on, I will refer to (33) and (34) together as the SMC analysis. The classic cases of superiority (typified by (1b)), the problematic cases provided by L&S (typified by (7)), and the “pure” cases (typified by (15)) are all accounted for by the SMC analysis, suggesting that the arguably less natural ODC is eliminable. 12

Before I conclude this section, I should address one question in particular regarding the reduction of the ODC to the SMC. Notice that the chain-based SMC says that in order for a wh-in-situ Y to adjoin to a wh-phrase (chain) X in LF, every member of the chain X must c-command Y at S-Structure. By contrast, the ODC, which the SMC is intended to supplant, entails only the following: in order for a wh-in-situ Y to adjoin to a wh-phrase X in LF, it is sufficient that merely a single trace of X c-commands Y at S-Structure. This is sufficient because the presence of such a single trace of X is enough to prevent local Â-binding under hypothetical indexing. In this respect, the ODC and the SMC would appear to diverge. One relevant type of case to consider in this regard is the grammatical (36).

\[(36) \{_{CP} \, \text{who}, \, [_{IP} \, t'_i, \text{seems to whom} \, [_{IP} \, t_i, \text{to be intelligent}]]\}\]

Clearly, the ODC makes the right prediction. Who is not O-disjoint from whom; consequently, whom can adjoin to who in LF and these two operators can undergo Absorption, as the ODC analysis requires.

By contrast, it appears that the SMC analysis (wrongly) predicts that (36) is ungrammatical. Since the 3-membered chain of who, (who, t′i,t′i), does not c-command whom at S-Structure (i.e., it is not the case that every member of the chain c-commands whom), it appears as though this chain is not a scope marker for whom. Consequently, in LF either whom does not adjoin to this chain, resulting in nonconvergence at LF, or whom does adjoin to the chain, violating the SMC (34).

The problem is only apparent, however. There is in fact no chain of the form (who, t′i,t′i) in (36). Following the analysis presented by Chomsky and Lasnik (1993), I assume that syntactic movement of who results in the appearance of two chains in structures like (36), namely, those in (37).

we could restate the SMC (34) as a purely chain-theoretic condition, eliminating reference to “...a wh (i.e., a wh-phrase) in situ...” This might be done as follows:

(i) In the LF component, a wh-chain in situ Y can adjoin to a wh-chain X only if X c-commands Y at S-Structure. Of course, we would then need to assume that the operation of adjoining a chain Y to a chain X appears identical to the operation of adjoining the head of Y to the head of X. In what follows, I will continue to assume (33)–(35) as given in the text.

12 It might seem at this point that the SMC analysis is empirically equivalent to Pesetsky’s (1987) Path Containment Condition (PCC) analysis of superiority phenomena: the PCC disallows “interlocked” Â-chains, and this is apparently what the SMC prohibits as well. However, the SMC and the PCC are in fact empirically (quite) distinct. As an illustration, note that the PCC (incorrectly) disallows the LF representation of (7b), in which the Â-chain headed by what, and the Â-chain headed by who, “interlock” (see (23)). By contrast (as we have seen), the SMC (correctly) allows such representations containing “interlocked” Â-chains; in other words, the fact that the who chain and the what chain interlock in the LF representation (23) is irrelevant in the eyes of the SMC, a principle permitting LF adjunction of who to the who chain that c-commands it at S-Structure.
(37) \[ C_1 = (\text{who}_i, t_i') \]
\[ C_2 = (t_i', t_i) \]

Chain \( C_1 \) is an operator-variable (\( wh \)) chain; chain \( C_2 \) is an argument chain. (These two chains are said to be ‘‘linked’’ at the position \( t_i' \).) Given that (36) exhibits the chain structure (37), the SMC correctly predicts that in the LF component \( \text{whom} \) can indeed adjoin to the \( wh \)-chain \( C_1 \). This is a licit operation because, at S-Structure, the \( wh \)-chain \( C_1 \) does indeed c-command \( \text{whom} \).

Thus, Chomsky and Lasnik’s (1993) analysis of chain structure allows us to maintain the SMC (in the face of examples like (36)) precisely because the SMC is formulated in terms of the arguably natural notion ‘‘c-commanding chain.’’ In the following section I investigate certain other aspects of the SMC analysis.

4 The Scope-Marking Condition

Consider the following example:

(38) *Why did Bill buy what?


Like Huang, I (and those I have informally polled) find such cases ungrammatical (i.e., worse

13 See Kitahara 1992, 1996, for other arguments directly supporting this analysis of chains. For an alternative analysis of superiority phenomena, see Kitahara 1993a.

14 Before proceeding, I should note that the SMC and ODC also diverge with respect to the following type of derivation:

(i) a. S-Structure \[ [\text{CP} [\text{IP} [\text{who}_i \text{wonders} [\text{CP} [\text{who}_j [t_j \text{left}]]]]]] ]

b. LF \[ [\text{CP} [\text{IP} [t_i \text{wonders} [\text{CP} [\text{who}_i [\text{who}_j [t_j \text{left}]]]]]]] ]

The ODC allows such LF movement of \( \text{who} \), since \( \text{who} \), and \( \text{who}_j \) are not O-disjoint in (ia). By contrast, the SMC prohibits such movement and is therefore partially redundant with the Proper Binding Condition applying to \( t_j \) at LF (and/or the ban on vacuous quantification/Full Interpretation). If this redundancy is in fact undesirable (see Epstein 1990a for arguments that redundant analyses can be empirically motivated), we could eliminate it by adding (iib) to the existing formulation of the SMC (given informally in (iiia)).

(ii) At S-Structure a \( wh \)-chain \( C \) is a scope marker for a \( wh \)-in-situ \( X \) iff

a. every member of \( C \) c-commands \( X \) (see (34)/(35)), or

b. no member of \( C \) c-commands \( X \).

With the addition of (iib), the SMC, like the ODC, would permit derivation (i), thereby eliminating the redundancy.

The addition of (iib) would also allow ‘‘inversely linked LF \( wh \)-movement’’ (allowed by the ODC), as in (iii).

(iii) a. S-Structure \[ [\text{CP} [\text{NP} [\text{whose picture of whom}_i [t_i \text{fell}]]]] ]

b. LF \[ [\text{CP} [\text{NP} [\text{whose picture of whom}_i [\text{NP} [\text{whom}_j [\text{IP} [t_j \text{fell}]]]]]]] ]

That is, since no member of the chain ([whose picture of whom] \( i,t_i \) \( C \)) c-commands \( \text{whom} \) at S-Structure, \( \text{whom} \) can adjoin to the chain in LF, as in (iib).

Notice finally that the disjunction in (ii) can be eliminated by restating this condition as a condition under which a \( wh \)-chain is not a scope marker.

(iv) At S-Structure a \( wh \)-chain is not a scope marker for a \( wh \)-in-situ \( X \) iff one member of \( C \) does c-command \( X \) and some other member of \( C \) does not c-command \( X \).

I believe that the addition of (iib) to the existing analysis (iiia) is irrelevant in what follows.
than the grammatical *Why did Bill buy a house?*, but not strongly so. For example, ECP violations such as (39) seem worse (as Huang notes).

(39) **What did Bill buy why?**

The contrast between (38) and (39) follows from the ECP. The LF representation of (38) is (40).

(40) \[ CP \text{what}_i \text{why}_j [\text{C}_j (\text{did})] [\text{IP Bill buy t}_j t_i] \]

The trace of *why* (*ti*) is antecedent-governed (16c) by the X' C that (as a result of specifier-head indexing) binds *ti*, which is in turn subjacent to C. The trace *tj* is lexically properly governed. Thus, (38) satisfies the ECP.

By contrast, (39), the LF representation of which is given in (41), violates the ECP since *ti* is neither antecedent-governed nor lexically properly governed.

(41) \[ CP \text{why}_i \text{what}_j [\text{C}_j (\text{did})] [\text{IP Bill buy t}_j t_i] \]

Thus, (39) violates the ECP.

Although the ECP predicts a contrast between (38) and (39), it leaves the ungrammaticality of (38) unexplained. Regarding degree of ungrammaticality, (38) seems on a par with (1b), the classic case of superiority, suggesting that we should pursue a unified analysis of these two types of examples.

Further motivation for such unification is provided by the following facts. Recall from section 1 that when the classic case of superiority (1b) is (in effect) embedded, as in (7a), the result is grammatical and, moreover, the *wh*-in-situ requires a wide scope reading.

(1) b. *What did who buy?*

(7) a. Who wonders what who bought?

Example (38) displays exactly the same properties. When it is (in effect) embedded, as in (42), the result is grammatical and, furthermore, the *wh*-in-situ requires wide scope interpretation (as evidenced by (43)).

(42) Who wonders why Bill bought what?

(43) a. #Fred wonders why Bill bought what.
   b. Fred wonders why Bill bought a car.

Given this similarity, it seems reasonable to seek a unified analysis of (1b) and (38).\textsuperscript{15}

Can the SMC analysis provide a unified account? To begin, recall the SMC analysis of the S-Structure representation (1b).

(1) b. [what, did [who buy t]]

\textsuperscript{15} Although I have abandoned Chomsky’s (1973) Superiority Condition, it is interesting to note that it cannot account either for the deviance of (38) or for its similarity to (1b). In (38), unlike in (1b), the presumably superior phrase *why* was indeed selected to undergo *wh*-movement. Hence, (38) satisfies the Superiority Condition whereas (1b) violates it. The ECP is similarly ineffective. As noted, (38) satisfies the ECP and (as discussed in section 2) (1b) does too.
The SMC analysis predicts ungrammaticality: the wh-chain (what, t_j) is not a scope marker for who since it is not the case that every member of the chain c-commands who.

Consider now the S-Structure representation of (38).

(44) \[ \text{[CP why}_j \text{ did [IP Bill buy what t}_j]] \]

The SMC analysis could exclude (44) exactly as it excludes (1b) if we could maintain that, at S-Structure, t_j fails to c-command the wh-in-situ what and therefore the chain (why_j, t_j) is not a scope marker for this wh-in-situ. However, it seems implausible that t_j fails to c-command what. Whatever the exact location of the adjunct trace, it presumably c-commands the direct object what.\(^{16}\) In particular, if t_j is a sister to the V or higher, it will c-command what. Thus, the chain (why_j, t_j) presumably does c-command what at S-Structure and therefore is indeed a scope marker for it. In this way, the SMC allows what to adjoin to this chain in LF and, incorrectly, no principles are violated. Hence, the SMC analysis apparently fails to account both for the deviance of (38) and for its presumed similarity to (1b).

Notice however that the SMC analysis would yield exactly the right result if the S-Structure representation of (38) were not (44), but (45), in which no adjunct trace is present.

