Remarks on `Conditions on Transformations' (with George M. Horn)

Emmon W Bach
George M Horn

Available at: https://works.bepress.com/emmon_bach/29/
Remarks and Replies

Remarks on “Conditions on Transformations”
Emmon Bach, George M. Horn

1. Introduction
In recent years a number of linguists have tried to formulate constraints on transformational and other rules. Such constraints are of two kinds: (1) They can be part of the definition of a grammar and thus limit the class of formal systems available to the language acquirer. A constraint on the systems of structural analyses available for transformations—for example, that they contain a finite set of terms—is such a condition. We call these constitutive constraints. (2) The constraints may have nothing to say about the constitution of a grammar but instead govern the applicability of rules freely formulated. For example, Ross (1967) proposed a number of constraints on the applicability of various types of transformations. We call such constraints applicability constraints.¹

Most recently, N. Chomsky (1973), drawing on earlier work of his own and others, has proposed a set of conditions governing the applicability of transformational and interpretive rules. In this article, we will examine a number of Chomsky’s conditions and, in passing, several other proposals from recent literature or oral tradition. We will try to show the inadequacies of Chomsky’s proposals. More positively, we will argue that some of the conditions are reflections of constitutive conditions, while others should be replaced by a more uniform and general applicability condition that seems to come somewhat closer to making the right predictions.

In principle, applicability constraints can be part of the general theory about possible human grammars, that is, part of universal grammar, or they can be particular to a language. We assume that a universal constraint is preferable to a language-particular one. Hence, our attack will be two-pronged. We will show that in some instances cross-linguistic data contravene the universal nature of the condition in question. And we will show that the conditions in question cannot be maintained even as being particular to English.² Of course, if the conditions are not correct for English,

¹ We ignore here the distinction drawn by Chomsky between absolute prohibitions and constraints that are part of the evaluation metric. Notice that both constitutive and applicability constraints can be either formal or substantive in nature (hence our innovative terminology for Chomsky’s “formal” and “functional”).
² Chomsky (1973, 233) makes a specific disclaimer as to the universal validity of his conditions.
they are not universal. But if it could be shown that some constraint had strong cross-linguistic validity, we would be inclined to consider reanalyses of English grammar of the sort undertaken by Chomsky in order to bring it into line with the universal condition. As we will see, a number of the constraints are extremely powerful. If constraints of such power are part of the machinery of particular grammars, general linguistic theory is considerably weakened.

We will take up in turn the Specified Subject Condition, Subjacency, the Subject Condition, the Tensed Sentence Condition, and Chomsky’s formulation of Strict Cyclicity. Since a number of Chomsky’s conclusions rest on his analyses of such rules as Wh Movement, we will devote a separate section to those analyses. In each case we will briefly summarize Chomsky’s statement of the condition or analysis in question and his supporting evidence. In the first three instances we will argue that a generalized NP Constraint (Horn (1974)) avoids the problems of Chomsky’s conditions and gives better predictions. In the case of the Tensed Sentence Condition, we will show that the condition fails and conclude that the support given to it by Chomsky rests on an incorrect analysis of sentences involving Raising into object position (Postal (1974)). Finally, we will take up some problems involving the principle of Strict Cyclicity and suggest a revision that seems to come somewhat closer to the correct predictions about the applicability of cyclic rules.

2. The Specified Subject Condition

The Specified Subject Condition is stated as follows (Chomsky (1973, 244); henceforth we cite this article merely by page number):

\[(1) \text{ No rule can involve } X, Y \text{ in the structure:} \]
\[\ldots X \ldots [\alpha \ldots \bar{Z} \ldots -W Y V \ldots \ldots] \ldots\]
where \( \bar{Z} \) is the specified subject of \( W Y V \) in \( \alpha \)

The symbol \( \alpha \) represents NP or S; a specified subject is a subject NP that contains lexical items or a pronoun that is not controlled by the minimal major category containing \( X \). In Chomsky’s framework, NP and S are cyclic categories that differ in that sentences contain a complementizer, while noun phrases do not.

Chomsky uses this condition to prevent a rule like Each Movement from applying in a structure like (2a) to derive the unacceptable (2b):

\[(2) \text{ a. the men each expected [sCOMP the soldier to shoot the other]} \]
\[\text{ b. *The men expected the soldier to shoot each other.}\]

3 We will show below that phrases like a picture of John can instantiate two structures.

4 There is some difficulty in understanding what Chomsky means by involve. For the time being we take it to mean 'mention as constant terms'.

5 Since Chomsky intends his conditions to cover both transformational rules and interpretive rules, it does not matter whether we (or he) consider the sentences of (2) to be transformationally related or not (see (238, fn. 17), but also Postal (1974, 71–76)).
Each Movement can apply in the following examples:

(3) a. the candidates each expected \([8\text{COMP PRO to defeat the other}]\)
   
   b. The candidates expected to defeat each other.

The subject of the complement sentence in (2a), *the soldier*, is a specified subject, while
the subject of the complement sentence in (3a), *PRO*, is not, because it is controlled
by the minimal major category containing \(X\).

Extending the notion of subject to the possessive (italicized) NP in the following
examples, Chomsky, uses the same principle to account for the ungrammaticality of
(4a) as opposed to (4b):

(4) a. *Who did you see John’s pictures of?
   
   b. Who did you see pictures of?

The underlying structures of these examples are assumed to be roughly as follows:

(5) a. COMP you saw \([\text{NP John’s pictures of WHO}]\)
   
   b. COMP you saw \([\text{NP pictures of WHO}]\)

If \(\alpha = \text{NP}, X = \text{COMP}, Z = \text{John’s}, \text{and } Y = \text{WHO}, \) it is easy to see how the
Specified Subject Condition will block (4a) but allow (4b).

Chomsky distinguishes examples like (4a) from ones like (6a) by a special clause
allowing extraction over a specified subject by movement into a COMP node and
further movement from that node to other COMP nodes. This mode of escape also
operates to override other constraints.

(6) a. Who did we believe that Bill hit?
   
   b. COMP\(_1\) we believe \([8\text{COMP\(_2\) Bill hit WHO}]\)

The WH-word in (6b) can move over the specified subject *Bill* into the COMP\(_2\)
node and from there to COMP\(_1\), while no such movement can take place in (5a) since
the NP has no COMP (we take up the analysis of *Wh Movement* and other such rules
in section 7).

The Specified Subject Condition, together with the COMP-to-COMP analysis,
predicts that movement rules that do not involve movement into a COMP node and
controlled deletion rules can never operate over a specified subject, even in an em-
bedded sentence containing a COMP. This prediction, however, is false. Such rules
actually display the same characteristics as Wh Movement. We will show that the
rules of Too-Adj-Deletion and Ready Deletion must be analyzed as applying over a
specified subject in examples like the following, which we will show to have under-
lying structures like those in (8):\(^6\)

---

\(^6\) Lasnik and Fiengo (1974) discuss a number of constructions under the heading of Complement Object
Deletion. They do not consider sentences with *ready*, which seem to differ markedly from the examples they
study. For example, although there may be some question about the grammaticality of deletion into passive
agents with *too* and *enough* (as well as *easy*), examples with *ready* seem completely uninterpretable:
(7) a. That tribe is too unimportant for scientists to study.
   b. The house is ready for John to buy.

(8) a. that tribe is too unimportant [s for scientists to study it]
   b. the house is ready [s for John to buy it]

The Specified Subject Condition imposes an analysis like the one proposed by
Lasnik and Fiengo (1974) on such examples, that is, one in which the phrases for
scientists and for John are not the subjects of the complement structures as in (8):

(9) a. that tribe is too unimportant [pp for scientists] [vpto study it]
   b. the house is ready [pp for John] [vpto buy it]

In these structures, the deletion rule would apply to the object of a subjectless VP and
would hence not be blocked by the Specified Subject Condition.

There are a number of arguments against the assumptions embodied in (9) and in
favor of those of (8). Note that there are structures that contain independent for-
phrases, such as the following, assigned the structure (10b) (Chomsky, 263):

(10) a. It is pleasant for the rich to do the hard work.
   b. it is pleasant for the rich [s COMP PRO to do hard work]

But the sentences of (7) differ in a number of respects from sentences like (10a).
First, in examples like (10a) the for-phrase can be moved freely, while this is not
the case in examples like (7):

(11) a. For the rich, it is pleasant to do hard work.
   b. *For John, the house is ready to buy.
   c. For scientists, that tribe is too unimportant to study.

Example (11c) is well-formed, but it has a different meaning from its putative source
(the structure underlying (7a)).7 This difference is consistent with the difference
between complement subjects and independent for-phrases in clear cases like these:

(12) a. I arranged for the manuscripts to be sent to Rome.
   b. *For the manuscripts, I arranged to be sent to Rome.

