Chomsky (1993) suggests a return to the “bottom-up” structure-building algorithm of a generalized transformation (GT), applications of which are restricted by the Extension Requirement (ER) discussed below. Chomsky notes that two central processes—head movement via adjunction (or adjunction generally) and LF movement—appear to violate the ER. Recent work has attempted to derive the exceptional nature of these two processes from independent principles of grammar. We contend that the seemingly exceptional nature of head movement results from an unmotivated and unrecognized assumption that effectively prohibits what we will call interarboreal operations. If this assumption is abandoned, head movement can be seen to obey the ER. In this squib we therefore offer a formal solution to a theory-internal problem, in the hopes of suggesting a direction for further research.

An example of an interarboreal operation is schematized in (1). In this derivation (which we discuss in more detail in section 2), the verb adjoins to $P^0$ prior to concatenation of $P^0$ and the VP by the operation Merge (a subcase of GT), as in (1b). It is the resulting complex $P'$ that is subsequently merged with the VP, projecting an IP node. As a result, the phrase marker generated by this operation—(1c)—is isomorphic to that derived by the more traditional ER-violating version of head movement. Within the current framework, there is no principle that excludes operations like (1). Thus, we suggest that either such interarboreal operations should be admitted and exploited, or a principled reason to exclude them should be proposed. Here, we will take the former tack.

In section 1 we present Chomsky’s definition of the ER and examine the apparently exceptional nature of head movement. In section 2 we clarify the technical details of the ER-consistent interarboreal operation that derives head adjunction configurations. In section 3

This squib reports on ideas developed independently in Bobaljik 1995a and Brown 1995. For discussion of these ideas and comments on the squib, we would like to thank Noam Chomsky, Samuel D. Epstein, Danny Fox, Howard Lasnik, Suei Wurmbrand, and two LI reviewers. Errors are ours alone.
In other words, the ER requires that all operations add a sister to a root node. A new sister node cannot be created for any node that is already in a sisterhood relationship. This blocks downward and sideways movement and other noncyclic operations. However, Chomsky (by stipulation) restricts (3) to overt substitution operations only. The reason for this is clear: adjunction (especially head movement) and (most) LF movement violate (3). To see this, consider (4), a typical case of V-to-I raising.

(1)   a. VP I DP  
       b.  
       c. VP I V  V DP

we discuss how overgeneration is prevented by integrating a slightly modified Chain Condition (Rizzi 1986, 1990), similar to Chomsky's (1993b:253) C-Command Condition on chains. In section 4 we present our conclusions.

1 The Extension Requirement and the Problem of Head Movement

Chomsky (1993:22) suggests that the basic syntactic operation is a generalized transformation (GT), which may be broken down into discrete steps as in (2) (after Chomsky 1993:22, Kitahara 1994, 1995).

(2) a. Target a phrase marker \( \kappa \).
b. Add an empty \( \emptyset \) external to (i.e., as a sister to) \( \kappa \).
c. Substitute a phrase marker \( \beta \) for \( \emptyset \), forming the new phrase marker \( \gamma = \text{a projection of } \kappa \).

In subsequent work (Chomsky 1995a,b) the operation in (2) is called Merge, though the steps in (2) remain essentially unchanged. Move (or Move a) is a similar operation (Chomsky 1993:22), with the additional assumptions that \( \beta \) in (2c) is an element (nonroot phrase marker) within the target \( \kappa \) and that the operation involves a copying and/or chain formation procedure. Chomsky also proposes (1993:22) that GT is constrained by what has come to be called the Extension Requirement.

(3) "\( \emptyset \) must be external to the targeted phrase marker \( K \) [i.e., \( \kappa \)]. Thus, GT and Move a extend \( K \) [i.e., \( \kappa \)] to \( K' \) [i.e., \( \gamma \)], which includes \( K \) [i.e., \( \kappa \)] as a proper part."