(45) \[ \text{[CP why}_j \text{ did [IP Bill buy what]]} \]

Just like the wh-in-situ in (1b), the wh-in-situ in (45) has no scope marker since “only a c-commanding wh-chain is a scope marker for wh-in-situ” (see (35a)). In (45) why is not a wh-chain c-commanding what (since there is no variable). Thus, there is no wh-chain c-commanding what at S-Structure, with the result that the SMC disallows the required LF movement.\(^{17}\) Given the S-Structure representation (45), we obtain the unified analysis with (1b). But is (45), containing no wh-adjunct trace, in fact the S-Structure representation of (38)? It is, given the following independently motivated assumption (see Epstein 1987, 1991 and L&S 1992 for ways to deduce (46)): (46) S-Structure representations contain no traces of wh-adjuncts.

This assumption is independently motivated by (among other phenomena) the fact that (47) is better than the ECP violation (48) (see L&S 1984, 1992).\(^{18}\)

(47) *who, do you wonder [\text{CP whether [IP John said [CP t}_1 \text{ [IP t}_i \text{ left]]]}]

\(^{16}\) Rizzi (1990:47) suggests that reason adverbials are adjoined to either TP or AgrP.

\(^{17}\) I assume that a trace of why appears in the LF component. Given the existence of such a trace, in LF there will be a chain c-commanding what, unlike at S-Structure. This is the only empirical motivation I know of for (35b) = “scope marking is determined at S-Structure,” hence the only impediment to reformulating the SMC as a purely LF condition, consistent with the leading idea (Chomsky 1986, 1993) that there are no S-Structure conditions. If wh-adverbials like why in fact bind no trace in LF (see Rizzi 1990:47) or if, alternatively (perhaps by Procrastinate), once the trace of why is created, the LF representation has by definition been created (hence, wh-in-situ, by definition, cannot move), then scope marking could take place entirely in LF; that is, the S-Structure condition (35b) could by hypothesis be eliminated.

\(^{18}\) As L&S (1984) note, (47) is certainly better than the ECP violation (i).

(i) **who, do you wonder \[\text{CP whether [IP John said [CP t}_1 \text{, that [IP t}_i \text{ left]]]}}]
(48) **why_i do you wonder [CP whether [IP John said [CP t′_i [IP Sue left t_i]]]]

Informally, the contrast in grammaticality indicates that ‘‘short’’ movement (i.e., movement into a ‘‘local’’ [Spec, CP]) followed by ‘‘island-hopping’’ movement is worse with an adjunct (48) than it is with a subject (47). To account for the contrast, L&S (1984, 1992:52) formalize proper government (see (16a)) as feature assignment (49a–b) constrained by the stipulation in (49c).

(49) a. A properly governed trace is assigned [ + γ].
    b. A trace that is not properly governed is assigned [ − γ].
    c. Only an argument is assigned a γ-feature at S-Structure.

Under (49), (47) escapes the ECP as follows: at S-Structure, t′ (or, in conformity with (16), C⁰) assigns [ + γ] to t. t′, being a nonargument, receives no γ-feature at S-Structure and subsequently deletes in LF.¹⁹ The ECP, formulated as the following filter, is satisfied (regardless of its level of application):

(50) ECP
    *t
    [ − γ]

An analogous (ECP-satisfying) derivation of (48) is blocked by the stipulation (49c); that is, in (48) the nonargument t cannot be assigned [ + γ] at S-Structure (with subsequent deletion of t′ in LF). Rather, (49c) dictates that γ-assignment to the nonargument t in (48) cannot occur at S-Structure. Therefore, γ-assignment must be postponed until the LF level, assuming (51) (L&S 1984:92).

(51) γ-assignment occurs at levels of representation only.

At the LF level, t′ must be present to assign a γ-feature to t. But t′ then violates the ECP: it is neither lexically properly governed (16c) nor antecedent-governed (11). Thus, (49c) yields the desired results: (47) escapes the ECP, but (48) violates it precisely because the nonargument t is not allowed to receive a γ-feature at S-Structure. But the question that arises with respect to (48) is, Why is the adjunct trace invisible to γ-assignment at S-Structure? One truly explanatory answer is, Because it is not present at S-Structure. In order to maintain such an explanatory answer, I will therefore assume that the descriptive generalization in (46) holds.

To summarize, there is independent motivation concerning the ECP for assuming that traces of adjuncts are not present at S-Structure (46), the truth of which entails that (45) is indeed the S-Structure representation of (38). Given (45), the results we sought are obtained: the deviance of (38) is explained by the SMC analysis, and therefore a unified treatment of (38) and the classic

¹⁹ Chomsky (1991:441) deduces such trace deletion from Full Interpretation. That is, (whο_t′_t) is neither an operator-variable chain, nor a homogeneous/uniform Λ-chain, nor an argument chain. Deleting t′ yields the legitimate LF object (who_t), an operator-variable chain. See Browning 1987 for an earlier discussion of such chain uniformity.
superiority case (1b) is provided (i.e., (38) and (1b) receive identical analyses, as do (7a) and (42)). More generally, (46) entails (52).

(52) At S-Structure, a wh-adjunct occupying [Spec, CP] is neither a chain nor a chain member; hence, it never functions as a scope marker for wh-in-situ.

Having provided independent support for the SMC, formulated as a movement constraint, I turn to the formal mechanism, LF I-raising, whereby long-distance LF subject movement, forced by the SMC (see (7)), can satisfy the ECP at LF. I will present evidence favoring an analysis in which I undergoes movement at LF over one postulating I-in-situ at this level. The remainder of the article is concerned with determining the precise properties of such LF I-movement.

20 The data presented in this section are accounted for, as are the following types of examples:

(i) *Why did who fix the car? (*SMC analysis)
(ii) **Why did John fix the car how? (*SMC, *ECP)
(Huang 1982a:557) judges the following perfect, hence better than (38), to which he assigns ‘?’:
(iii) Tell me why you bought what.
No account for the purported contrast is given (see Huang 1982a:585–586). I am not sure that I find (iii) any better than (38), but if it is, the SMC might provide an explanation. Example (iii) is an imperative, the felicitous use of which is constrained by certain discourse factors. It may be, then, that (iii) forces or biases for a D-linked interpretation of what in the sense of Pesetsky (1987). If what is D-linked, then, under Pesetsky’s analysis, it does not undergo LF movement; hence, the SMC is irrelevant and lack of a scope marker in (iii) does not induce ungrammaticality. The inapplicability of the SMC to D-linked wh-phrases would also account for the suppression of superiority effects (noted by Pesetsky) in analogues to (1b) such as (iv).
(iv) Which woman did which man see?

21 The SMC is a constraint on, and hence presupposes the existence of, LF wh-movement. However, the SMC could be rendered compatible with an analysis in which wh-in-situ does not undergo LF wh-movement but is instead subject to an LF interpretive rule of scope assignment. That is, the SMC could be reformulated as the following interpretive (not movement) constraint:

(i) In the LF component, interpret a wh-in-situ W as having scope identical to that of the head of a wh-chain CH, where CH c-commanded W at S-Structure.

However, I will suggest below that certain directly relevant data can readily be explained under the LF wh-movement analysis presumed here, but not under an interpretive (nonmovement) reanalysis of wh-in-situ. Consequently, I will continue to assume LF wh-movement and therefore will retain the original SMC formulated as a constraint on LF wh-movement. The interpretive constraint (i) is not incorporated.

To begin with, consider the raising structure (ii).

(ii) [IP John i [AP likely [IP ti to win]]
(iii) [CP [AP how likely ti to win][CL ti][IP John, ti, tj]]

(iii) is an example of what might be called “the least controversial case of a predicate-internal subject”; that is, John originates internal to the ‘predicate’ AP and moves to [Spec, IP]. In (iii) the subject John has similarly raised from the position ti to [Spec, IP]; and, in addition, the AP has undergone wh-movement to [Spec, CP]. I will assume that ti, the trace of subject raising, is an anaphor. Applying the chain-binding algorithm proposed by Barss (1986, 1988), I further assume that, even though this anaphor is free at S-Structure, it is nonetheless licit by virtue of being chain-bound by its antecedent John. Informally, ti is licit because John is a local c-commander of the trace tj, the trace of the container of the anaphor. (I in fact depart here from Barss’s analysis of (ii), which assumes that tj is not anaphoric and is therefore not subject to the binding theory; see Barss 1986:409.)

Now consider the following minimal pair:

(iv) [CP who i [IP ti wonders [CP [AP how likely ti to win][IP John, ti, tj]]]
(v) *[CP who i [IP ti wonders [CP [AP how likely ti to win][IP who, ti, tj]]]
5 I as a Proper Governor at LF

How can the LF representation (23), derived by long-distance LF subject movement, in effect “forced” by the SMC, satisfy the ECP? Recall that L&S (1992) propose that subject position (coindexed with I via specifier-head indexing) becomes antecedent-governed at LF by virtue of (the disjunctive) (27).

(27) In the LF component, I can substitute into C or can adjoin to IP.

Thus, there are two well-formed LF representations associated with the S-Structure representation (7b), namely, (53) and (54).
Given (27), there should be no ECP effects with subject wh-in-situ. But consider the (I believe previously unnoted) contrast between the following two cases of embedded topicalization:

(55) [CP who, [IP ti thinks that [IP this problem, [IP John solved tj]]]]

(56) *[CP who, [IP ti thinks that [IP A who, I solved tj]]]

Following Baltin (1982) and L&S, I assume, as indicated in (55) and (56), that Topicalization is S- (= IP-) adjunction.

Regarding the contrast in (55) and (56), I should first note that some speakers find embedded topicalization, as in (55), somewhat marginal. Nonetheless, (56) seems to be significantly less grammatical. Importantly, the ungrammatical (56) differs only ‘minimally’ from L&S’s grammatical example (7). In (7) the object undergoes syntactic wh-movement to [Spec, CP]. In (56), by contrast, the object undergoes topicalization (not wh-movement); the result is strongly ungrammatical and arguably violates the ECP.22 (Clearly, the SMC is not implicated: obligatory LF adjunction of whok to the c-commanding wh-chain headed by whoi is licensed.)

How, then, can the ill-formedness of (56) be explained? I will (tentatively) adopt the LF I-
movement analysis postulated by L&S, but I will argue that their disjunctive analysis (27) must be reformulated in order to account for (56).\footnote{Tiedeman (1990) provides an alternate analysis, which, however, cannot account for the ungrammaticality of (56). For discussion, see Epstein 1993.}

First, with respect to (56) itself, I assume (with L&S and contra May (1985)) that each of the IPs in this representation, $IP_A$ and $IP_B$, is a distinct maximal projection. In other words, I assume (58).

(58) Adjunction to an $X^{\text{max}}$ yields two distinct maximal projections, as opposed to a single segmented maximal projection.

Thus, the more precise representation of (56) is (59).