(13) a. I bought for my son a 12-string Martin.
   b. For my son, I bought a 12-string Martin.

Second, examples like (7), unlike examples like (10), cannot contain two for-
phrases:

7 The difference in meaning can be brought out more clearly in examples like these:

(i) That tribe is too unimportant for anyone to study.
(ii) For anyone, that tribe is too unimportant to study.
REMARKS AND REPLIES

(14) a. It is pleasant for the rich for the poor to do the hard work.
   b. *That tribe is too unimportant for scientists for anyone to study it.\(^8\)
   c. *That house is ready for John for Bill to buy it.

Third, Right Node Raising applies in general to single constituents (Bresnan (1974)). If the structures of (8) are correct, we should be able to apply the rule to them:

(15) a. The pizza is too hot and the soup is too cold—for anyone to eat.
   b. The moussaka is ready and Mike says that the egg-lemon soup is almost ready—for us to eat.

If the *for*-phrase and the verb phrase are separate constituents as in (9), Right Node Raising should not apply to them.

Finally, Stillings (1975) has shown that proper formulation of Gapping requires the material following the gap in the right conjunct to be a single constituent. We predict that on the analysis (8), the following sentences will be grammatical:

(16) a. Bill is too tall for us to hire, and Will—for us to fire.
   b. The kidney pie is ready for us to put in the oven, and the salad—for you to put on the table.

For these reasons, we believe that the structure of examples like (7) is as in (8) and that they therefore constitute direct counterevidence to the Specified Subject Condition. Independent evidence for our conclusion is that free *for*-phrases in such sentences are restricted to animate nouns. But we can have deletion across such phrases with inanimate nouns:

(17) a. For a minor explosion, that building is too solid to demolish.
   b. That building is too solid for a minor explosion to demolish.

It is interesting to note that the object-deletion rules pattern just like movement rules such as *Wh* Movement in operating into NPs (in Chomsky’s analysis). Compare the following examples with examples (4a) and (4b) above:

(18) a. That tribe is too important for us to see pictures of.
   b. *That tribe is too important for us to see John’s pictures of.

(19) a. The house is ready for us to take pictures of.
   b. *The house is ready for us to take John’s pictures of.

Here it appears that the rules obey the Specified Subject Constraint when no COMP is involved. However, if they are deletion and not movement rules, the presence or

\(^8\) There is a grammatical reading of this sentence in which *for scientists* is a direct modifier of *unimportant*. Thus we can actually have three *for*-phrases in such sentences:

(i) For anyone with any brains, this effect is too important for the well-being of the species for us to ignore it.
absence of COMP should make no difference since there is no possibility of “escape” through COMP.

Another deletion rule that should obey the Specified Subject Condition is Comparative Deletion. A detailed discussion of this rule is found in Bresnan (1975), where it is shown that Comparative Deletion obeys various island constraints, including Chomsky’s conditions. The important point here is that this rule can apply into complement sentences with a specified subject. As above, no movement into a COMP node is involved and the rule should not apply. Chomsky’s analysis predicts that examples like the following should be ungrammatical:

(20) Bill saw more movies than Harry saw.
(21) Joe read more books than John read magazines.

These examples have structures something like the following when the rule(s) apply:

(22) a. Bill saw more movies [sCOMP Harry saw x-many movies]
    b. Joe read more books [sCOMP John read x-many magazines]

In each case, the deleting agent and the deleted material are separated by a specified subject and Comparative Deletion is blocked by Chomsky’s condition.

This rule, like the rules of TOO-Adj-Deletion and Ready Deletion, appears to obey the Specified Subject Condition in examples like these:

(23) a. We saw pictures of more people than they saw pictures of.
    b. *We saw pictures of more people than they saw John’s pictures of.

In example (23a), the rule operates into an NP (in Chomsky’s analysis) that does not contain a specified subject. It is blocked, however, in (23b), which does contain a specified subject. Thus, these deletion rules pattern like Wh Movement.

It is, of course, possible to reanalyze all deletion rules as movement rules involving movement into a COMP node in order to account for the similarities (see Chomsky (253, fn. 32)). But this move seems to us misguided, especially when we consider some additional problems.

We encounter one of these when we consider Chomsky’s rule of It Replacement (Tough Movement). This rule derives examples like (24b) from structures like (24a).

(24) a. It is hard for me to look at Walter.
    b. Walter is hard for me to look at.

To derive (24b), the rule of It Replacement must move Walter from the complement sentence to replace it in the matrix. The underlying structure is something like this:

(25) IT is hard for me [sCOMP PRO to look at Walter]

---

9 We follow Chomsky’s analysis here. Lasnik and Fiengo (1974) analyze such sentences as resulting from Complement Object Deletion. If they are correct, then the examples are irrelevant to the point being made here, but examples like (26) remain counterexamples to the Specified Subject Constraint.
In order to account for the grammaticality of sentences like (26), Chomsky analyzes *It* Replacement within the framework of the Subjacency Condition (see section 3 below) as a replacement of the PRO subject in each complement by the noun phrase being moved by the rule. On the highest cycle, this noun phrase replaces IT:¹⁰

(26) Walter is hard for me to stop Bill from looking at.

This example is derived from the following structure:

(27) IT is hard for me [S₁COMP₁ PRO₁ to stop Bill from [S₂COMP₂ PRO₂ looking at Walter]]

If Walter were to replace IT on the last cycle in one step, the rule would violate the Subjacency Condition, since the moved constituent would be passing over two cyclic nodes, S₁ and S₂. The rule cannot move Walter to COMP₁ and then substitute it for IT, since an element can move from a COMP node only into another COMP. Chomsky solves his problem by moving Walter to PRO₂ on the S₂ cycle, then to PRO₁ on the next cycle, and finally to IT on the S₀ cycle. *It* Replacement is a movement rule that does not involve movement into a COMP node. Therefore, the COMP-to-COMP escape hatch is inoperative and *It* Replacement, like the deletion rules previously discussed, should never apply over a specified subject. Some evidence to support this result is provided by Chomsky. He invokes the Specified Subject Condition to block examples like the following:

(28) *Walter is tough for me to stop Bill’s looking at.

This example is almost a paraphrase of example (26) above. It is derived from the following structure:

(29) IT is tough for me [SCOMP PRO to stop [SCOMP Bill’s looking at Walter]]

This structure differs from structure (27) in that it contains a lexically specified subject instead of a PRO as subject of the clause that contains Walter. Therefore Walter must jump from the PRO subject of the first embedded S over Bill, in violation of the Specified Subject Condition.

The Specified Subject Condition, however, will wrongly block examples like (30) and (31):

(30) *It is reckless for me to stop Bill from looking at Walter.
(31) *It is reckless for me Walter to stop Bill from looking at.

¹⁰ Notice that if the wrong predicate is present in the highest sentence in examples parallel to (26), some filtering mechanism or derivational constraint must be called upon to prevent results like these:

(i) It is reckless for me to stop Bill from looking at Walter.
(ii) *Walter is reckless for me to stop Bill from looking at.
(iii) *It is reckless for me Walter to stop Bill from looking at.

This difficulty is inherent in successive cyclic rules of the sort invoked by Chomsky here. We take up this general question in sections 7 and 9 below.
Each of these examples contains a lexically specified subject, not PRO, in the clause from which extraction takes place (and (31) involves a tensed sentence; see section 6 below). Thus in each case Chomsky’s rule must move Walter into the next clause in violation of his condition.

As with the other rules we have discussed, It Replacement appears to obey the constraint when it operates into NPs:

(32) John was tough for us to take pictures of.
(33) *John was tough for us to take Bill’s pictures of.

Since no movement into COMP is involved in any of these cases, it is not possible to account for the difference between the noun phrase examples and the embedded sentence examples. It appears once more as if the rule can operate over a specified subject in complement sentences but not in NPs.

We have considered a number of direct counterexamples to the Specified Subject Condition. We can also argue against it on the grounds that it is not strong enough to block examples like the following, in which movement is forbidden even when the NP in question contains no specified subject.

(34) a. *Who did they destroy pictures of?
   b. *Who did you see a book about?
   c. *Who did John burn a large green book about?

These examples contrast with the following ones, in which movement is allowed:

(35) a. Who did you see pictures of?
   b. Who did John write a book about?

Notice that all of the rules we have discussed pattern alike in this regard:

(36) a. *That tribe is too unimportant for us to destroy books about.
   b. *The house is ready for us to demolish the roof of.
   c. *Jill destroyed books about more people than Jack destroyed books about.
   d. *Harriet was tough for us to destroy books about.

11 We assume that anyone is the subject of the following complement rather than an object of imagine, since the entire phrase passes all tests for constituency. Compare also (i), suggested to us by Joan Bresnan:

(i) Walter is hard for me to imagine there being such a fuss about.
Compare the following:

(37) a. That tribe is too unimportant for us to write books about.
b. The house is ready for us to take a picture of.
c. Jill wrote books about more politicians than Jack wrote books about.
d. Harriet was tough for us to write books about.