1 The precise details of the copying procedure are different in Chomsky 1993 and subsequent work. In Chomsky 1993 it is assumed that movement "leaves behind a trace" (p. 22) identical to the moved element (p. 37). In the text we adopt the formulation of Move given in Chomsky 1995a:13.

(i) Move a

"[GIVES] the phrase marker \( \Sigma \) select \( \kappa \), a.s.t. the root of \( \kappa \) dominates or \( \kappa \)-commands \( \alpha \); form \( \alpha \) by merging \( \alpha \) and \( \kappa \); define the chain \( \Gamma = (\alpha_1, \alpha_2) \)."

For Chomsky's (1995a) purposes, the choice between this formulation and one in which the chain is formed as a part of movement (p. 399) is not important. For our purposes, it becomes important in section 3.
Consider again GT, (2). This can be viewed as the core of the
operations Merge and Move (e.g., Chomsky 1993:22, Epstein 1994).
In either case, it is assumed that the target (x in (2)) must be a root
node—a phrase marker that is not a subtree of any other phrase marker.
In the case of Merge (or binary GT; Chomsky 1993:22), the element
that is substituted (β in (2)) is, like x, a root node; it may be either
an independently constructed, arbitrarily complex phrase marker or a
form selected from the numenation. It is important to note that, within
this framework, multiple phrase markers often must be built up in
parallel, though this point is frequently omitted as an expository short-
hand. The operation Merge concatenates these phrase markers, applying
"often enough to leave... just a single [phrase marker]" (Chom-
sky 1995b:226, 243). To see that there must be multiple (complex)

situational position, "overt movement" may be characterized as the pronunciation
of the highest instance or copy, with "covert (i.e., LF movement)" pronunciation of the lower copy. Metaphorically speaking, all movement could be
overt, the apparent covert/LF distinction being an artifact of pronunciation.
Approaches along these lines are pursued in Bokajdyk (1995b), Brody 1995, Great
and O'Neil 1994, Poeppel, in preparation, and works referred to therein.
Chomsky (1993) and Kitaara (1993) consider two types of (overt) adjunction that apparently violate the ER. One of these is head movement,
which we consider in detail in the text. The other is an asymmetry between relative clauses and complements clauses concerning Condition C—"recon-
struction" effects. Thus, given a pair like (i) and (ii) from Freidin 1994, coo,
ference between Johns and he is possible if John is in a relative clause (i), but impossible if John is in a CP complement to the noun recommendation (ii).
(i) [Which recommendation that John, revised] did he recover submit $\lambda y$?
(ii) [Which recommendation that John, should be promoted] did he submit $\lambda y$?

As a reviewer points out, an appeal to inter arboreal movement along the
lines we suggest for head movement does not seem to be a plausible route to
an account of the contrast between (i) and (ii). However, as Chomsky (1993:
44) notes, there are a number of alternative accounts of (i) and (ii) that do not
appeal to counter cyclicity adjunction. Freidin (1994:383–384) suggests that
in (i), "the relative clause modifies the variable and therefore remains as a
part of the quantifier structure, as in (ii)"; counter cyclicity does not lead to
a Condition C violation. In the (abbreviated) LF representation of (ii) shown in
(iv), the complement clause is a part of the variable, not a modifier of it, and
complementation therefore violates Condition C.

(iii) (for which x such that John revised x $\lambda x = \text{a recommendation}$ did he submit $\lambda x$)
(iv) (for which x $\lambda x = \text{a recommendation}$ that John should be promoted did he submit $\lambda x$)

In other words, Freidin proposes that only the variable need reconstruc-
ted, and that a relative clause, as a modifier of the variable, need not. The asymmetry is captured without appeal to counter cyclicity adjunction, though other questions are raised. It is worth noting that Chomsky (1993:44) cites Freidin's analysis as evidence that binding conditions hold at LF. Other accounts of the contrast are possible, though such discussion is well beyond the scope of a brief.
Squibs and Discussion

Copy. This copy then serves as the input (6) to the substitution step of GT (2c). We illustrate this schematically in (6), where a DP is raised from complement of V to specifier of IP. First, Copy operates (6a); then GT/Move substitutes the copy (6b).