(59) $\ast[CP \ \text{who}_i \ [IP \ t_i \ \text{thinks that} \ [IP_A \ \text{who}_k \ I_k \ \text{solved} \ t_j]]]$  

Evidence for the existence of two separate maximal IPs in topicalized structures is provided by the following contrast adapted from L&S 1992:96:

(60) a. $[CP \ \text{who}_i \ [IP \ t_i \ \text{likes this book}]]$  
   b. $\ast[CP \ \text{who}_i \ [IP \ \text{this book} \ j \ [IP \ t_i \ \text{likes} \ t_j]]]$  

Under May’s (1985) segmentation theory of adjunction, the “distance between” $t_i$ and its antecedent (who$_i$) is identical in (60a) and (60b); that is, in each case there is exactly one maximal projection, IP, dominating $t_i$ and excluding who$_i$. The grammaticality contrast is thus (arguably) unexplained. By contrast, under a (standard) duplication theory of (Chomsky-) adjunction, there is only one IP “separating” $t_i$ from who$_i$ in (60a), whereas there are two IPs separating $t_i$ and who$_i$ in (60b). Exploiting precisely this distinction provided by the standard, duplication theory of adjunction they adopt, L&S analyze (60b) as a violation of (Subjacency and) the ECP. As in the standard theory of (English) bounding nodes, their analysis entails that two (distinct maximal projection) IPs form a barrier for Subjacency (and for antecedent government). Thus, the “island-forming” property of topicalization is accounted for. The relevant definitions from L&S 1992 (pp. 74, 87, 102, 183) are as follows (supplementing those cited in (11) and (16)):

(61) B is subjacent to A if for every C, C a barrier for B, C $m$-commands A.

(62) A $m$-commands B iff the minimal maximal projection dominating A dominates B. (See Aoun and Sportiche 1983.)

(63) C is a barrier for B if
   a. C is a maximal projection, and  
   b. C is not an $\hat{A}$-binder, and  
   c. C is not L-marked, and  
   d. C dominates B. (See also Chomsky 1986.)

(64) A $L$-marks B iff A is a lexical category that $\theta$-governs B. (See Chomsky 1986.)
(65) A $\theta$-governs B iff A is an $X^0$ that $\theta$-marks B, and A and B are sisters. (See Chomsky 1986.)

Under this analysis, (60a) is predicted to be grammatical. By contrast, (60b) violates Subjacency and the ECP. The lower IP in (60b) is a barrier for $t_i$ (i.e., it is not L-marked, it dominates $t_i$, it is not an $\tilde{A}$-binder, and, under the duplication theory of adjunction, it is a maximal projection in its own right). Since the lower IP is a barrier for $t_i$ and it fails to m-command who$_1$ (precisely because the upper IP is also a maximal projection, given “duplication adjunction”), $t_i$ is not subjacent to who, violating Subjacency. Since this lower IP similarly fails to m-command C$_0$ (which, under specifier-head indexing, binds $t_i$), $t_i$ also fails to be antecedent-governed (= subjacent to a binding head) and thereby violates the ECP.

With this analysis of adjunction, Subjacency, and the ECP in hand, let us return to (59). Recall, given (27), subject wh-in-situ should not give rise to ECP violations, yet (59) is, by hypothesis, precisely such an example. Under (27) and the duplication theory of adjunction, one possible LF representation of (59) is (66); that is, who$_k$ undergoes long-distance LF movement, and I$_k$ adjoins to IP$_A^{\text{max}}$, creating a second (duplicate) IP$_A^{\text{max}}$.

\[
(66) \quad [\text{CP who}_k \text{who}_1 [\text{IP } t_i \text{ thinks } [\text{CP that } [\text{IP}_X \text{ this problem} \text{ IP}_X \text{ I}_k \text{ I}_k \text{ t}_k \text{ t}_k \text{ solved } t_j]]]]
\]

Later I will point out another logically possible LF representation of (59), but (as I will show) it also yields the wrong result; that is, like (66), it does not violate the ECP. Suppose, following L&S (all of whose evidence, I believe, concerns syntactic adjunction), that (67) is correct.

(67) Syntactic adjunction to $X^{\text{max}}$ creates a second $X^{\text{max}}$.

As discussed, this accounts for the contrast between (60a) and (60b). At the same time it could well be the case, following May (1985) (all of whose evidence, I believe, instead concerns LF adjunction), that (68) holds.\(^{24}\)

(68) LF adjunction to $X^{\text{max}}$ segments the $X^{\text{max}}$; that is, it does not create a second $X^{\text{max}}$.

Thus, contra standard assumptions, I believe there is in fact no incompatibility between the theory of adjunction proposed by L&S and that proposed by May (1985). In other words, I tentatively suggest that the former (duplication) characterizes syntactic adjunction, whereas the latter (segmentation) characterizes LF adjunction.\(^{25}\)


\(^{25}\) Poole (1996) points out that if Chomsky’s (1993) extension version of the strict cycle constrains not only substitution but also XP-adjunction (see Chomsky 1993:23), then syntactic segmentation adjunction would be disallowed, since a dominator of the node targeted by the generalized transformation GT is not created. Thus, Poole notes, duplication adjunction is forced by the extension version of the strict cycle in the syntax. If, as Chomsky (1993) proposes (cf. Kitahara 1993b), the extension version of the strict cycle does not apply at LF, then segmentation adjunction would be allowed (but not ensured), as Poole further notes.
Even if this “mixed theory of adjunction” is correct, it provides no solution to the problem of capturing (59) as an ECP violation (at LF). That is, even if LF I-adjunction segmented IP$_A$ (as opposed to duplicating it, as indicated in (66)), there would still be no violation.

\[(69) \text{[CP who$_k$ who$_i$ [IP$_t$ thinks [CP that [IP$_{\text{max}}$ I$_k$ [IP$_{\text{max}}$ t$_k$ t$_i$ solved t$_j$]]]]]}
\]

In (69), as in (66), t$_k$ is antecedent-governed—the wrong result. That is, in both (66) and (69) the minimal barrier dominating the subject t$_k$ is IP$_A^{\text{max}}$. Since IP$_A^{\text{max}}$ (in both representations) m-commands I$_k$, t$_k$ is (by definition) subjacent to this binding head; equivalently, t$_k$ is antecedent-governed by I$_k$, thereby satisfying the ECP.

How can (66) and (69) be blocked? A natural proposal is that I$_k$ cannot be adjoined to IP$_A$. I therefore propose that the constraint on adjunction shown in (70) (the relevant aspects of which will be deduced below) prevents I$_k$ from adjoined to IP$_A$ in (56). (The question of whether such adjuction (now prohibited) would segment or duplicate IP$_A$ disappears.) Hence, (66) and (69) are not generable.

\[(70) \text{Affect } \alpha, \text{ applying to the adjunction configuration}
\]
\[
\ldots [X_P, [X_P, \ldots] \ldots
\]
\[
cannot \text{ adjoin a category to } X_P.
\]

Given (70), when the S-Structure representation (59) enters the LF component, LF I-adjunction to the maximal projection IP$_A$ is prohibited. Consequently, the ECP-satisfying LF representations (66) and (69) cannot be derived from (59). Rather, under (70), the “closest” that I$_k$ can get to the subject trace (while binding it) is a position adjoined to the “outer” IP—namely, IP$_B$, the IP (maximal projection) created by syntactic topicalization. In order to have the worst possible chance of succeeding in analyzing this as an ECP violation, let us assume the mixed theory of adjuction under which LF I-adjunction to IP$_B^{\text{max}}$ results in the segmentation of IP$_B^{\text{max}}$ (putting I$_k$ closer to t$_k$ than it would be if each IP$_B$ were a distinct maximal projection, as would be the case under the duplication theory of adjunction applying at LF). This yields the following LF representation, derived by long-distance LF movement of who$_k$ and segmentation adjunction of I$_k$ to IP$_B^{\text{max}}$.

\[(71) *[CP who$_k$ who$_i$ [IP$_t$ thinks [CP (that) [IP$_{\text{max}}$ I$_k$ [IP$_{\text{max}}$ this problem$_j$ [IP$_{\text{max}}$ t$_k$ t$_i$ solved t$_j$]]]]]]
\]

As desired, (71) violates the ECP. The subject t$_k$ is not subjacent to (hence not antecedent-governed by) the binding head I$_k$. Recall the definitions given earlier, repeated in (72).

\[(72) a. \text{B is subjacent to A if for every C, C a barrier for B, C m-commands A.}
\]
\[(72) b. \text{A m-commands B iff the minimal maximal projection dominating A dominates B.}
\]

26 I leave open the possibility suggested by Howard Lasnik (personal communication) that (70) might be derivable from the A-over-A Constraint discussed in Chomsky 1972:51.
For the subject trace $t_k$ to be antecedent-governed in (71), it must be subjacent to a binding head. The only head binding $t_k$ (under c-command) in (71) is $I_k$. In order for $t_k$ to be subjacent to $I_k$, every barrier for $t_k$ must m-command $I_k$. The maximal projection IP$_\lambda$ is a barrier for $t_k$. But this barrier does not m-command $I_k$: the minimal maximal projection dominating IP$_\lambda$ is the bisegmental maximal projection IP$_B$, but this bisegmental maximal projection IP$_B$ does not dominate $I_k$, given (73).

(73) $A$ is dominated by $B$ only if it is dominated by every segment of $B$. (Chomsky 1986)

Thus, the subject $t_k$ is not antecedent-governed (nor is it lexically properly governed). The ECP is therefore violated, as desired, and we now have an account of the ungrammaticality of cases like (56).

6 Syntactic versus LF I-Raising within Checking Theory

In the preceding sections I have investigated in detail some consequences of LF I-raising, in particular the resulting antecedent government of subject position by I at LF (L&S 1992). Antecedent government of subject position by virtue of LF I-raising must somehow be restricted to the LF component since subject position is not in fact antecedent-governed at S-Structure, even when syntactic I-raising applies. This is shown in Rizzi 1990 by, for example, the ill-formedness of a matrix wh-question displaying both I-to-C movement and heavy NP shift of the subject.

(74) *[CP where [L$_t$ has]] [IP $t$, $t$, gone][NP the man from Cincinnati]]

The ill-formedness of (74) is not predicted by the analysis assumed here; that is, (74) satisfies the ECP. Given S-Structure specifier-head indexing (see (8)), the subject trace and the I-trace are coindexed. Consequently, I and the subject are coindexed. The subject trace is therefore antecedent-governed by I$_i$ in C. How can it be that LF I-raising renders the subject position antecedent-governed but syntactic I-raising does not?