To summarize, the Specified Subject Condition has the effect of allowing movement out of complement sentences with lexical subjects via their COMP nodes, or by PRO hopping in the case of *It Replacement, while blocking movement out of NPs with possessive NP determiners because they contain no COMP. However, we have seen that with respect to this putative condition all transformations, whether they are movement or deletion rules and whether or not they involve movement into a COMP node, display the same syntactic properties. They can extract from, or delete in, sentences with specified subjects but not NPs with specified subjects. These facts suggest that the presence or absence of a specified subject is not crucial to the full range of extraction and deletion facts we have been considering.

After discussing the Subjacency and Subject Conditions, we will propose an alternative account for the whole range of extraction and deletion phenomena; in particular, for the differences between (34) and (36), on the one hand, and (35) and (37), on the other. The constraint to be proposed will allow us to account for facts like those illustrated in (38) without resort to a Specified Subject Condition:

(38) a. Who did you see pictures of?
b. *Who did you see John’s pictures of?

3. Subjacency

The notion of Subjacency is defined as follows (247):

(39) If $X$ is superior to $Y$ in a phrase marker $P$, then $Y$ is subjacent to $X$ if there is at most one cyclic category $C \neq Y$ such that $C$ contains $Y$ and $C$ does not contain $X$.

(“A category $A$ is ‘superior’ to the category $B$ in the phrase marker if every major category dominating $A$ dominates $B$ but not conversely” (Chomsky, 246).) If $Y$ is subjacent to $X$, then $Y$ is either in the same cyclic structure—NP or $S$—or $Y$ is only one cyclic structure “down” from $X$. In structure (40a) $WHO$ is subjacent to both COMP nodes, while in structure (40b) it is subjacent only to the second, since it is separated from the first by two cyclic nodes:

(40) a. COMP he believes $[sCOMP John saw WHO]$
b. COMP he believes $[NP the claim [sCOMP John saw WHO]]$
Having defined Subjacency, Chomsky places the following condition on the applicability of transformations:

\[(41) \text{No rule can involve } X, Y, X \text{ superior to } Y, \text{ if } Y \text{ is not subjacent to } X.\]

The Subjacency Condition accounts for many of the facts that led to the formulation of Ross's Complex NP Constraint. The condition will block movement of the WH-word to the initial COMP in structure (40b) but will allow movement in (40a):

\[(42) \begin{align*}
\text{a. } & \text{Who did he believe that John saw?} \\
\text{b. } & \text{*Who did he believe the claim that John saw?}
\end{align*} \]

Extraction from relative clauses such as the one shown in the following structure is likewise prohibited, accounting for the ungrammaticality of (43b):

\[(43) \begin{align*}
\text{a. } & \text{COMP we saw [NP the man [sCOMP who knew WHO]]} \\
\text{b. } & \text{*Who did you see the man who knew?}
\end{align*} \]

In addition, according to Chomsky, the Subjacency Condition will block sentences to which the Complex NP Constraint does not apply:

\[(44) \begin{align*}
\text{a. } & \text{*Who did you write articles about pictures of?} \\
\text{b. } & \text{*What do you receive requests for articles about?}
\end{align*} \]

But it will allow examples like these:

\[(45) \begin{align*}
\text{a. } & \text{What do you write articles about?} \\
\text{b. } & \text{What do you generally receive requests for?}
\end{align*} \]

Chomsky assumes that the underlying structures for the examples in (44) are something like this:

\[(46) \begin{align*}
\text{a. } & \text{COMP you write [NP particles about [NP pictures of WHO]]} \\
\text{b. } & \text{COMP you receive [NP requests for [NP particles about WHAT]]}
\end{align*} \]

In each case, the WH-word is two cyclic categories removed from the initial COMP, and Wh Movement cannot apply because of the Subjacency Condition. On the other hand, the structures of examples (45a,b) allow movement:

\[(47) \begin{align*}
\text{a. } & \text{COMP you write [NP particles about WHO]} \\
\text{b. } & \text{COMP you generally receive [NP requests for WHAT]}
\end{align*} \]

Examples like (48) seem to show that Wh Movement is not always constrained by Subjacency:

\[(48) \text{Who did Bill believe that John told Ralph to kill?}\]

This example has the following structure:

\[(49) \text{COMP Bill believes [sCOMP John told Ralph [sCOMP PRO to kill WHO]]}\]
REMARKS AND REPLIES

To derive (48), the WH-word must move from its original position over two cyclic nodes. Chomsky accounts for this by successively moving the WH-word first into the COMP of the most deeply embedded S and then from COMP to COMP on each cycle, thus in accordance with the Subjacency Condition. We will deal with this analysis below (section 7).

As mentioned before, It Replacement is also formulated in such a way as to be consistent with the Subjacency Condition. However, even the PRO-hopping re-analysis of this rule cannot provide an account of grammatical sentences such as (50), which would be blocked by the Subjacency Condition:

(50) Harriet is hard for me to imagine Bill wanting to kiss.

The source of (50) is (51):

(51) IT is hard for me [s₁COMP PRO₁ to imagine [s₂COMP Bill wanting [s₃COMP PRO to kiss Harriet]]]

On the S₃ cycle Harriet replaces PRO. PRO-hopping cannot apply on the second cycle because the subject of S₂ is not PRO but Bill. Therefore, to derive (50) we must allow Harriet to move from S₃ to S₀ in one step, thus violating the Subjacency Condition (as well as the Specified Subject Condition).

Another rule that presents problems for the Subjacency Condition is Cleft Movement (Chomsky (1970)). This rule derives examples such as (52a) from structures like (52b), by moving the NP John to the empty (△) position in the matrix sentence:

(52) a. It was John that Bill hit.
   b. it BE △ [sCOMP Bill hit John]

It seems clear that the rule of Cleft Movement does not involve cyclic movement from COMP to COMP, since the NP in question does not end up in COMP position but rather after the copula.

Now consider the following example:

(53) It was Harriet that Bill wanted John to kiss.

(53) is derived from the following structure:

(54) it BE △ [sCOMP Bill want [sCOMP John to kiss Harriet]]

To derive (53) we would have to violate the Subjacency Condition, since substitution of Harriet for △ would reach across two cyclic nodes.

Although they seem to be included in Chomsky's first statement of the Subjacency Condition, deletion rules are excluded in his final formulation (247–248):

(55) Consider the structure α, where α is a cyclic node:
    [s ... X ...] △ ...

If ... T is an extraction rule moving some category in ... to the position
\( X \), then some constant term of \( \Sigma \) [the structural description of \( T \)] must hold of a category subjacent to \( X \) in \( \alpha \) for \( T \) to apply to \( \alpha \).

Indeed, the following examples indicate that deletions are not constrained by subjacency:

\[
\text{(56) a. That tribe is too unimportant for the government to pay scientists to study.}
\]
\[
\text{b. The house is ready for John to persuade Bill to buy.}
\]
\[
\text{c. Bill saw more movies than Henry believed that Ralph saw.}
\]

In each case, the deletion rules in question cross two cyclic nodes.

However, there are some sentences in which it appears that deletion rules obey the Subjacency Condition. Like movement rules, it looks as if these rules cannot operate into complex NPs.

\[
\text{(57) a. *That tribe is too unimportant for us to find a man willing to study.}
\]
\[
\text{b. *The house is ready for us to find a man willing to buy.}
\]
\[
\text{c. *Jill interviews people who know more interesting things than Jeff interviews people who know.}
\]

Nor can these rules operate into structures such as the following ones, which have NPs embedded in NPs:

\[
\text{(58) a. *That tribe is too unimportant for us to write books about pictures of.}
\]
\[
\text{b. *The house is ready for us to take pictures of the basement of.}
\]
\[
\text{c. *Jill wrote articles about the exploits of more politicians than Bill wrote articles about the exploits of.}
\]

If the Subjacency Condition is applied to these rules in these particular cases, the previous examples in which they appeared to violate that principle become real counterexamples. If, on the other hand, the Subjacency Condition applies only to extraction rules, we cannot invoke it to explain the ungrammaticality of (57) and (58).