(6) a. 

\[ \text{IP} \]
\[ \text{DP} \]
\[ \text{I} \]
\[ \text{V} \]
\[ \text{arrives} \]
\[ \text{DP} \]
\[ \text{D} \]
\[ \text{N(P)} \]
\[ \text{monster} \]

Copy
\[ \text{the} \]
\[ \text{arrives} \]
\[ \text{DP} \]
\[ \text{D} \]
\[ \text{N(P)} \]
\[ \text{monster} \]

(6) b. 

\[ \text{IP} \]
\[ \text{DP} \]
\[ \text{I} \]
\[ \text{V} \]
\[ \text{arrives} \]
\[ \text{DP} \]
\[ \text{D} \]
\[ \text{N(P)} \]
\[ \text{monster} \]

At PF the lower instance of the DP [the monster] is deleted (Chomsky 1993:35).

As noted above, Chomsky makes another assumption about Move, namely, that the source of the copy must be contained in the target (see 1993:22 and 1995a:399 fn. 13; also see fn. 1 above). Obviously, this assumption precludes interaroreal operations like (1); however, we are aware of no independent motivation for it. Moreover, though the assumption is tacitly maintained in Chomsky 1995b, there it is apparently taken to follow from the requirement that every instance of a given element must either c-command or be c-commanded by all other instances of that element (see Chomsky 1995b:250, 253). In fact, as we will show in section 3, this is not correct. The C-Command Condition (Rizzi's (1986) Chain Condition) does not itself rule out interaroreal operations; they are excluded by stipulation only. We now consider a potential advantage to abandoning this stipulation.

2 Interaroreal Operations

Consider again the derivation schematized in (1). The phrase marker in (1a) is the result of the following operations: GT targets a root node \( \kappa \) (= V') and adds an empty position \( \theta \), projecting \( \gamma \) (= VP). In this instance, the element substituted for \( \theta \) is DP (independently constructed by GT if complex, or provided from the numeration by Select if trivial/lexical). The next step derives the phrase marker in (1b): GT targets a root node \( \kappa \) (= I') and adds an empty position \( \theta \), projecting \( \gamma \) (= complex I'). As in the first operation, the target is a trivial phrase marker—in this case I'—provided from the numeration by Select. We take this step to be adjunction; hence, \( \gamma \) is a segment of I' (i.e., it does not project to a higher bar level). In this step the element that is substituted for \( \theta \) is V'—a copy of an element contained in a more complex phrase marker. That is, this operation is Move (i.e., Copy + Merge); it obeys the ER and all aspects of the characterization of movement except for the unmotivated stipulation that Move must operate internally to a single phrase marker. Finally, yet another operation of GT targets a root node \( \kappa \) (the complex I' from (1b)) and substitutes the VP for \( \theta \), projecting \( \gamma \) (= IP). All three operations obey the ER. Importantly, the result of the final step (1c) is exactly the structure that is derived via the ER-violating movement operation in (4)—the standard characterization of head movement. The problem, then, is not to derive the exceptional nature of head movement, but to show that interaroreal operations can be adequately constrained.

3 Constraining Interaroreal Operations

Many questions arise at this point, and we will address a few of them here. First, we consider the status of XP-movement with respect to interaroreal operations. We also consider the status of requirements such as Shortest Movement—in particular, the Head Movement Constraint (Travis 1984). We argue that Rizzi's (1986) Chain Condition,