In Chomsky’s (1991, 1993) checking theory framework, there is a crucial structural difference between syntactic I-raising in English, as in (74), and LF I-raising. (The latter, I will argue, is in

---

27 The trace of I-movement in (71) also violates the ECP since it, like the subject trace, is not subjacent to $I_k$. Hence, it fails to be antecedent-governed. Although L&S (1992) do not discuss X$^0$-movement with respect to the ECP, notice that I-to-C movement ‘over’ topicalization as in (i) is ungrammatical and is similar to L&S’s core case of subject extraction (ii).

(i) *[cp[cl will]] [IP this book, [IP John, buy $t_1$]]

(ii) *[cp who, [c _]] [IP this book, [IP, bought $t_1$]]

This suggests that L&S’s formulation of the ECP might in fact extend at least partially to head movement (see Epstein 1990b for discussion). But I will not pursue this issue here.

28 Some potentially interesting consequences of the proposed analysis regarding the application of subject Quantifier Raising in LF are explored in Epstein 1993.

29 See also Koopman 1983, where (i) is analyzed as an ECP effect (not unlike (74)).

(i) *Who did see Mary?
fact covert verb-second.) I believe this independently motivated structural difference between the
syntax and LF explains, in a principled way, the contrast in antecedent government.

6.1 Background Assumptions: Checking Theory

First, regarding the category I have been calling “I,” Chomsky (1993:7) (elaborating proposals
in Pollock 1989) proposes the clausal structure (75), in which there exist at least three inflectional

(75) \[ CP C \{ AgrP \ A grS \} [ TP T \{ AgrP \ A grO [ VP V] ] ] \]

Regarding the movement of these three inflectional heads, Chomsky makes the (nonstandard)
assumption in (76).

(76) There is no affix lowering (i.e., Agr-lowering or T-lowering) onto the verb.

Rather, the verb exits the lexicon fully inflected, bearing agreement and tense features. However,
these features must be checked. An inflectional feature of a verb (e.g., the verb’s tense feature)
can be checked only if the verb becomes a member of the checking domain of T, which it does
if it adjoins to T. Thus, agreement and tense checking is achieved in structure (75) by adjoining

---

30 The relevant definitions from Chomsky 1993 are as follows:

(i) The domain of a head \( \alpha \) = the set of nodes contained in Max(\( \alpha \)) that are distinct from and do not contain \( \alpha \).
(p. 11)

a. A category \( \alpha \) dominates \( \beta \) if every segment of \( \alpha \) dominates \( \beta \). (p. 11)

b. Max(\( \alpha \)) = the least full-category maximal projection dominating \( \alpha \). (p. 11)

c. A category \( \alpha \) contains \( \beta \) if some segment of \( \alpha \) dominates \( \beta \). (p. 11)

Two subsets of the domain of a head \( \alpha \) are then defined: the complement domain of \( \alpha \) and the residue of \( \alpha \).

(ii) The complement domain of \( \alpha \) = the subset of the domain reflexively dominated by the complement of the
construction. (p. 11)

a. A category \( \alpha \) reflexively dominates all categories dominated by \( \alpha \), and \( \alpha \) itself.

b. Complement of the construction is not defined in Chomsky 1993.

(iii) The residue of \( \alpha \) = the domain of \( \alpha \) minus the complement domain of \( \alpha \). (p. 11)

As a simple illustration of the three sets just defined, consider (iv).

(iv) 

\[
\begin{align*}
XP \\
\quad ZP \\
\quad \quad X' \\
\quad \quad \quad Z' \\
\quad \quad \quad \quad X = \alpha \\
\quad \quad \quad \quad \quad YP \\
\quad \quad \quad \quad \quad \quad Z \\
\quad \quad \quad \quad \quad \quad \quad Y' \\
\quad \quad \quad \quad \quad \quad \quad \quad Y
\end{align*}
\]
the verb to the inflectional heads, AgrO, AgrS, and T (rendering the verb a member of the checking domain of each), that is, by "V-raising." There is, of course, parametric variation concerning the level at which V-raising (checking) occurs. For example, in French the verb raises (i.e., inflectional checking is performed) in the syntactic component, whereas in English this happens at LF.31

This picture is however complicated by the fact that, in English, certain verbs raise in the syntax (finite have/be) whereas main verbs raise in LF. Why should this be? Chomsky (1993: 24) assumes the following:

(77) In general, English performs V-raising in LF; that is, inflectional features are checked at LF.

\[ a. \ Dom(\alpha) = \{ZP, Z', YP, Y, Y'\} \]
\[ b. \ Dom(\alpha) = \{YP, Y', Y\} \]
\[ c. \ Dom(\alpha) = \{ZP, Z, Z'\} \]
\[ d. \ Dom(\alpha) = \{ZP, Z, Z'\} \]

31 This analysis is attractive in a number of respects. First, there is no affix lowering, hence no free trace at S-Structure (cf. Chomsky 1991). Second, the analysis is consistent with the leading idea that natural language grammars incorporate the same instantiations of Affect \( \alpha \) (in this case, V-raising) but differ with respect to the level at which these operations apply. Third, the analysis permits a French sentence and an English sentence that are synonymous to have structurally identical LF representations, a natural if not necessary result that is precluded if French has V-raising and English has affix lowering.
Why, then, do *have/be* exhibit the exceptional property of undergoing syntactic raising? Following a long tradition within which it is assumed that the “deficient” semantics of these verbs are responsible for their undergoing syntactic movement (see, e.g., Pollock 1989, Chomsky 1993, and the references cited there), Chomsky (1993:31) assumes LF Movement Visibility (78).

(78) **LF Movement Visibility**

In the LF component, semantically deficient (underspecified) categories are not visible to Move $\alpha$; hence, they cannot be moved by Move $\alpha$.

Thus, finite *have/be* cannot be moved in the LF component. That these verbs raise in the syntax now follows from (78) and the above assumption that checking of the verb’s inflectional features is obligatory and can be achieved only by raising the inflected verb and adjoining it to the functional inflectional head. But this raises one other question. Why is checking obligatory? Chomsky (1993:27) proposes the following explanation. Full Interpretation (FI)/economy of representation requires (naturally) that every symbol in an LF representation have an interpretation. Agr and (I will assume) T are semantically vacuous. They therefore must delete in order for a legitimate LF representation to be generated.32 Suppose now that (79) obtains.

(79) Functional categories can be deleted only after performing their checking.

It now follows that checking (achieved only in the configuration generated by V-raising) must occur at some point prior to LF; that is, raising/checking is obligatory. Notice, however, that (79) leaves unspecified the precise point in a convergent derivation at which a functional category (e.g., Agr$_S$, Agr$_O$, T) checks and deletes. (This will be crucial below.) Regarding this issue, Chomsky in effect assumes that economy of representation holds at intermediate levels as well, entailing that (79) must be reformulated as (80).

(80) Functional categories can be deleted only after performing checking, and they must delete immediately after performing checking (since they have no function other than checking).

---

32 T might not be semantically vacuous (see Chomsky 1991 regarding *finite* T). I nonetheless assume that T$^0$ is absent from LF representations. This is deducible from FI/economy of representation, which would prohibit an LF representation containing, redundantly, both the tense features lexically specified on the verb and the same features on the functional checking category T$^0$, as in (i), for example.

(i) \[ [cp \ldots [ T \ V ] \ldots ] \]

\[ [+ \text{past}] [+ \text{past}] \]

Given that a sentence has only one tense, (i) cannot be an LF representation. As demonstrated in Epstein 1987, 1991, features of a head cannot be deleted (e.g., the masculine feature of *himself*, the [+ γ] feature on a trace, the tense feature on the verb). Rather, only entire syntactic categories are subject to deletion. In (i) V cannot delete since it has semantic content not recoverable from T. However, T can delete since its feature is fully recoverable from the verb’s (identical) tense feature. If T can delete, it follows that it is an unnecessary symbol in the LF representation. By economy of representation, then, it must delete.
6.2 A Difference between Syntactic V-Raising and LF V-Raising

One of the central contrasts Chomsky (1991:sec. 3, 1993:30) seeks to explain is that syntactic V-raising can violate the Head Movement Constraint (HMC) but LF V-raising cannot. This is illustrated by the following S-Structure representations:

(81) \[ \text{[AgrP John has, ... [NegP not t, seen Mary]]} \]

(82) \[ \text{[AgrP John ... [NegP not see s Mary]]} \]

In (81) syntactic movement of has violates the HMC; that is, movement over the intervening head of NegP (not) has occurred. But the ECP is (somehow) satisfied. In (82) the (lexically inflected) verb see + s must raise in LF in order to have its inflectional features checked and thereby effect the deletion of the functional heads, as required by FI/economy of representation. For some reason, this LF movement, which would seem to correspond exactly to the licit syntactic movement in (81), violates the ECP. How is the contrast between syntactic and LF V-raising to be explained?

Consider first the case of syntactic raising in (81). The D-Structure representation of (81) is (83).

(83) \[ \text{[AgrP John AgrS [TP T [NegP Neg [AgrO, AgrO [VP has [VP seen Mary]]]]]} \]

In order to have its inflectional features checked, has must raise in the syntax, since it is invisible to LF Move $\alpha$ (see (78)). First, it adjoins to AgrO, creating $[\text{AgrO, V has} + \text{AgrO}]$. Even if the agreement features in has are now checked by AgrO, AgrO does not delete at this point since it has not in fact completed all of its checking. (Chomsky assumes that AgrO is also involved in checking the accusative Case feature with which the NP Mary exits the lexicon. That is, he assumes that NPs exit the lexicon bearing structural Case (e.g., accusative), just as Vs exit the lexicon bearing inflectional features. In the LF component, Mary moves to [Spec, AgrO], (= “object shift”), in which position the category $[\text{AgrO, V + AgrO}]$ (or, more precisely, the trace of this category) can check accusative Case (in a specifier-head configuration.) Next, in order to check the tense features on has, the category $[\text{AgrO, has + AgrO}]$ must adjoin to T. But this violates the HMC because the head Negation (Neg$^0$) intervenes. However, this movement is movement of AgrO; hence, it leaves an AgrO-trace in the head position of AgrO$^P$. It is this trace that would appear to violate the ECP. However, recall that after LF object shift of Mary to [Spec, AgrO], AgrO (or, more precisely, the trace of AgrO) will check accusative Case on Mary, thereby completing its checking of both inflectional and Case features; it will consequently delete automatically. The offending trace is therefore absent from the LF representation, and the ECP (applying only to LF; see, e.g., Epstein 1987, 1991) is therefore satisfied. Hence, checking theory explains why syntactic V-raising over negation is allowed.

33 I omit VP-internal subject raising, which is irrelevant here.
But then, why isn’t the same (ECP-satisfying) movement allowed in LF in (82)? (That is, why does LF movement in (82) violate the ECP?) I will now propose an analysis within which this corresponding movement is in fact precluded in LF by an independently motivated principle, with the result that the ECP is violated. The articulated D-Structure representation of (82) is (84).