In addition to these problems, there are indications that the Subjacency Condition is too weak. As we have seen, it correctly blocks examples like (59a) while allowing ones like (59b):

\[
\text{(59) a. *What did you write articles about pictures of?}
\]
\[
\text{b. What did you write articles about?}
\]

However, it is not the case that NPs such as those in (59b) always allow extraction. Consider the following examples:

\[
\text{(60) a. *What did you request an article about?}
\]
\[
\text{b. *What did you write about a picture of?}
\]
\[
\text{c. *What did they destroy a book about?}
\]
In Chomsky’s framework, the sources of these examples are \((61a,b,c)\) respectively:

\[
(61) \begin{align*}
\text{a. } & \text{COMP you request } [\text{NP an article about WHAT}] \\
\text{b. } & \text{COMP you write about } [\text{NP a picture of WHAT}] \\
\text{c. } & \text{COMP they destroy } [\text{NP a book about WHAT}]
\end{align*}
\]

In all of them, the WH-word is subjacent to the higher COMP and there is nothing in Chomsky’s analysis to prevent the application of \(Wh\) Movement. Thus, the Subjacency Condition will not account for the ungrammaticality of the examples in \((60)\). Note also that the other movement and deletion rules pattern in the same way:

\[
(62) \begin{align*}
\text{a. } & \text{*It was John that they destroyed a book about.} \\
\text{b. } & \text{John is hard for us to destroy books about.} \\
\text{c. } & \text{*That tribe is too unimportant for us to destroy books about.} \\
\text{d. } & \text{*The house is ready for us to renovate the basement of.} \\
\text{e. } & \text{*John destroyed books about more politicians than Hank destroyed books about.}
\end{align*}
\]

In all these cases, an NP of the form \([\text{NP NP P NP}]\) blocks extraction or deletion of its internal elements even though they are subjacent to the constant \(X\) of the rule in question.

We have seen that elements can be freely extracted from and deleted in sentences, even ones several cyclic categories removed from their final positions or controllers. The only real effect of the Subjacency Condition is to prohibit extraction from NP configurations such as the following:

\[
(63) \begin{align*}
\text{a.} & \quad \text{NP} \\
& \quad \text{NP} \quad \text{S} \\
& \quad \triangleleft \text{X} \\
\text{b.} & \quad \text{NP} \quad \text{NP} \\
& \quad \triangleleft \text{X}
\end{align*}
\]

Structure \((63a)\) represents complex NPs with lexical heads, and structure \((63b)\) represents NPs embedded in NPs with no intervening S node (examples \((44a,b)\)).

In fact, deletion rules cannot operate into such structures either, even though the Subjacency Condition is not supposed to apply to them. Moreover, this restriction is not sufficient to account for the fact that no extraction or deletion is allowed from even less complex NPs where subjacency is not violated. All of this suggests that there
is a possible generalization about extraction and deletion from NPs that has nothing to do with subjacency. Before presenting this generalization, we consider a third condition.

4. The Subject Condition

Chomsky formulates the Subject Condition to account for the ungrammaticality of examples like (64a,b), which contrast with (65a,b):

(64) a. *Who did stories about terrify John?
   b. *Who were books about written by Bill?

(65) a. Who did John hear stories about?
   b. Who did Bill write books about?

He assumes that the examples in (64) are derived from structures like the following:

(66) a. COMP [NPstories about WHO] terrified John
   b. COMP [NPbooks about WHO] were written by Bill

In each case the WH-word is moved from the subject of the sentence. In (65) the WH-word is not in the subject. Chomsky proposes that this difference is crucial and formulates the Subject Constraint to account for it:

(67) No rule can involve $X$, $Y$ in the structure
     \[ \ldots X \ldots [a \ldots Y \ldots] \ldots \]
     where $a$ is a subject phrase properly containing the minimal major category containing $Y$.

The Subject Condition is not strong enough to block extractions from nonsubject NPs such as the following, which are of the same type as those in the examples above:

(68) a. *Who did they destroy a book about?
   b. *Who did you believe stories about?

Nevertheless, such sentences are ungrammatical. These examples suggest that extraction from subject NPs is only part of a more general restriction.

It should be mentioned that even to account for the range of data that Chomsky considers under the heading of the Subject Condition, he is forced to deny the existence of a rule raising from subject into object position. Thus the italicized NPs in the structures of (69) must still be subject of their complement clauses when extraction takes place if it is the Subject Condition that is to block them:

(69) a. COMP you expect stories about WHO to terrify John
   b. COMP you believed a book about WHO to have been written by John

(70) a. *Who did you expect stories about to terrify John?
   b. *Who did you believe a book about to have been written by John?
If there is a rule of Raising, and there is a considerable amount of evidence that there is (Postal (1974), Bach (to appear)), then the Subject Condition will not block the ungrammatical examples of (70). (More on Raising in section 6 below.) The proposal made in the next section is indifferent to the existence of Raising.

5. The NP Constraint

To summarize briefly, the total effect of the Specified Subject Condition, the Subjacency Condition, and the Subject Condition, taken together with the COMP escape hatch and the reanalysis of other rules, is to block extraction from the following configurations:

A. Noun phrases with possessive NP determiners
B. Complex NPs
C. NPs embedded in higher NPs (with no intervening S node)
D. Subject NPs

An example of each structure is given here along with the ungrammatical sentence that would be derived by extraction from it:

(71) a. COMP you saw [NPJohn’s pictures of WHO]
   b. *Who did you see John’s pictures of?

(72) a. COMP he believed [NPthe claim [SCOMP John killed WHO]]
   b. *Who did he believe the claim that John killed?

(73) a. COMP they wrote [NParticles about [NPpictures of WHO]]
   b. *Who did they write articles about pictures of?

(74) a. COMP [NPa book about WHO] was written by John
   b. *Who was a book about written by John?

Extraction rules can operate across a specified subject in a sentence if they move the constituent into a COMP node. They cannot operate across a specified subject in an NP, since NPs do not contain COMPs. Deletion rules, which by definition involve no movement into COMP, and other movement rules that involve no movement into a COMP node should be blocked by the presence of a specified subject in an NP or S. Extraction rules cannot operate into configurations (B) and (C) above because of the Subjacency Condition and because these configurations have no COMP. Since the Subjacency Condition does not apply to deletion rules, there is no such restriction on them.

Contrary to these predictions, all the extraction and deletion rules we have examined display the same syntactic behavior. Both rule types can operate into sentences with specified subjects but not into NPs with specified subjects. Neither type can apply into structures (B) and (C).
The three conditions prohibit extractions from all types of NP except those of the form \([\text{NP}\cdot\text{NP}\cdot\text{PP}]\) with no POSS determiner. However, even these do not allow extraction in many cases. This fact is illustrated by the following examples in which the object of the preposition and the entire PP in such structures is extracted.

(75) a. *Who did they destroy a book about?
   b. *About whom did they destroy a book?

(76) a. *Who did he lose a picture of?
   b. *Of whom did he lose a picture?

(77) a. *What did Einstein attack a theory about?
   b. *About what did Einstein attack a theory?

(78) a. *Which city did Jack search for a road into?
   b. *Into which city did Jack search for a road?

(79) a. *What did Kissinger prevent a war over?
   b. *Over what did Kissinger prevent a war?

Extraction is also prohibited from nominalizations:

(80) a. *What did you discuss the growth/growing of?
   b. *Who did Chomsky review the criticism of?

Extraction of the head NP is blocked by the Left Branch Condition or the A-over-A Condition. Deletion rules cannot operate into any of these structures either.

We can account for all of these facts if we quite generally prohibit extraction or deletion from a noun phrase. To do this we propose the following condition on applicability (Horn (1974)):

\[\text{The NP Constraint}\]

No constituent that is dominated by NP can be moved or deleted from that NP by a transformational rule.

(If there are such things as free deletions, this constraint does not apply to them).

In the remainder of this section we will discuss briefly a set of apparent counterexamples to this generalization, of which the following are representative:

(82) a. Who did John write a book about?
   b. Who did Bill take a picture of?
   c. Who do government employees see pictures of every day?

For a more complete discussion of the NP Constraint, see Horn (1974). We will not attempt to present a final formulation of it here.\(^{12}\)

\(^{12}\) In Horn (1974) an extended discussion and reanalysis of complement structures are given, which also provide a number of apparent counterexamples to the NP Constraint. Note that (81) is an applicability condition (sections 1, 9).
The examples in (82) contrast with examples (75)–(80) in that the object of the preposition, presumably dominated by the object NP, may be extracted in (82) but not (75)–(80). There are other differences as well.

First, the entire PP can be fronted in (82) but not in the other examples:

(83) a. About whom did John write a book?
   b. Of whom did they take a picture?
   c. Of whom do government employees see pictures every day?
   d. *About whom did they destroy a book?
   e. *Of whom did they lose a picture?

Moreover, in examples like (82) the NP direct object can move independently of the PP, while this is impossible in examples like (75)–(80):

(84) a. A book was written about Nixon by John.
   b. A picture was taken of Sam by Jim.
   c. Pictures are seen of Gerald Ford every day.
   d. *A book was destroyed about Nixon by John.
   e. *Pictures are lost of Gerald Ford every day by government employees.