---

1This follows from the assumption that "[a]junction differs from substitution only in that it forms a two-segment category rather than a new category" (Chomsky 1995a:402; see also May 1985 and, for potential problems with this view, Epstein 1989).
instances of some term $\alpha$, the two instances are interpreted as a chain iff (7) is satisfied. Taking, for example, the interarbororeal operation in
(1), an operation of Merge creates the phrase marker in (1b). This operation satisfies (7) since the newly created phrase marker—the complex $P'$ in (1b)—contains only one $V$. A subsequent operation of Merge joins the two phrase markers (1a) and (1b). The newly created phrase marker (1c) contains two $V$; thus, (7) is satisfied—the operation of Merge is legitimate—only if these are in the appropriate relationship. As noted above, under the segmentation analysis of adjunction (May 1985) the required c-command configuration obtains if interarbororeal head movement is analyzed as adjunction. In addition, Baker (1988) provides evidence that heads c-command out of adjunction structures; that is, a verb adjoined to $I$ governs its trace. Thus, we take interarbororeal head movement to be adjunction. The Chain Condition, as part of the definition of Merge, is satisfied.

Turning now to XP-movement, consider a schematic example of an interarbororeal operation of XP-substitution. In (8) the complement DP of some head X moves to become the specifier of an independent (root) phrase marker, YP (i.e., the DP is copied and the copy merges with $Y$), satisfying the ER. $Y$ projects to YP, and YP in turn is merged with the root node containing the source element DP, again satisfying the ER. The result looks as if “sideways” movement had occurred. (8) clearly runs afoul of the Chain Condition. The final operation derives a phrase marker containing multiple instances of a single DP (i.e., the original as complement to X and a copy as the specifier of D, with “global” evaluation procedures, which may “look ahead” and consider subsequent steps of the derivation in determining economy. Formally, it is similar to Lasnik and Saito’s (1992:90) Generalized Proper Binding Condition—the principle that traces must be bound at every step of the derivation. As a reviewer notes, to the extent that the principle yields output, it can be restated as such an “everywhere” condition, the derivational versus representational distinction is blurred, and the case for a “strictly derivational” approach weakened.

As David Pesetsky (personal communication) points out, many hypothetical derivations involving interarbororeal XP-movement superficially resembling (1) will run afoul of some version of a feature-checking theory. A reviewer addresses (i) as an example that plausibly meets all feature-checking requirements, and for which we would thus need to appeal to the Chain Condition or something similar.

(i) *(That John believed me, to have appeared that Mary is sincere*) convinced them that it was told t.

The pronoun me is merged into the passive VP with told, where it receives a 0-role but no Case. Next, a copy of me is merged into the 0-less subject position of the raising predicate to have appeared, where it checks Case features via exceptional Case marking. Finally, Merge introduces the dummy pronoun it into the 0-less Case-checking position, Spec. [DP] of the passive was told. Note that this does not invalidate the point in footnote 9—namely that, following Chomsky (1995b), we take the C-Command Condition on chains to be part of the definition of Merge (Chomsky’s Move $\alpha$).
(8) Interarboresial XP-movement via substitution (hypothesisal)

YP, in which neither \( c \)-commands the other. Thus, the operation is illegitimate. The independently motivated Chain Condition (Chomsky's 1995b:235 C-Command Condition) is alone sufficient to exclude interarboresial operations of XP-substitution, and to significantly constrain interarboresial operations involving adjunction.\(^{12}\)

The second part of Rizzi's Chain Condition, (7b)—the "relativized minimality" condition—is discussed at length in Rizzi 1990. Within the framework current at the time, this was stated in terms of government. We state it now solely in terms of \( c \)-command: the requisite relationship between multiple instances of a single element must be of minimal length (i.e., local). The notion "shortest" or "closest" cannot be defined on interarboresial operations per se; rather, it must be defined on the links of the chains created by such operations. Thus, locality must be stated in terms of Relativized Minimality (7b) or the Minimal Link Condition (Miyagawa 1993, Chomsky 1993). Again, the definition of locality need not be a condition on final representa-

\(^{12}\) Note that certain interarboresial XP-adjunction structures would not run afoul of the Chain Condition in this way. For instance, if the DP in (8) were to adjoin to YP—instead of substituting for the specifier of YP—then the DP would \( c \)-command the other instance of DP in the complement-to-X position. Again, a feature-checking theory may suffice to exclude the relevant derivations, although we leave the matter open.