(84) \[ \text{[Agr}_P \text{ John Agr}_S [\text{TP T [NegP Neg [Agr}_O \text{P Agr}_O \text{ [VP see + s Mary]]]]]}
\]

In the LF component, \text{see + s} first adjoins to \text{Agr}_O, yielding \text{[Agr}_O \text{ see + s + Agr}_O]. In this configuration the agreement features of \text{see + s} are checked by \text{Agr}_O. We must now somehow prevent what happened in the well-formed syntactic derivation from happening in LF: namely, raising of \text{[Agr}_O \text{ see + s + Agr}_O] to T (violating the HMC) and subsequent deletion (by virtue of object shift) of the offending \text{Agr}_O-trace left by this movement. I propose that movement of \text{[Agr}_O \text{ see + s + Agr}_O], although allowed in the syntax, is in fact prevented in LF by an independently motivated principle. Under LF Movement Visibility, the \text{Agr}_O category \text{[Agr}_O \text{ see + s + Agr}_O] cannot be moved. Since \text{Agr}_O is a semantically vacuous category, Move \text{a} applying in LF cannot see it, hence cannot move it. If no further movement applies, FI/economy of representation is violated since, without further raising, the functional heads T and Agr_S cannot check the tense and agreement features in \text{see + s} and therefore cannot delete. These illegitimate LF objects would consequently be present in the LF representation, but FI/economy of representation disallow this. In other words, this derivation crashes at LF.

There is one other derivation of the ungrammatical (82) that must be blocked. Suppose that, after \text{[V see + s]} adjoins to \text{Agr}_O in LF, yielding the movement-	extit{invisible} \text{[Agr}_O \text{ see + s + Agr}_O], object shift of \text{Mary} to [Spec, Agr_O] occurs. Given this, \text{Agr}_O (to which the V is adjoined) can mediate checking of accusative Case on \text{Mary}. Since \text{Agr}_O has now checked both the (verbal) agreement features in \text{see + s} and the (nominal) accusative Case on \text{Mary}, its checking is completed and it therefore automatically deletes. I assume that deletion of \text{Agr}_O results in the “excision repair” (sub)structure (85) in which \text{Agr}_O is absent and V comes to be immediately dominated by \text{Agr}_O'; that is, the deletion of \text{Agr}_O, to which V was adjoined, puts V in the head of \text{Agr}_OP (see Epstein 1993).

(85)
```
     Agr_O
    /   |
   NP   Agr_O'
      /    |
Mary (acc) V
     /     |
    see+s
```
In general, I will assume (crucially) that if a head Z adjoins to a head Y, as in (86a), and Y subsequently deletes, then the derived constituent structure is (86b), in which Z occupies the head position of YP; that is, Z is immediately dominated by Y′ (just as it was prior to Y’s deletion).  

(86) a. \[
\text{YP} \\
\text{Y′} \\
\text{Y} \\
\text{Z}
\]

Crucially, (86b) is inconsistent with X-bar theory (Y′ fails to immediately dominate Y0), an important fact that I will take up below. For the moment I return to (85), derived by (a) V-to-AgrO movement, (b) object shift of Mary, and (c) AgrO-deletion. Given AgrO-deletion, which has produced the (intermediate) representation in (85), the V, now in the head position of AgrO P, is a category that is indeed visible for movement by LF Move a. So that its tense features can be checked, the V now adjoins to T. Crucially, this movement of a V leaves a V-trace.

(87) \[
\text{AgrS P John . . . [see+s + T] [NegP not \text{AgrO P Mary [AgrO [tV . . . ]]]]}
\]

The V-trace, unlike an AgrO-trace, has semantic content (it is, or is part of, a legitimate LF object) and is therefore not deletable (Chomsky 1991). The ECP is therefore violated and the ungrammaticality of cases like (82) is explained.  

The contrast between syntactic V-raising over negation (81) and LF V-raising over negation (82) is accounted for. Syntactic raising is, in fact, movement of the deletable functional category AgrO. By contrast, the independently motivated and natural principle of LF Movement Visibility dictates that LF raising must be raising of the nonfunctional (undeletable) lexical category V. In what follows, I will exploit this crucial distinction.

6.3 An Answer

I am now in a position to answer the question with which this section began: How is it that LF I-raising renders the subject position antecedent-governed but syntactic I-raising does not? The analysis depends crucially on the independently motivated presence of a functional category in the syntax, and on its absence at LF. I will begin with the question of why syntactic I-to-C raising

---

34 Notice that, under May’s (1985) segmental theory of adjunction, the immediate dominator (see (73)) of Z is Y′, in both (86a) and (86b).

35 For a treatment adopting aspects of the analysis presented here, see Kissock 1995:chap. 4.

36 Notice that, in the analysis just presented, order of “rule” application, stipulated in the standard theory (entailing concomitant learnability problems), is deduced from the presumably universal principle of LF Movement Visibility, an independently motivated and arguably natural (as opposed to “purely formal”) principle of the LF component.
does not cause the subject to be antecedent-governed in (74). Recall that syntactic V-raising first adjoins the V to AgrO, yielding \([\text{Agr}_O \text{V has}] + \text{Agr}_O\). Recall also that further syntactic raising leaves an \(\text{Agr}_O\)-trace in the head of \(\text{Agr}_O\)P. This trace does not delete in the syntax, but only in LF. I assume that \(\text{Agr}_O\) itself (to which the V \text{has} is adjoined) deletes only when the \(\text{Agr}_O\)-trace is deleted in LF. The fact that \(\text{Agr}_O\) itself is not deleted before LF and hence is not deleted in the syntax is deducible from the Proper Binding Condition, which, following L&S, I will assume applies to the output of each application of Affect \(\alpha\). That is, if \(\text{Agr}_O\) itself deleted prior to LF, the \(\text{Agr}_O\)-trace (occupying the head position of \(\text{Agr}_O\)P) would be unbound, violating the Proper Binding Condition. Thus, since the \(\text{Agr}_O\)-trace does not delete until LF, \(\text{Agr}_O\) itself does not delete until LF. It now follows that in the syntax the raised V is still adjoined to \(\text{Agr}_O\).

Now, I also assume that after syntactic adjunction of V to \(\text{Agr}_O\), the resulting \([\text{Agr}_O \text{V has}] + \text{Agr}_O\) next adjoins to AgrS. As in Chomsky 1993, AgrS has T adjoined to it in the syntax. This happens because T (having strong N-features) must check Case on an NP \text{in the syntax} and can do so only if adjoined to AgrS.\(^{37}\) Thus, in the syntax, T adjoins to AgrS and then checks the Case on an NP occupying [Spec, AgrS].\(^{38}\) That is, under the VP-Internal Subject Hypothesis, the subject must raise in the syntax to [Spec, AgrS] so that T can do what it must, namely, check Case on an NP in [Spec, AgrS] in the syntactic component.

Thus, in the part of the derivation of (74) shown in (88), \([\text{Agr}_O \text{V has}] + \text{Agr}_O\), occupying the head of \(\text{Agr}_O\)P, now adjoins to AgrS (which has T adjoined to it).

\(^{37}\) See Jonas 1992 and Jonas and Bobaljik 1993 for arguments that T-to-AgrS movement is subject to parametric variation and that in (e.g.) Icelandic, T-in-situ can check nominative Case on an NP occupying [Spec, TP], an operation underlying transitive expletive constructions, apparently unavailable in English.

\(^{38}\) The obligatoriness of Case checking by T of an NP occupying [Spec, AgrS] constitutes, in large part, the Extended Projection Principle.
Notice that movement b. over the trace of T violates the HMC. However, the ECP (a constraint on LF representations) will not be violated, since the offending trace of AgrO deletes in LF. Finally, I-to-C (i.e., AgrS-to-C) movement applies to (88), moving the entire AgrS complex to C. 39

Movement of the AgrS complex to C in (88) yields the S-Structure representation in (89).

(89) \[ \text{CP where } C \{ \text{AgrS} P, t_i \ldots \text{gone} \} \{ \text{the man from Cincinnati} \} \]

\[ \text{AgrS} \]

\[ \text{AgrO} \]

\[ \text{AgrS} \]

\[ \text{V} \]

\[ \text{AgrO} \]

\[ \text{T} \]

\[ \text{AgrS} \]

a. AgrS-to-C movement

b. heavy NP shift (repeated here)

Notice that, as a result of S-Structure specifier-head coindexing, the subject trace of heavy NP shift and the trace of AgrS bear index i, and so does AgrS itself (occupying C). Assuming with L&S (1992) that subject position is indelibly γ-marked at S-Structure, the subject trace is assigned a γ-feature in (89). As usual, the head of AgrS P (= the trace of AgrS) does not properly govern its specifier. But AgrS in C does antecedent-govern the subject; that is, the subject trace is indeed subjacent to the binding head AgrS in C. Therefore, the subject trace is assigned [+γ] and the ECP is satisfied. This, of course, is the wrong result. Suppose then that we supplement L&S’s definition of antecedent government (repeated in (90a–c)) with Rizzi’s (1990) condition on proper

39 There is a potential question regarding which category moves from the head of AgrS P to C in I-to-C movement. This question arises because it is not altogether clear which, if any, of the functional heads check and delete in the syntax. In fact, this potential unclarity is irrelevant to the proposed analysis. There are only four logical possibilities: only (a) a V, (b) a T, (c) an AgrO, or (d) an AgrS could be the category that moves to C. Under any of these possibilities the analysis will succeed. It will explain why syntactic “T”-to-C movement (i.e., syntactic V-, T-, AgrO-, or AgrS-movement to C) does not yield proper government of the subject.

In the text, I assume that the entire AgrS complex moves to C. I believe this is what Chomsky’s (1993) framework predicts; that is, since AgrS, AgrO, and T need not check V-features in the syntax (i.e., English is not a syntactic V-raising language), economy of representation predicts that they cannot. Thus, these categories do not delete in the syntax, since they have checking to do at LF.
head government (90d), a condition Rizzi motivated precisely to prevent I-to-C movement from causing a subject to be properly head-governed in English.\footnote{There is a redundancy between clauses (90b) and (90d). Each prevents a head X from antecedent-governing the specifier position \textit{YP} in (i).}

\begin{quote}
(90) \textit{Antecedent government}
\begin{enumerate}
\item A antecedent-governs B iff
\begin{enumerate}
\item A = X, and
\item A binds B, and
\item B is subjacent to A, and
\item A and B are both dominated by the single-bar projection of A.
\end{enumerate}
\end{enumerate}
\end{quote}

Given (90d), \textit{Agr}_{S} in C is now prevented from antecedent-governing the subject trace. Even though the subject trace is indeed subjacent to the binding head \textit{Agr}_{S}, (90d) prevents antecedent government since it is not the case that both \textit{Agr}_{S} in C and the subject trace are dominated by \textit{Agr}_{S}'. The subject trace is therefore indelibly marked \footnote{\textit{[g]} in the S-Structure representation (89) and the ECP is therefore violated at LF. Thus, adopting (90d) gives exactly the result Rizzi intended: syntactic I-raising does not render subject position properly governed. In fact, this analysis, including (89) with \textit{Agr}_{S} in C, is for all intents and purposes identical to the I-to-C movement analysis Rizzi assumes.}\footnote{This brings us to the final question: Why is it that \textquoteleft I\textquoteleft'-raising in LF \textit{does} yield antecedent government of the subject, as for example in the derivation of (7a)?}

This brings us to the final question: Why is it that \textquoteleft I\textquoteleft'-raising in LF does yield antecedent government of the subject, as for example in the derivation of (7a)?