These facts, especially the latter one, indicate that the NP direct object and the PP are independent constituents in examples like (82). Otherwise the Left Branch Condition or A-over-A Principle would prohibit movement. In addition, Ross (1967) noted that the movement of entire PPs from dominating NPs is not generally permissible. Examples (83a–c) are violations of this generalization.

We will therefore assign structures like the following to examples such as (82):

(85) S
    /            
   NP           VP
    /      
   V    NP  PP

Now the movement facts shown in examples (82), (83a–c), and (84a–c) are no longer counterexamples to the NP Constraint. None involve extraction from a higher NP.

The fact remains, however, that the following passives are also possible:

(86) a. A book about Nixon was written by John.
   b. A picture of Ralph was taken by Bill.
   c. Pictures of Gerald Ford are seen every day by government employees.

For this reason let us also assign the following structure to these sentences:
This is identical to the (only) structure assigned to examples like (75)-(79).

There is independent evidence that examples such as those in (82), unlike those of (75)-(79), may be associated with two structures. When certain quantifiers occur in these examples, they are clearly ambiguous. Consider the following sentence:

(88) John wrote his first five books about Nixon in 1965.

On one reading, the scope of the quantifier includes books, and on the other, it includes books about Nixon. Thus, on the first reading we are talking about the first five books John ever wrote, which happened to be about Nixon; but on the second reading, we might be talking about John's sixteenth through twentieth books, but the first five about Nixon that he wrote.

When the same quantifiers occur in sentences like those of (75)-(79), the sentences are unambiguous:

(89) John destroyed his first five books about Nixon.

Here the scope of the quantifier can only be books about Nixon.

Thus the two possible scopes for (88) correspond to the two possible structures for the sentence: (85) and (87). The one scope possible for (89) corresponds to its one possible deep structure, (87). Within the two-structure analysis we can say that the scope of such quantifiers includes the rest of the NP that immediately dominates the quantifier. 13

Notice also that the following sentence is unambiguous:

(90) Who did John write his first five books about?

The scope of the quantifier can only be books and not books about WHO. This fact follows from the principle of interpretation of quantifiers and from the NP Constraint, since we can only front the object of the preposition in structures like (85).

A second argument in favor of the two-structure analysis is provided by pronoun facts. It is not generally possible for a pronoun to be the head of an NP containing modifiers. Thus, the following examples are ungrammatical:

13 Here and in the subsequent discussion it should be noted that the structures in question pattern like clear cases of NP PP sequences (as in (i)) and [NP NP PP] constituents (as in (ii)):

(i) John put his five oldest children to bed.
(ii) I have identified the three most prominent constellations of the southern skies.
(91) a. *I saw the large it yesterday.
b. *We met him who left early.

This holds true for the sentences to which we assign only structure (87):

(92) a. *John destroyed it about Nixon.
b. *Einstein attacked it about relativity.
c. *It was destroyed about Nixon.

In sentences with structure (85), however, definite pronouns can occur:\(^{14}\)

(93) a. Bill wrote it about Nixon.
b. Einstein formulated it about relativity.
c. Henry took it of John.
d. It was written about Nixon.

The two-structure analysis and the NP Constraint will account for most of the data that Chomsky's conditions are intended to explain, as well as accounting for facts not covered by his conditions. Since the analysis applies to deletion rules, general movement rules, and movements into COMP, we predict that the different types of rules will display the same syntactic behavior, and indeed this is the case.

It is easily seen that the NP Constraint will block extraction from configurations (B), (C), and (D) above, and that it will thus block examples (72)–(74), as do Chomsky's conditions. In addition, it will block examples (75)–(80) (see also (34) and (36)), while allowing examples like (82) (see also examples (35) and (37)), which are derived from structure (85). Examples (26), (30a), and (31a), repeated below, are not blocked because they involve movement from complement sentences and not NPs. Examples (30a) and (31a), which are fully grammatical, are blocked by the Specified Subject Condition. Example (28) involves movement from a Poss-Ing construction, which is a basic NP (see Horn (1974, 1975)).

(26) Walter is hard for me to stop Bill from looking at.
(28) *Walter is hard for me to stop Bill's looking at.
(30) a. Walter is hard for me to imagine anyone looking at.
(31) a. Walter is hard for me to imagine that anyone would look at.

Examples like (64a,b) and (74b) are blocked because the NP PP sequence in subject position must be dominated by a single NP node. There is no way to derive them from structure (85), in which the NP and PP are separate constituents. These contrast with examples like (65), which can be derived from structure (85).

\(^{14}\) We are not claiming that sentences like (93) are acceptable in any context, but only that there is a marked contrast between them and those of (92), which seem to allow no mitigating context:

(i) What did you write your book about?
   I wrote it about Nixon.
(ii) I destroyed a book about someone.
    *Oh, who did you destroy it about?
(64) a. *Who did stories about terrify John?
   b. *Who were books about written by Bill?
(65) a. Who did John hear stories about?
   b. Who did Bill write books about?
(74) b. *Who was a book about written by John?

We are left, then, with the task of distinguishing between examples like (94a) and (94b):

(94) a. What did you see pictures of?
   b. *What did you see Bill's pictures of?

If we say that such possessive determiners, and probably other nominal premodifiers, cannot occur in structures of the form (85) with these particular lexical combinations (write a book about, take a picture of, etc.), but freely occur in structures like (87), then (94b) can have only (87) as its source and involves extraction from a NP in violation of the NP Constraint. Note that this is consistent with the interpretation of quantifier scope in phrases like Bill's first five books about Nixon. On the other hand, example (94a) can be derived from structure (87). At present, a more general account of such cooccurrence restrictions has not been developed, nor has a method for distinguishing the possessive determiners in examples like (94b) from pronominal possessives like the his in example (88), and these remain problems for the proposed analysis. But, at any rate, since the NP Constraint and our analysis of the relevant sentences account for a wider range of data in a more general way, they are to be preferred to Chomsky's analysis.15

6. The Tensed Sentence Condition

We have already encountered several counterexamples to the Tensed Sentence Condition. Chomsky states the condition as follows (257):

15 (81) is clearly still too strong. In particular, prepositional phrases containing of seem to be much freer in extractability:

   (i) The table that we waxed the top of was blue.
   (ii) *The table that we waxed the shoes on was blue.

(ii) is grammatical only if the phrase on the table is interpreted as an independent PP in the verb phrase. Measure phrases in of are also quite free:

   (iii) What did you drink a quart of?

Further, speakers who accept sentences like (iv) may be interpreting them on the analogy of sentences like (v):

   (iv) ?Who would you disapprove of John's marrying?
   (v) Who would disapprove of John marrying?

If a verb takes only a Poss–Ing complement and does not allow Acc–Ing, the results of extraction seem much worse:

   (vi) *Who did you defend John's hitting?
   (vii) *We defended John hitting Henry.
REMARKS AND REPLIES

(95) No rule can involve \( X, Y \) (\( X \) superior to \( Y \)) in the structure
\[ \ldots X \ldots \{ \ldots \alpha \ldots \} \ldots \]
where \( Y \) is not in COMP and \( \alpha \) is a tensed S

Chomsky’s initial motivation for (95) comes from examples like these:

(96) a. Everyone believes the dog to be hungry.
    b. The dog is believed to be hungry by everyone.

(97) a. Everyone believes the dog is hungry.
    b. *The dog is believed is hungry by everyone.

(98) a. The candidates each expected the other(s) to win.
    b. The candidates expected each other to win.

(99) a. The candidates each expected that the other(s) would win.
    b. *The candidates expected that each other would win.

(96) and (97) are cited to show that the Tensed Sentence Condition blocks the applicability of Passive, while (98) and (99) show that Each Movement (or an interpretive rule; see fn. 5 above) is governed by the same constraint.

Linguists have often assumed that there is a difference in structure between pairs of sentences like (96a) and (97a) and (98a) and (99a) in that a rule of raising into object position has applied. Chomsky denies this claim. For him, all of the examples have a common structure roughly as follows:

(100) \[
S \\
NP \quad VP \\
V \\
S \\
NP \quad VP
\]

If Chomsky is correct, then the only possible explanation must lie in the difference between a tensed clause (that is, a clause with a finite verb) and one with an infinitive or other nonfinite verbal affix. We will not enter into this controversy directly here (see Postal (1974) for an extended discussion of the issues and Bach (to appear) for an evaluation of the pros and cons of Raising). Rather, we will examine a number of other cases where it appears that the Tensed Sentence Condition leads to incorrect predictions. If our arguments are correct, then we can take the conclusion as evidence against Chomsky’s hypothesis about the structure of sentences like (96a) and (98a).

The most obvious counterexamples to the Tensed Sentence Condition are unbounded extraction rules like Wh Movement, Relative Clause Formation, and the like:

(101) Who do you think should get the prize for linguistic trivia?
I saw the man who Harry said Mary was meeting secretly. As we have seen, Chomsky deals with these counterexamples in a special way. We will defer discussion of them to section 7. But note that Chomsky’s rule of It Replacement, which does not involve movement through a COMP, is also not affected by the condition, as we noted in example (31a):

(31) a. Walter is hard for me to imagine that anyone would look at.