For an alternative analysis of interarboresial XP-movement, see Nunes's (1993:chap 4, 1996) analysis of parasitic gaps and across-the-board extractions. Although similar in spirit to ours, it differs in its assumptions about chain formation and the Chain Condition.

4 Conclusions

The theory developed in Chomsky 1993, 1995a,b does not exclude the possibility of interarboresial operations for any principled reason. Nevertheless, this possibility has remained unexplored in the current literature. If interarboresial operations are not permitted by Universal Grammar, then this is a gap in the theory that awaits a principled explanation. We have suggested instead that this is not a flaw of the theory. In particular, we have shown that admitting interarboresial operations provides a solution to the apparently problematic nature of head movement for the ER. Furthermore, we have shown that the Chain Condition constrains interarboresial operations in such a way as to exclude interarboresial XP-substitution, thereby significantly restricting the generative power of these operations. Further empirical consequences of our proposals are left for future research.

References


In the preminimalist framework (see Chomsky 1986), it was standardly assumed that structural Case licensing took place in three distinct configurations. First, nominative Case in constructions such as (1a) was licensed in the specifier-head relation with I. In the split-l framework (e.g., Pollock 1989, Belletti 1990, Chomsky 1991), in which I is split into several distinct heads, nominative Case licensing took place under specifier-head agreement with the complex Agr0 + Tense functional head, formed by raising Tense to Agr0.

(1) a. He likes her.
   b. \[ \lambda_{\text{he}} \text{Agr}_0 + \text{Tense}_0, [\text{vp} \quad \{ \text{vp likes her}\}] \]

Second, structural accusative Case in simple transitive constructions such as (1a) was assumed to be licensed in the head-complement relation with V. Third, exceptional accusative Case was assumed to be licensed in the head-specifier of the complement configuration. Thus, the matrix verb in constructions such as (2) was assumed to Case-license the embedded-clause subject across the maximal projection that functions as the complement of the matrix verb.

(2) the DA [\text{vp proved} \lambda_{\text{the defendants}} \text{to be guilty}]\]

Chomsky's (1991) modification of Pollock's (1989) split-l hypothesis, which split I into Tense and two Agr phrases, opened up the possibility of reducing all structural Case licensing to a single structural configuration, which was fully exploited in Chomsky and Lasnik (1993). Chomsky and Lasnik suggested that structural accusative is checked under specifier-head agreement with the complex Agr0 + V head, formed by V-movement to Agr0. To be Case-checked, the accusative NPs in (1) and (2) then undergo movement to [Spec, Agr0], a process I will refer to as object shift. (The Agr0 + V complex is assumed to undergo further head movement from the Agr0 position.)

(3) a. he [\text{Agr}_0 + \text{likes}_1, \lambda_{\text{he}} \{\text{vp} \quad \text{its}\}] 
   b. the DA [\text{Agr}_0 + \text{proved}, \lambda_{\text{the defendants}} \{\text{vp} \quad \text{to be guilty}\}] \]

Under Chomsky and Lasnik's proposal, all structural Case licensing takes place in a single configuration of specifier-head agreement, mediated by an Agr head—an analysis that is conceptually clearly preferable to the one in which structural Case licensing takes place in three distinct configurations. The crucial step in reducing all structural Case licensing to a single relation is the proposal that accusative Case licensing is achieved via object shift, which in itself makes the adoption of object shift desirable on conceptual grounds.

Recent work has uncovered more empirical arguments for object shift (see, e.g., Bošković, to appear a, Branigan 1992, Hornstein 1994, for helpful discussion and comments, I thank Howard Lasnik and anonymous reviewers. This work was supported in part by NSF grant SBR-951088.