(7) a. Who wonders what who bought?

If we assume the VP-Internal Subject Hypothesis, the embedded CP of (7a) appears as follows, prior to movement:

\begin{equation}
(91) \ldots [\textit{CP} C [\textit{Agr}_{S} P \textit{Agr}_{S}] [\textit{TP} T [\textit{Agr}_{S} P \textit{Agr}_{O} [\textit{VP} \textit{who bought what}]]]]
\end{equation}

In the mapping to S-Structure, \textit{who} raises to [Spec, \textit{Agr}_{S}], in which position obligatory syntactic checking of nominative Case by T adjoined to \textit{Agr}_{S} can take place. Moreover, \textit{what} moves to [Spec, CP] so that obligatory syntactic checking of [+wh] by C can occur. Thus, the embedded CP of (7a) appears as (92) at S-Structure.

\begin{equation}
(92) \ldots [\textit{CP} \textit{what} C [\textit{Agr}_{P} P \textit{who} C [\textit{Agr}_{S} P \textit{T}_{k} + \textit{Agr}_{S}] [\textit{TP} \textit{t}_{k} [\textit{Agr}_{P} P \textit{Agr}_{O} [\textit{VP} \textit{bought t}_{j}]]]]]
\end{equation}

(Notice that, by specifier-head coindexing, the trace of \textit{who} (\textit{t}_{j} in [Spec, VP]) and \textit{bought}, are coindexed, and therefore \textit{who} and \textit{bought} are coindexed as well.)
In the LF component, *who* must undergo long-distance movement, adjoining to the *wh-* phrase in the matrix [Spec, CP] and leaving a trace in the embedded [Spec, AgrS]. LF V-raising (checking) also applies. Recall that, given LF Movement Visibility, functional categories cannot be moved in the LF component. This means that, in this case, each LF movement must be movement of a verb. LF movement therefore proceeds in the following order:

(93) a. V adjoins to AgrO (the resulting AgrO category is now invisible to LF Move α).
   b. Object shift (of the direct object trace of *what*, tj) takes place to [Spec, AgrO].
   c. AgrO checks both Case features on tj and V-features on *bought*.
   d. AgrO automatically deletes, leaving a (movement-visible) V in the head position of AgrO P.
   e. V adjoins to AgrS (‘‘hopping over’’ the trace of T, in violation of the HMC).

Thus, in conformity with LF Movement Visibility, each movement is movement of a V (leaving a V-trace). The V-movement in (93), plus long-distance LF movement of the subject *who*, yields the intermediate representation (94) of the embedded CP in the LF component (irrelevant details omitted).

(94) . . . [CP *what* C [AgrO P t [AgrS [TP t [AgrO P t [VP t . . . ]]]]]]

d. V-to-AgrO movement (LF)
a. V-to-AgrS movement (LF)
b. T-to-AgrS movement (syntax)
c. long-distance movement of *who* (LF)

Notice that LF movement of the verb *bought* from the head of AgrO P to AgrS-adjoined position (= b.) violates the HMC owing to the intervening trace of T. The ECP would appear to be violated because this LF movement (of a V) leaves a V-trace, which is not deletable. Thus, just like the LF movement of a V over negation, this LF movement of a V over T-trace ‘‘should’’ violate the ECP. But in fact the ECP, a constraint on LF representations, will be satisfied. This is so because T (and T-trace) is, as I have assumed, a functional checking category that undergoes deletion in the LF component; hence, the intervening T-trace, the presence of which induced the HMC violation, will be absent in the LF representation (see below). The ECP will then be satisfied. (Thus, there are at least two types of movement violating the Head Movement Constraint that can nevertheless satisfy the ECP: (a) the movement leaves a deletable trace (Chomsky 1991) and
(b) the intervening head (more generally, the barrier to movement) deletes. I will return to T-trace deletion, in particular the derived constituent structure, momentarily.)

Now, suppose that LF “‘I’-to-C (i.e., AgrS-to-C) movement applies to (94). Once again, under LF Movement Visibility, the functional category AgrS cannot be moved. Thus, it must check and delete. However, this deletion leaves both the V\(^0\) \textit{bought} and T unattached, and it is not clear to me what the resulting structure is. I will simply assume (although I believe this is not crucial) that in (94) both AgrS and T check and delete, leaving only the V in the head of AgrS\(P\). The V, a category visible for LF movement, is then moved to C.\(^{41}\)

\[(95) \quad \ldots [\text{CP what } C_{AgrS} t t \ldots] \]

\subsubsection*{41 a. long-distance movement of who (repeated here)}
\subsubsection*{41 b. V-to-C movement}

If (95) were in fact the LF representation (of the embedded CP of (7a)), the ECP would be violated: the subject trace is not antecedent-governed. Although clauses (90a–c) of the definition of antecedent government are satisfied (i.e., the subject trace is subjacent to the binding head \textit{bought}), clause (90d) is not satisfied since the subject trace is not subjacent to \textit{bought within V\textsuperscript{′}}. But (95) is not in fact the relevant LF representation. Given that C\(^0\) is a purely functional category in this structure, FL/economy of representation dictates that it too must delete before a legitimate LF representation is obtained (I will return to the question of whether C\(^0\)-deletion (like Agr- and T-deletion) is a result of C\(^0\)-checking). The deletion of C\(^0\) yields an intermediate representation in which [\textit{v bought}] is immediately dominated by C\textsuperscript{′}.

\[(96) \quad \ldots [\text{CP what } C_{\text{V, bought}} [Agr_{S, P} t t, t, \text{TP } \ldots]] \]

This is still not an LF representation. As a result of C\(^0\)-deletion, X-bar theory is now violated since CP does not have a C head, but a V head. Similarly, the functional X\textsuperscript{max} categories AgrO\(P\) and AgrS\(P\) are also headed by a verbal category (namely, a V-trace of \textit{bought}), in violation of X-bar theory. (Recall, in LF, it was the movement-visible verb that moved from the head of AgrO\(P\) (to AgrS-to-joined position) and from the head of AgrS\(P\) (to C), leaving V-traces in the head of both AgrO\(P\) and AgrS\(P\).) In order to produce an LF representation that is consistent with X-bar theory, the V-headed AgrO\(P\), the V-headed AgrS\(P\), and, most crucially, the V-headed CP

\(^{41}\) If T had not deleted prior to movement to C, it would delete after movement to C, leaving the V \textit{bought} in C, exactly as in (95). This is the crucial property, yielding proper government of the subject trace in LF, as we will see momentarily.
must each become VP projections. This yields the (still intermediate) representation (97) of the embedded CP of (7a) in LF, in which all projections (except TP) have “become” VP projections.

(97) . . . VP (was CP)

Thus (as a result of functional head deletion (rendering the projection V-headed) and requirements imposed by X-bar theory), all projections (except TP) are transformed into VP projections. But what about TP? Recall that LF V-movement over T never moved *bought* through (i.e., never adjoined it to) T₀, so the head position of TP, unlike the head of the two Agr projections, is not verbal (i.e., not a V or V-trace). In fact, I assume the head of TP is, at this point in the derivation, ε, that is, an empty head devoid of features. I make this assumption because T (adjoined to Agrₛ) has checked and deleted in the LF component. Under the Proper Binding Condition, if T deletes,
its trace must too; otherwise, it would be unbound. (Alternately, the $T^0$ chain is the object that undergoes deletion.) I assume this deletion leaves $e$, a category devoid of features (Chomsky 1991) in the head of TP. But there is now (arguably) an unwanted ECP violation in (97) since the V-trace in the head of what was $AgrS_P$ is not close enough to the V-trace in the head of what was $AgrO_P$, because of the intervening $e$ in the head of TP (assuming $e$ ‘counts’ as an intervening head). Over and above this possible ECP violation, the representation is not consistent with X-bar theory since $T'$ fails to immediately dominate a T head. Following Chomsky (1991), I assume that T-deletion, leaving $e$ in the head of TP, entails that the TP projection must be transformed as follows in order to satisfy X-bar theory:

\[(98) \[ eP \[ e' e \]]\]

But now, if (as in Chomsky 1991) $eP$ has no features, then I assume (contra Chomsky) that FI/economy of representation prohibits it from appearing in an LF representation since it presumably has no interpretation. I suggest that a legitimate LF representation is derived by substituting the V-trace occupying the head position of (what was) $AgrO_P$ into the empty $e$ (leaving a V-trace in the departure site).\footnote{See Chomsky 1991 and Kitahara 1994, 1997, for proposals regarding the movement of $X^{\text{max}}$-traces.} In accordance with X-bar theory, the $eP$, now bearing a verbal head, becomes VP, just like the other functional projections did.\footnote{Other analyses of the problem raised by V-movement over the intervening $T^0/e^0$ are imaginable. First, as suggested in Thráinnsson 1996, it might be that English, in contrast to (e.g.) Icelandic, has an unsplit I, as in analyses preceding Pollock 1989. If this is correct, the problem simply disappears. Second, even if English has both $AgrS$ and $T$, it is not altogether clear that LF V-movement from the head of $AgrO_P$ to $AgrS$-adjoined position, over the head of TP/eP, is in fact a violation. Given Ferguson and Groar’s (1994) formulation of the ‘Shortest Movement’ Condition, such V-movement over the tail of the $T^0/e^0$ chain would be allowed, since no closer V-checking position was ‘skipped over.’ Third, if $e^0$, being semantically null, is (perhaps naturally) invisible to LF operations, its presence during the movement would be irrelevant. Finally, an LI reviewer suggests that ‘[i]t may be possible to assume that the empty projection is pruned away, excised.’ This, too, is a possibility, but notice that the entire empty ‘projection’ (i.e., the syntactic categories labeled $T$’ or TP in (97)) of course cannot be deleted (in any standard sense of deletion). By contrast, deleting just the nodes labeled $TP$, $T$, and $e$ in (97) would constitute deletion (Affect a) of entities that are not syntactic categories (a type of operation whose existence I leave open for further research).} This yields the (fully interpretable, convergent, and X-bar-theory consistent) LF representation (99) of the embedded CP of (7a). Notice that, given the transformation of TP/eP to VP, the V-chain now unequivocally satisfies the ECP. Moreover, the central result we have been seeking is now obtained: the subject trace of long-distance LF movement of who (occupying the specifier position of the former $AgrS_P$) is antecedent-governed in this LF representation. That is, not only is the subject trace subjacent to the binding verb bought (90a–c), but also, in conformity with (90d), the subject trace is subjacent to the binding verb bought, within the single-bar projection (the $V'$ projection) of bought.\footnote{Recall that bought and who are coindexed via specifier-head coindexing at S-Structure. LF specifier-head coindexing coindexes the trace of who in the specifier of what was $AgrS_P$ with the trace of bought in the head of what was $AgrO_P$. Thus, the traces of who and bought are indeed coindexed, so that binding does in fact obtain. Further, binding within the $V'$ projection of bought obtains because, as a result of C0-deletion, CP is transformed into VP. The (complex) index-sensitive head government condition (90) will be reformulated as a more natural condition in section 6.5.}