The next large class of rules that violate the Tensed Sentence Condition are anaphora rules of all kinds. Since Chomsky considers his constraint to apply not just to transformations but to interpretive rules as well, it does not matter whether we treat them in one or the other way. First, we should consider ordinary definite pronominalization:

(103) Harriet said that she was going to Ucluelet.

Second, One Pronominalization:

(104) I bought a bicycle bigger than the one that you bought.

Third, VP Ellipsis:

(105) I caught a trout before you did.

Fourth, Do-So:

(106) Mary proved the theorem before John could do so.

Fifth, the rule or rules that give or interpret reflexives of the sort discussed in Ross (1970):

(107) Bill said that the book had been written by Justine and himself.

Sixth, so-called picture-noun reflexives:

(108) Gerald was delighted that a picture of himself was hanging in the hall.

These last two examples are especially damaging to the condition because it is supposed to be the Tensed Sentence Condition that blocks the interpretation or formation of reflexives in sentences like (109) as opposed to sentences like (110):

(109) *Michael believes himself is a millionaire.
(110) Michael believes himself to be a millionaire.

Seventh, the deletion or interpretation of the subjects of gerundives (Tavakolian (1973) has shown that “gerundive-Equi” works in a manner quite distinct from Equi in to-infinitivals):

(111) Dean said that perjuring himself was out of the question.

Chomsky (298, fn. 16) mentions the fact that Coreference Assignment must escape the effect of the condition. Notice also that all rules relating coordinate structures are excluded by the superiority part of the condition.
Now it is true that the examples given may all be characterized as involving anaphora of one sort or another and the rules for anaphora very likely exhibit special properties (see Wasow (1972a)). But so do several of the rules invoked by Chomsky as evidence for his condition. And in any case, if we survey other rules, we find that they violate this condition as well. A miscellany of such rules we have found in the literature follows:

Neg Transportation

(112) I don’t believe he’ll arrive until three.

The can’t seem to construction (Langendoen (1970))

(113) Jerry can’t seem to keep his tenses straight.

Some/Any (in fact a whole range of negative polarity items, see Baker (1970a))

(114) I doubt that he’s got any money.

Comparative Deletion and Comparative Ellipsis (Bresnan (1975))

(115) Life has more pleasures than the beach has grains of sand.

(116) Harry lost more of his hair than Sam did.

Tense Agreement (!; this example suggested to us by Barbara Hall Partee)

(117) John said that he had had a good time.

We cannot claim to have examined every rule of English that the Tensed Sentence Constraint might govern, but among those we have examined the only ones where the Tensed Sentence Constraint yields the correct results are precisely those where Chomsky adopts the controversial view that there is no raising into object position.

The Tensed Sentence Constraint has a curious history, partly documented in the Chomsky (1973). In Chomsky (1965, 146) it was suggested that there was a universal constraint against the insertion of “morphological material into sentences that have already been passed in the cycle” (1973, 234). Chomsky considers the Tensed Sentence Constraint and the Specified Subject Constraint, taken together with their various sub-conditions and exclusion clauses, to represent extensions and sharpenings of this idea. The Insertion Prohibition was quickly accepted as a confirmed hypothesis about universal grammar and actually used as a basis for arguments concerning alternative analyses. For example, Bresnan (1970, 1972) uses this putative universal to argue against the transformational introduction of complementizers.

If it could be shown that the conditions Chomsky proposes are valid for a wide variety of languages, we would have good grounds for accepting reanalyses that would make them work for English. Chomsky (233) specifically disclaims any cross-linguistic validity for his principles, but it remains true that if they are confined to particular languages they represent a considerable weakening of linguistic theory. We will now consider an example from another language that shows that the Tensed Sentence
Constraint has no cross-linguistic validity (other examples have been given for Zapotec by Rosenbaum (1974) and for Swahili by Keach (1975)).

Consider an Amharic sentence like (118):

(118) Yohannis inen habtam now ala.
John me rich is said
‘John said I was rich.’

By any definition we can think of, the internal clause of this sentence is tensed: now is the third person masculine singular form of the verb meaning ‘to be’. The form inen is marked with the definite object marker. (If the speaker is female, (118) is inappropriate and a feminine form must be used.) Moreover, Amharic has regular nonfinite forms of verbs that are used in other constructions.

There are two possible analyses for sentences like (118). The most obvious one is that there is a rule of raising into object position that can apply to the subject of the embedded sentence. This would explain why the pronoun is now marked as the object of ala. It would also explain why Passive can apply to this new object, as it can. If this analysis is correct, we have another example of the incorrectness of the Tensed Sentence Condition as a universal prohibition.17

But even if we adopt the alternative, Chomskyan analysis of (118) and say that no raising has taken place, the Tensed Sentence Condition must fall. For on the one hand “morphological material” (in the form of the definite object marker -n on inen) has been introduced into a tensed clause. Notice that it is presumably the Tensed Sentence Condition that prevents rules from blindly producing or allowing (119) and (120):

(119) *I believe myself is a crackpot.
(120) *I imagine him is here.

And, on the other hand, as we have noted, Passive can apply to the structures immediately underlying (118).18 Thus, if there is a Tensed Sentence Condition, it must be stated individually for the grammar of English.

Returning to examples (96)-(97)

(96) a. Everyone believes the dog to be hungry.
   b. The dog is believed to be hungry by everyone.
(97) a. Everyone believes the dog is hungry.
   b. *The dog is believed is hungry by everyone.

17 This analysis would then also constitute counterevidence to another putative universal, first stated, we believe, in Kiparsky and Kiparsky (1970), to the effect that infinitives are the automatic result of the presence of verbs in clauses having no subject, either as the result of raising or deletion.

18 Another fact of Amharic, discussed in Bach (1970,) is evidence against the Tensed Sentence Condition (as well as the Insertion Prohibition mentioned above). When a relative clause is formed on a noun within a prepositional phrase, the preposition is carried down into the relative clause as a prefix on the verb.
we see that the Tensed Sentence Condition has no independent justification in English (or any language as far as we know) and conclude that Chomsky’s premise that no raising has occurred in (96a) is false.¹⁹

7. The COMP-to-COMP Escape Hatch
At several points in the preceding sections we have noted that counterexamples to constraints on extraction or “involvement” are taken care of by a clause on the conditions that allows movement of an element from a COMP position to another COMP position. For example, we find sentences like (121), involving the Specified Subject (and Tensed Sentence) Conditions, and (122), involving the Subjacency Condition:

(121) Who do you believe that Bill saw?
(122) Who did John say Mary told to come to the party?

Many linguists (Jackendoff (1972), Bresnan (1971), Baker (1970b)) have assumed that rules like Wh Movement or Relative Clause Formation involve a single movement (copy and deletion). Thus the immediately underlying form for (121) would be something like this:

(123) Q you believe that Bill saw WHO

Chomsky, on the other hand, offers an analysis in which the movement occurs in a “successive-cyclic” fashion (Postal (1972)). For example, (121) has in its derivation an intermediate stage in which the interrogative element has been moved into COMP position in the lower sentence:

(124) Q you believe [s[COMP,WHO] Bill saw]

Further, Chomsky argues that the single unbounded movement analysis is inconsistent with an independently needed principle of strict cyclicity (we take up strict cyclicity in the next section). Under this analysis, then, examples like (121) and (122) do not constitute counterexamples to the conditions in question; more precisely, clauses are added to the conditions that allow movement (“involvement”) from a COMP position to a COMP position to override the relevant constraints. There are thus two parts to Chomsky’s claims about the analysis of such examples: (1) movements of this type occur successive-cyclically; (2) once an element has been moved into COMP position, it can move further only into another COMP position.

As Chomsky notes (243, fn. 22 but with no references), a number of linguists have argued specifically against his analysis (Postal (1971, 1972), Bach (1971), now also Bach (1975)). There is no point in repeating these arguments here, but it is worthwhile

¹⁹ We have nothing to say about the Each Movement cases. Postal (1974) concludes that facts about each provide no evidence for or against Raising. In the same place Postal has given a number of arguments against the Tensed Sentence Condition.
checking at least one to see whether, as Chomsky claims, it is inapplicable to his formulation.

Postal (1972) points out that rules like Raising and Tough Movement (or Tough Deletion) would have to be complicated if the WH-element has been moved prior to their application, as in sentences like these:

(125) Who seems to know the answer?
(126) Who is tough for me to stop Bill from looking at?

Specifically, every such rule that deletes or raises out of sentences with COMP elements will have to include a special structural condition to apply to the case where the relevant NP to be raised or deleted has been moved into COMP position (necessarily on an earlier cycle). Further, every such restatement of a rule to include this special case will involve a violation of Chomsky’s second claim that involvement can only hold between two COMP positions once an element is moved into a COMP.

Note that the COMP-to-COMP restriction that is supposed to prohibit the derivation of (127) from (128) (243, fn. 24) must be relaxed to allow the derivation of (129):

(127) *What was asked to read by John.
(128) John asked [s[COMP what] to read]
(129) Who was believed to have taken the caviar?

To see this, consider the source of (129):

(130) PRO believe [sCOMP WHO to have taken the caviar]

On the lower cycle, Wh Movement must take place and on the top cycle Passive must apply (in a new way) to give (129), in contradiction to the principle that is supposed to explain the ungrammaticality of (127) on the intended reading. The only alternative is to assume that Wh Movement can be bypassed on the lower cycle, and then to apply Passive and finally Wh Movement on the highest cycle. But this possibility of optional Wh Movement destroys the only remaining positive argument for successive-cyclic application (to which we return below).

The proposed successive-cyclic analysis (and any general principle of cyclicity it is supposed to follow from) requires ad hoc conditions on all rules like Tough Movement and Raising.20

Let us now ask what independent evidence is given to support the notion that Wh Movement takes place into COMP position (this idea is independent of the successive-cyclic nature of the rule). The first reason apparently comes from the putative Complementizer Substitution Universal of Baker (1970b) and Bresnan (1970):

---

20 This argument does not apply to a postcyclic rule of the sort that Chomsky considers later in the article. But a number of the arguments that have appeared in the literature apply to postcyclic formulations as well. For an evaluation of the Postal–Chomsky interchange on this matter, see Bach (1975).
REMARKS AND REPLIES

(131) Only languages with clause-initial COMP permit a COMP-substitution transformation. (Chomsky (1973, 234))

This principle was one of a number of hypotheses purporting to explain the correlation between interrogative movement and word order types first brought into prominence by Greenberg (1963). It appeared that no SOV language exhibited Wh Movement and that no SOV languages had clause-initial complementizers. The reason that the connection between Wh Movement and order type was explained by a principle like (131) was that in English and many other languages only one movement can take place per question. Chomsky (originally, we believe, in Chomsky (1964)) extended this observation to relative clause formation, since it is impossible in English to form a question from within a relative clause or to form a (restrictive) relative clause from a question. In the present context, this means that once a COMP position has been filled by one element it cannot be filled by another. This explanation works only if Wh Movement is a successive-cyclic, obligatory movement into COMP position. As we have just seen, there are English sentences in which in Chomsky’s framework of hypotheses the movement must be optional.

It is now known that a number of the factual claims underlying the above general hypotheses (and some others like them)21 are false. Frantz (1973) has reported on the Panoan language Sharanahua, which is a verb-final language and does not have clause-initial complementizers but which still has obligatory Wh Movement. Wachowicz (1973) has shown that several interrogative movements can take place in Polish in the same clause. Thus a good deal of the initial plausibility of the idea that movement to COMP position is essential to an analysis of Wh Movement drops away, at least insofar as that analysis appeals to principles of universal grammar.

We might still want to retain this principle in the description of English, even if it has no universal basis; however, we will now try to show that the one argument for such a proposal also fails. Recall that the assumption that Wh Movement takes place into COMP position is supposed to explain why we can have only one Wh Movement, since once the COMP node is filled no other movement can take place into that position (let us call this the “COMP-jamming” hypothesis). Then we would have an explanation for the ungrammaticality of examples like these (some must also involve Subjacency):

(132) *Who do you like the man who is living with?
(133) *The man who who likes is here.
(134) *Who what saw?

We will argue against this account on two grounds.

First, a wide variety of movement rules such as Topicalization (see Ross (1967)) obey the same constraints and escape hatches as Wh Movement:

---

21 In particular, the hypothesis of Bach (1971) that the Question Movement rule must be universally formulated with an initial verb as part of the structural condition.
Now, we can account for these examples in an entirely different way, in which case this new explanation would cover the *Wh Movement cases as well. Or we can extend Chomsky's analysis to say that these movement rules also involve movement to COMP position. But there is no other reason to think that these rules have anything to do with COMP. Notice that languages with no *Wh Movement, like Japanese, do have leftward movement rules of the Topicalization type. Japanese, of course, has clause-final complementizers.

Second, we observe the same patterns involving rules that cannot possibly be movement rules, for example, Comparative Deletion (Bresnan (1973 and 1975)):

139) She seemed as happy on Thursday as she seemed sad on Friday.
140) *When did she seem as happy on Thursday as she seemed sad?

In fact, it is not at all clear that movement of a WH-phrase has taken place in relative clauses with *that; there is even some evidence that it has not, that such clauses simply arise by deletion of the shared NP or pronoun in the relative clause.22 For example, it follows from the deletion analysis of *that-clauses (as it does not from the movement analysis) that no prepositions will appear at the head of *that-relatives:

141) *The man to that I said that is no novice.

*That-relatives, of course, act just like the genuine movement rules with respect to island constraints.

We have argued in this section, once again, that Chomsky's successive-cyclic account of unbounded movement rules is incorrect. If we have been successful, then we must conclude that the movement rules in question remain genuine counter-examples to a number of his constraints. We now take up the principle of cyclicity that Chomsky claims to require successive-cyclic rule application.

8. Strict Cyclicity

As noted, Chomsky states that the principle of strict cyclicity requires that *Wh Movement operate successive-cyclically.23 The principle is defined as follows (243):

22 This is not a new idea (see Emonds (1970), Morgan (1972)). Bresnan (forthcoming) has shown that so-called headless relatives involve simple deletion (and obey the constraints we have been looking at) and has argued (class lectures at the University of Massachusetts, Amherst) that *that-relatives also involve simple deletion. See also Grimshaw (1975a) on Middle English relatives (with some consequences for Modern English). Keyser (1975) argues for a uniform movement analysis as the best account for the history of relative clauses in English.

23 Chomsky writes simply "cyclic". We conclude that he means successive-cyclic from the actual analysis he proposes and from the fact that the arguments he alludes to (243, fn. 22) include arguments against successive-cyclic applications of the rule.
REMARKS AND REPLIES

(142) No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.

There seems to be a good deal of evidence for some principle limiting the applicability of rules within proper cyclic subdomains of a structure. For example, whenever we explain certain facts by rule ordering, we need to prevent the later application on a subsequent cycle of a rule that is ordered earlier than some other rule. For example, if we explain the failure of Passive to apply in sentences like (143) to give (144) (on the intended interpretation) by saying that (143) is derived from structures like (145) by a rule of Conjunct Movement (Lakoff and Peters (1969)), with Conjunct Movement ordered after Passive, we must prevent Passive from “going back down” on a later cycle to apply illegitimately:

(143) Mary married John.
(144) *John was married by Mary.
(145) Mary and John married.

But even if we do not hold to a theory of linear ordering of rules, we must prevent cyclic rules like Passive from applying after rules like Conjunction Reduction or Gapping:

(146) The Department of Public Safety arrested Mary, and the FBI, Bill.
(147) *Mary was arrested by the Department of Public Safety, and the FBI, Bill.

It is not altogether clear to us why Chomsky’s statement of Strict Cyclicity requires successive cyclic application of Wh Movement. Since he considers the rule to be a movement into COMP position, one could apply the rule across several cyclic boundaries at one “swoop” without violating his requirement (although there would be a violation of the Subjacency Condition). But let us suppose that his interpretation does require successive-cyclic application, which we have seen to be incorrect. A minor change in (142)—namely, changing affect to involve, would allow us to apply rules on the first cycle on which the structural description was met even if the rules themselves only affected a proper cyclic subdomain. This is essentially the idea behind Postal’s term higher-trigger.24

However, there remain a number of counterexamples to strict cyclicity, even in the weaker version followed here. We will discuss a number of these counterexamples, note that they follow a common pattern, and then suggest a further revision.

8.1. German Word Order

In Standard German the finite verb must stand at the end of subordinate clauses with

---

24 We are not putting forward this reformulation as a new idea. We believe it is just an attempt to make precise the notion of cyclicity that many linguists have implicitly followed (in fact, Wasow (1972b) formulates the notion in exactly the same way, attributing the formulation to Chomsky).
overt complementizers such as daβ, but if the daβ is deleted, the verb must stand in its “root” position as the second element (or first in yes–no questions):

(I 48) Er sagte, daβ er morgen komme.
‘He said that he was coming tomorrow.’

(I 49) Er sagte, er komme morgen.

In general, deletion of daβ depends on higher structure: the nature of the verb governing the daβ-clause, presence of negatives, position in the sentence, and so on:

(I 50) Ich glaube nicht, daβ er kommt.
‘I don’t think he’ll come.’