We now have an answer to the question, Why does LF ‘‘I’’-raising to C0 properly govern the subject position? In a nutshell, LF ‘‘I’’-raising is in fact V-raising to C, that is, covert V2. This yields a (‘‘category-neutral’’) LF representation like (99) in which all projections are (non-
a. V-to-C movement (LF)
b. long-distance movement of the subject who (LF)
c. V-trace movement

functional) VP projections, as (I have argued) is independently dictated by economy of representation at LF (prohibiting purely functional checking categories) and X-bar theory. Therefore, in LF, unlike in the syntax, clause (90d) of antecedent government is satisfied: the fronted verb does antecedent-govern the subject within the V′ projection of the verb. Consequently, long-distance LF subject movement ( unlike long-distance syntactic subject movement) can satisfy the ECP. Under the analysis proposed here, English has V2 (albeit in LF), a process known to render subject position properly governed (see Torrego 1984 concerning Spanish and Rizzi 1990 concern-
ing languages like German, for a proposal regarding proper government of subject position at S-Structure by a syntactically fronted verb). By contrast, syntactic ‘‘I’’-raising in English does not render subject position properly governed. As in Rizzi 1990, I-to-C movement (here, AgrS-to-C movement) does not result in the subject’s being antecedent-governed. Since C and its projections remain present at S-Structure, the fronted AgrS in C does not bind the subject within AgrS. Rather, the only single-bar projection dominating both the fronted AgrS and the subject is C’, and antecedent government therefore does not obtain at S-Structure.

Thus, like Chomsky’s analysis of V-raising over negation, the analysis proposed here rests on the (independently motivated) presence of certain functional categories in the syntax and on their (principled) absence at LF. The significant aspects of the differences produced by syntactic ‘‘I’’-to-C movement and LF ‘‘I’’-to-C movement are illustrated in (100).

(100) a. S-Structure

```
CP
   /
  /   /
C'  /   /
   C  /   /
     /   /
     CAGRSP
     /   /
   AGRS  /
   /   /
   /   /
AGRO  AGRS
    /     /
   /     /
   /     /
   V     T
  
has
```

b. LF

```
VP (was CP)
   /
  /   /
V'  /   /
   V    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /    /
   /7
```
Under Rizzi’s (1990) analysis, in the S-Structure representation AgrS does not antecedent-govern the subject since it is not the case that they are both dominated by AgrS’. By contrast, in LF, a fronted V does indeed antecedent-govern the subject. This is a direct result of (a) functional-head deletion (demanded by Fl/economy of representation), which yields a CP with a V head, and (b) X-bar theory, which requires the V-headed CP to be transformed into a VP. In fact, conforming to X-bar theory (but see section 7) requires each V-headed functional projection (produced by functional-head checking and deletion) to be transformed into a VP. This yields, at LF, ‘‘category-neutral’’ VP-recursion structures such as (99) (similar in certain respects to those Larson (1988) proposes for double object constructions).46

In the following section I examine the arguably unnatural definition (90) of antecedent government, upon which I have been relying.

6.4 Toward the Elimination of Index-Sensitive Antecedent (Head) Government Conditions

Consider again the index-sensitive (hybrid) definition of antecedent (head) government used thus far.

(90) Antecedent government
A antecedent-governs B iff
a. A = X0, and
b. A binds B, and
c. B is subjacent to A, and
d. A and B are both dominated by the single-bar projection of A.

46 Several issues arise regarding the analysis presented in this section. The first concerns LF V2 and parameterization of that-t effects. Law (1991) provides a (pre–Minimalist Program) analysis of West Flemish (WF) complementizer agreement postulating [V[+I]] movement to C in LF. (I thank Daiko Takahashi for bringing this article to my attention.) Within Law’s analysis, the suppression of syntactic that-t effects, as illustrated by the grammaticality of (i), is ascribed to the head government of the subject trace produced by LF movement of [V[+I]] to C. ((i) is adapted from Law 1991: (14), (16.).)

(i) a. den vent da Pol peinst da Valère gezien heet
   ‘the man that Pol thinks that saw Valère’
   b. S-Structure
   [CP O, [da [µ Pol peinst [CP t’; [da [µ t Valère gezien heet +I]]]]]]
   c. LF da (that)-replacement
   [CP O, [peinst [Pol [CP t’; [heet +I]] t Valère gezien t]]]

Thus, contra Rizzi 1990:39, where it is proposed that the proper head government condition must be met at S-Structure, and contra L&S 1984, 1992, in which argument (subject) traces are indelibly s-marked at S-Structure, and contra the ‘‘*-marking’’ analysis proposed in Chomsky and Lasnik 1993:546, Law postulates that the subject trace in (ib), which violates the head government requirement at S-Structure (and at the moment of its creation), can be ‘‘saved’’ by LF head movement of [V[+I]] to C, which renders the subject trace head-governed in LF.

Given his analysis, Law argues that English must lack [V[+I]] movement to the overt finite complementizer that in LF; if English had such movement, that-t configurations as in (ii) would be incorrectly predicted to be grammatical, being salvageable by LF head movement, just like the WF case (i).

(ii) *[CP who do [µ you think [CP t, [that [t I, left]]]]]

However (leaving aside his discussion of negative inversion), Law’s proposal that English (otherwise) lacks LF
As noted above, (90b) and (90d) each serve to block proper government of a specifier by its head in the following configuration:

(101) [XP Spec [X′ X YP]]

Proper government is blocked under (90b) because X does not c-command (see (26)) Spec, and under (90d) because Spec is not dominated by X′. The question I would like to address here is this: Can the index-sensitive head government condition (90) be reduced to a more natural condition (as dictated by the Minimalist Program (Chomsky 1993))? Notice that Spec in (101) is a member of the checking domain of X⁰, but is not a member of the internal domain of X⁰. Given this observation, I propose the following sufficient condition for satisfaction of the ECP:

(102) *Proper government*

t is properly governed (i.e., satisfies the ECP) if t is a member of the internal domain of an X⁰ chain.

As desired, (102) does not predict that Spec in (101) is properly governed. Moreover, this definition allows us to unify three seemingly distinct cases of proper government:

(103) a. Direct object (V-sister) complements

b. The (non-V-sister) “inner subject” in VP shell configurations—for example, the book in

\[ \ldots [\text{VP Sue} [V^\prime \text{ put} [\text{VP the book} [V^\prime t \text{ on the table}]]]] \]

c. [Spec, Agr_s P] (= “subject position”) after V₂ applies (e.g., t₁ in the LF representa-

(100b))

Given (102), cases (103b) and (103c) are equated under the proposed analysis; that is, the LF representation (100b) is identical in the relevant respects to one derived by V-raising (in a VP movement to finite complementizers (that and 0) leaves not only the central example (7a) unexplained, but also the following type of case, which is certainly far better than (ii):

(iii) who, thinks [[CP that [IP who I left]]]

(I find (iii) grammatical, but see L&S 1992:116 and Aoun, Horstein, and Sportiche 1981:fn. 14 (and references cited there) for discussion of the grammaticality status and analysis of such examples.)

So that the long-distance trace of who₂ created in LF can be properly head-governed, LF V-to-C movement presumably applies in (iii). But then the following question emerges: If (a) Law’s analysis is correct, and (b) WF has LF head movement that can “salvage” a subject trace created in the syntax, and (c) English indeed has LF head movement properly governing a subject trace created in LF ((7a) and section 5), then why is it that English syntactic that-t violations like (ii) cannot be salvaged by LF head movement, as (Law’s analysis proposes) they are in the WF case (i)?

One way (out of perhaps many) of resolving this possible problem (which arises only if Law’s analysis and the one presented here are both on the right track) is to assume a subtle and by no means unprecedented type of distinction between the two grammars: in English the head government requirement is in effect a constraint on movement (Chomsky and Lasnik 1993) or perhaps an S-Structure requirement (L&S 1984:92, Rizzi 1990). Thus, for (ii), there is no salvation. In WF the constraint is representational, applying only in LF. It is interesting in this context to note that “the ECP” as
shell). That is, both V-substitution and V-adjunction to a checking \( X^0 \) with subsequent deletion of the checking \( X^0 \) (103c) are processes whereby a specifier becomes a member of the internal domain of an \( X^0 \) chain and as such is nondistinct from a direct object complement (103a)—the position-type canonically satisfying the ECP.

But even though (102) equates (103b) and (103c) while reducing each to (103a), it confronts a potentially serious problem. (102) would seem to be overly permissive in that it apparently incorrectly predicts proper government in (89)/(100a): the \( t_i \) subject (= [Spec, Agr\(_S\)P], trace of heavy NP shift) is apparently a member of the internal domain of the Agr\(_S\) chain. This incorrect result would obtain even if Agr\(_S\) were instead adjoined to C\(^0\). Depending on whether such Agr\(_S\)-movement is substitution or adjunction to C\(^0\), there is a possible way of preventing proper government from obtaining in (89)/(100a). Recall that the internal domain of the Agr\(_S\) chain is the minimal subset of the domain of the Agr\(_S\) chain reflexively dominated by the complement of \( \alpha_1 \) (where \( \alpha_1 \) is the head of the Agr\(_S\) chain, itself). The question, then, is this: In (100a) (= (89)), or in the corresponding structure in which Agr\(_S\) is instead adjoined to C\(^0\), what is the complement of Agr\(_S\)?

I suggest the two following alternative definitions of complement:

(104) \( \beta \) is the complement of \( \alpha^0 \) iff the set of categories dominating \( \beta \) is equivalent to the set of categories dominating \( \alpha^0 \).

(105) \( \beta \) is the complement of \( \alpha^0 \) iff the set of categories containing \( \beta \) is equivalent to the set of categories containing \( \alpha^0 \).

If Agr\(_S\) is substituted into C\(^0\) (100a)/(89), then under either (104) or (105), Agr\(_S\) has no complement whatsoever. (Crucially, Agr\(_S\)P is not the complement of Agr\(_S\), since C dominates Agr\(_S\) but not Agr\(_S\)P.) Consequently, the internal domain of the Agr\(_S\) chain is the null set. As desired, [Spec, Agr\(_S\)P] (= the trace of heavy NP shift in (89)) is therefore not a member of the internal domain of the Agr\(_S\) chain and is therefore not properly governed under (102). (If Agr\(_S\) is adjoined to C, and if we assume, contra May (1985), that adjunction to C duplicates C (as in L&S 1992), then, once again, either (104) or (105) suffices to block proper government of [Spec, Agr\(_S\)P] under (102).) If, on the other hand, Agr\(_S\) is adjoined to C and May’s segmental theory of adjunction obtains (as in Chomsky 1993), then only (105) (i.e., not (104)) suffices to block proper government under (102). Given (105), Agr\(_S\)P is not a complement of Agr\(_S\) (C is a container of Agr\(_S\) but is not a container of Agr\(_S\)P). In fact, the internal domain of the Agr\(_S\) chain is (again) the null set.

formulated in Chomsky and Lasnik 1993 has both a derivational property (*-marking) and a representational aspect (the * filter applying to LF representations).

The second issue that arises from the discussion in this section involves syntactic V2 and parametric variation regarding the “timing” of C\(^0\)-deletion. Within the (checking-based) analysis proposed here, parametric variation concerns the level at which checking by a functional category occurs and critically involves Procrastinate (a principle that, I assume here, entails that operations apply as late as possible to yield convergence). Here I simply mention this issue; for a more complete discussion, see Epstein 1993.
Therefore, as desired, [Spec, AgrS] is not a member of the internal domain of the AgrS chain and is not properly governed under (102). 47

To summarize this section: I have sought to eliminate the index-sensitive antecedent (head) government condition (90) and have suggested that membership in the internal domain of an X0 chain is sufficient for (and in fact represents the canonical case of) satisfaction of the ECP. This analysis unifies the three cases in (103) and incorporates only independently motivated primitives. Moreover, it makes no appeal to indices, “θ-marking by an X0,” or “head government.” 48

7 The Bare Phrase Structure Theory

The purpose of this section is to suggest that the analysis proposed here (which in one respect depends upon a certain aspect of (stipulated) X-bar theory) is entirely consistent with (if not entailed by) the theory of bare phrase structure (BPS) outlined in Chomsky 1994.

I have argued that when a V adjoins to a functional checking head (which heads a functional projection), the checking head checks and deletes, leaving the V, immediately dominated by the single-bar projection of the (deleted) checking head (as it in fact was before the checking head deleted). For example, I have postulated that for (7a), V-raising to AgrS and then to C, with checking and deletion of AgrS and C, yields the intermediate structure (106) in the LF component.

\[
\begin{align*}
(106) \quad \text{who} & \quad \text{wonders} \\
\quad \text{C}(P) \\
\quad \text{N}(P) & \quad \text{C}^{(r)} \\
\quad \text{what} & \quad \text{V} \\
\quad \text{bought} & \quad \text{N}(P) \\
& \quad \text{AgrS}(P) \\
& \quad \text{AgrS}^{(r)} \\
\quad \text{who} & \quad \text{V}^{\text{copy}} \\
& \quad \text{T}(P) \\
& \quad t
\end{align*}
\]

47 Notice that when the internal domain is the null set, the checking domain will (consequently) be equivalent to the minimal domain itself. The empirical consequences of this require further investigation.

Notice also that, within the account I am proposing, the bifurcation of the minimal domain into internal and checking domains differs from the bifurcation assumed in Chomsky 1993. However, the minimal domain of the chain is the same in my account and Chomsky’s. Consequently, my analysis is consistent with Chomsky’s (1993:12) “equidistance” analysis of apparent Relativized Minimality violations induced by object shift, an analysis resting only on the notion “minimal domain” (and hence not sensitive to the particular bifurcation of the minimal domain into checking and internal domains).

48 For an alternative analysis of the LF suppression of subject ECP/wh-island effects, see Martin 1996. Martin also postulates a crucial role for LF head movement. In his analysis, [+wh] C is an LF affix and undergoes LF adjunction to the higher V. In (7a) the embedded [+wh] C adjoins to the V wonder in LF. Such movement renders the embedded [+wh] [Spec, CP] (= what) and the matrix VP-adjoined position equidistant (in the sense of Chomsky (1993)) from
(In (106), \(P\) and \(\prime\) appearing in category labels are parenthesized since, within BPS, such notations are eliminated given that phrase structure status (minimal projection, maximal projection, minimal and maximal projection, neither minimal nor maximal projection) is relationally determined (see Muysken 1982, Freidin 1992).)

In order to exclude such representations (at least at the interface levels, but perhaps altogether), I have claimed that such (postchecking and -deletion) structures, displaying a V-headed C projection and a V-headed Agr\(_S\) projection, are "excluded by X-bar theory," which Chomsky (1993) construes as a constraint on the output of the application of the generalized transformation GT. BPS, I believe, excludes these structures as well. Within BPS, the category labels \(C(P)\) and \(C(\prime)\), as well as \(Agr_S(P)\) and \(Agr_S(\prime)\), lack the parenthetical material (as noted) and are in fact identical to the head; in other words, "the head . . . the label . . ." (Chomsky 1994:16). Consequently, if the head (C or Agr\(_S\)) checks and deletes, these "higher" projection labels must (automatically) delete since they are identical to the head, from which they are projected. Thus, within BPS too, structures like (106) cannot exist; in particular, if the C head checks and deletes, the identical \(C(\prime)\) and \(C(P)\) labels delete too. Similarly, checking and deletion of the Agr\(_S\) head entails deletion of the identical \(Agr_S(\prime)\) and \(Agr_S(P)\) projection labels. But then, if these four projection labels are absent, what are the projection labels appearing in (106) as generated within BPS? Let us begin with \(Agr_S(\prime)\). This label must be either the head V-copy trace or the head of (its sister) TP, but this too is identically V-copy trace. Thus, the label \(Agr_S(\prime)\) must be V; there is no other possibility. Exactly the same pertains to the label \(C(\prime)\); this label, too, must be identical to the head of one of its two daughters, but again "the two are identically" V. What about the maximal projection labels \(Agr_S(P)\) and \(C(P)\)? Each of these labels must be V as well. This follows since (as Chomsky (1994) shows) in cases of substitution (in (106), specifier substitution), it is deducible that a moved category (in (106) who and what) can never project; hence, the head of the N(P) sister in (106)—namely, the category V—must project. This \(Agr_S(P)\) and \(C(P)\) are (like \(Agr_S(\prime)\) and \(C(\prime)\) labeled V in BPS. Thus, it appears that BPS (like stipulated X-bar theory) precludes structures like (106) and moreover forces their "transformation" into VP-recursion structures, just as I have proposed.

There are, however, two respects in which my analysis might appear to be inconsistent with BPS. First, the VP-recursion structures produced in LF appear to be derived by a type of movement that Chomsky (1994) refers to as "self-attachment." Although (as far as I can tell) Chomsky’s (1994:19, 20) discussion of self-attachment provides no direct empirical support for prohibiting it, the empirical matter need not concern us here since my analysis does not in fact involve (Greed-violating) self-attaching movement; that is, the V was not in fact adjoined to a VP (labeled V) that it headed prior to movement, but rather was adjoined to Agr/C (i.e., a checking head), thereby

\footnote{I am indebted to Noam Chomsky for discussion of this and other points raised in this section. Any errors are mine.}
satisfying Greed and thereby not representing a case of self-attaching movement (although the derived VP-recursion structures look as if they were derived by self-attaching movement). As a result, Chomsky’s (1994) conceptual arguments against self-attachment (namely, the unambiguous deduction of target (not mover) projection for cases of movement) are unaffected by my analysis, which invokes only (familiar) V-adjunction to a checking head.

The final respect in which my analysis might appear to be inconsistent with BPS is that I assume that the checking head literally deletes (as in Chomsky 1991, 1993), whereas Chomsky (1994:13, fns. 23, 24) suggests the possibility that checked functional heads might not delete but might instead be rendered “invisible” (noting (fn. 24) that “A question arises about implementation; deletion raising technical issues…”). There seem to me to be at least two potential problems with this alternative to literal deletion. First, conceptually, there is always a potential problem with explaining why and how it is that symbols are present in a representation yet the computational system and interface treat them exactly as if they were absent, a fact that could be readily (but not necessarily uniquely) explained if they were indeed absent, as is assumed here and in Chomsky 1991, 1993. Second, there is an empirical problem. If a checking head (e.g., AgrS) is merely rendered “invisible” (i.e., not deleted), then given that the projection labels AgrS(↑) and AgrS(P) in, say, (106) are identical to the AgrS head in BPS, then they too, being identical, would remain but would be rendered (like the head itself) “invisible.” But this presumably is the wrong result, since (e.g.) the category labeled AgrS(P) in (106) is (I assume) in fact a maximal projection and is therefore visible both for computation and at the interface, as is captured under the literal deletion analysis within which this category is a (fully visible) maximal projection bearing the category label V.50

Beyond the investigation presented here, a more comprehensive determination of the precise formal properties and empirical content of checking and checking-induced deletion within a BPS-based Minimalist Program analysis awaits further research.

8 Summary

The central aspects of the analyses presented here can be summarized as follows:

- The Operator Disjointness Condition (L&S 1992) can be replaced by the arguably more natural Scope-Marking Condition, the empirical adequacy of which relies on the theory of linked chains.
- A new theory of adjunction, the mixed theory of adjunction, perhaps obtains.
- Consistent with a leading idea concerning parametric variation, English generates covert verb-second.

50 For a critical examination of this analysis, see Groat 1994.
• (Certain) LF representations/subtrees are “category-neutral” VP-recursion structures, lacking “functional” checking categories and their projections.

• Index-sensitive head government ECP requirements may be reducible to simpler, more natural conditions, given certain independently motivated principles of checking theory and the derived constituent structures I have proposed for checking-induced deletions.

• The analysis of checking-induced deletion proposed here is argued to be consistent with, if not “forced by,” the bare phrase structure analysis proposed in Chomsky 1994.

References


Epstein, Samuel David. 1990b. X\(^0\) movement, X\(^{\text{max}}\) movement and a unified definition of ‘‘antecedent government.’’ Ms., Harvard University, Cambridge, Mass.


---

*The University of Michigan*

*Linguistics Program*

*1076 Frieze Building*

*105 South State Street*

*Ann Arbor, Michigan 48109–1285*

*sepstein@umich.edu*