(I 51) *Ich glaube nicht, er komme morgen.

But now, no matter what we take to be underlying order of elements in German, we run into violations of strict cyclicity. Suppose we take the verb-final order to be basic and assume a verb-shift rule putting the verb into second position (or a sequence of rules having this effect). Then the derivation of (149) proceeds as follows:

(I 52) Underlying [ser sagte [sdaβ er morgen komme]]

So Cycle
Daβ-Deletion er sagte er komme morgen
Verb Shift er sagte er komme morgen

The last step cannot take place until we have deleted daβ, which takes place on the highest cycle. But the last step involves only elements within a proper cyclic sub-domain of the whole sentence. Similarly, if we take the SVO (or VSO) order as basic and let our rule putting the verb at the end be contingent on the presence of a complementizer, we must either delay application of the rule until we have had a chance to delete daβ on a higher cycle or reorder once more into “root” order, in either case violating strict cyclicity. Exactly parallel problems arise if we introduce complementizers transformationally.

8.2. Do Support

Wasow (1972b) has argued against the existence of a rule of Not Transportation, since it would entail operation of Do Support in a way that would violate strict cyclicity (as well as the Insertion Prohibition mentioned in section 6). One could also conclude that the strict cyclicity principle is incorrect. And in fact the problem is much more general. Any rule that has the effect of setting up conditions for Do Support to apply and which may be triggered by higher structure will entail a violation of strict cyclicity. Among such rules, for example, are VP Deletion and Comparative Deletion:

(I 53) I got up before you did.
(I 54) I caught more fish than Harry did.
8.3. Adjective Shift

J. Bresnan has brought to our attention the following situation. Consider these sentences:

(155) I don’t know a taller man than Greg.
(156) No less a person than Dick told me that.
(157) He is as terrible a linguist as he is a good cook.

It appears that there is a rule that repositions an adjective after the article just in case the -er has been affixed to the adjective and the remainder of the quantifier phrase is empty. Thus the rule applies in (155) but is blocked in (156) (Bresnan (1973, 1975)). For various reasons the comparative construction should remain intact until the operation of the comparative rules (i.e. the comparative rules are “higher-trigger” rules). But in (157) the Adjective Shift rule applies in violation of strict cyclicity, since the conditions for its application are not met until we have applied the comparative rules in a higher cyclic structure.

8.4. Clitic-pronoun Movements

If a rule of pronominalization exists in languages like French or Spanish with clitic movement rules, violations of strict cyclicity will occur whenever pronominalization depends on the presence of an antecedent in a higher structure:

(158) Jean a dit à Marie qu’il l’a vue à Paris.
‘John said to Mary that he had seen her in Paris.’

And in fact, any movement rules referring essentially to pronouns will present a problem for strict cyclicity (see Jacobson (forthcoming) on the necessity of deriving some pronouns by rule in English).

8.5. Adverb Placement

J. Bresnan has pointed out to us that possibilities for adverb placement in English can be affected by changes that happen necessarily on later cycles than the one defined by the sentence in which the adverb stands:

(159) Who did you say John believes fervently to be the right candidate?
(160) *John believes fervently Bill to be the right candidate.
(161) *John believes Bill fervently to be the right candidate.

All of the above situations, which can be easily multiplied, seem to entail violations of strict cyclicity. Various conclusions could be drawn: that the rules do not exist, that the problems should be handled by output conditions, that the rules are last cyclic or postcyclic. But common to all of them is the fact that the conditions for the operation of the rules in question change as the result of the presence of higher structure. In other words, there is “involvement” in each case between higher structure and the proper cyclic subdomains affect, but this involvement is indirect. We suggest
that the principle of strict cyclicity be replaced by something like the following principle:

(162) Rules must apply on the first cycle on which the conditions of their applicability are met.

This condition seems to us to exclude and allow just what we want. If we accept (162), we can formulate unbounded deletion and movement rules that are triggered by higher structure (as we can under the revised “involvement” formulation of strict cyclicity). The illegitimate operation of rules like Conjunct Movement and Passive is properly blocked, and all of the cases just considered fall out quite naturally.

9. Conclusions

We have examined a number of the conditions proposed by Chomsky: the Specified Subject Condition, Subjacency, the Subject Condition, the Tensed Sentence Condition, Chomsky’s formulation of Strict Cyclicity, and his decomposition of unbounded movement rules into successive-cyclic movements. We have seen that the effect of the first three conditions is to prohibit extraction from noun phrases while allowing (via COMP) extraction from complement sentences, and we have proposed a more general condition, the NP Constraint, which blocks the application of movement and deletion rules into NPs quite generally. One apparent class of counterexamples (a book about John) was shown to be structurally ambiguous. We saw that if we analyze such sequences as being derived from two sources (in appropriate contexts), the correct predictions about extraction and deletion follow from the NP Constraint, and we were able to adduce independent evidence for the two-structure analysis. In the case of the Tensed Sentence Condition, we saw that there is no evidence for the constraint in English or any other language and considerable evidence against it, and we were driven back to the conclusion that Chomsky’s motivation for the condition rests on a misanalysis of sentences with raising into object position. We took up as a separate issue Chomsky’s treatment of Wh Movement as a successive-cyclic operation and attempted to defend the more traditional view that such rules are in fact unbounded movement rules triggered by higher environments. Finally, we discussed some problems concerning Chomsky’s formulation of a principle of Strict Cyclicity and suggested, rather tentatively, a general condition on the applicability of cyclic rules that allows rules to apply in proper cyclic subdomains of higher structures just in case the conditions for their applicability have been changed as the result of higher structure.

The NP Constraint is a condition on applicability and not a constitutive constraint on the formulation of rules. There are many well-motivated rules that break up noun phrases: Raising rules, rules in various languages that float quantifiers from noun phrase, extraposition rules, and the like. Of the conditions proposed by Chomsky, the Subjacency Condition seems to be a real constraint of a constitutive sort. We argued against the constraint as a condition on applicability in section 3, but the facts remain that (1) there are many rules that raise or delete from complement sentences and (2)
such rules always obey the constraint that the controlling environment is just one clause “up”.\(^{25}\) It has been suggested elsewhere (Bach (1974a and b)) that this result follows from a more general constraint on grammar construction that says, in effect, that subsentences must be locally grammatical, in the sense that the subsentences formed at the end of each cycle must underlie possible grammatical sentences in their own right. It would follow from such a constraint that successive-cyclic applications of rules like \emph{Wh} Movement would be blocked. For example, in Chomsky’s derivation of a complex question like (163), there is an intermediate subsentence that can lead to no grammatical result, as we see in (164):

\begin{align*}
(163) & \text{Who did you say that Mary claimed had won the race?} \\
(164) & *\text{You said that Mary claimed who had won the race?}
\end{align*}

It would also follow that the only successive cyclic applications of rules would be those in which at each stage of the cycle the conditions for, say, a raising rule would be met, as in (165) (cf. (166)):

\begin{align*}
(165) & \text{This claim is believed to have been imagined to be false.} \\
(166) & *\text{This claim is believed to have been regretted to be false.}
\end{align*}

We believe, finally, that something along the lines of Postal’s rule typology (1974) will turn out to be correct. Rules can be divided into two basic types: those with a finitely specifiable structural condition, and those using essential variables. Of the first type, the rules can contain specifications of at most one level of embedding (Subjacency as a constitutive condition). Applicability of rules will be governed by principles such as the NP Constraint and others we have not discussed, such as the Coordinate Structure Constraint (Ross (1967)) or A-over-A Principle (Chomsky (1973)). We hope to have made some contribution in this article to the understanding of the general nature of syntactic systems.\(^{26}\)

References


\(^{25}\) An exception to this claim is the rule of Object Deletion (if it is one rule) discussed in Lasnik and Fiengo (1974). Somewhat closer to the truth would seem to be something like Postal’s rule typology, mentioned below, according to which we might divide rules into those that are finitely specifiable and those that are not, the former applying within a clause, or exactly one clause “down”. A rule can then be thought of as applying to “\emph{x}-decker” structures, where \(x\) is precisely 1, 2, or \(n\) and no further bound may be placed on the structure of a rule.

\(^{26}\) We are grateful to a number of individuals for comments and criticisms: Joan Bresnan, Barbara Hall Partee, and two anonymous reviewers for \emph{LI}, none of whom, of course, should be held to share any of our views or be responsible for any errors.
Bresnan, J. (forthcoming) "Headless Relatives."
Grimshaw, J. (1975b) Papers in the History and Structure of English, University of Massachusetts Occasional Papers in Linguistics 1, University of Massachusetts, Amherst.


Bach
*Department of Linguistics*
*University of Massachusetts*
*Amherst, Massachusetts 01002*

Horn
*Adam Mickiewicz University*
*Institute of English*
*Poznán, Poland 61–784*