
Collection Permanent Link: [http://hdl.handle.net/10822/707741](http://hdl.handle.net/10822/707741)

© 1976 MIT Press.

This material is made available online with the permission of the author, and in accordance with publisher policies. No further reproduction or distribution of this copy is permitted by electronic transmission or any other means.
Trace Theory and Twice-Moved NPs*

1. The trace theory of movement rules was first outlined in Chomsky (1973) and has been developed in many subsequent works: Selkirk (1972), Wasow (1972), Vergnaud (1974), Fiengo (1974), and Chomsky (1975; 1976). The last three references characterize trace theory as a proposal that an NP, when moved, leaves behind a “trace” that it binds. Traces are then subject to two conditions: that either they be properly bound or they be obliterated. An NP “properly binds” a trace when it precedes and commands it. Hence any lowering or rightward movement rule will place an NP in a position such that it cannot properly bind its trace; therefore, that trace will have to be obliterated in some way. This can be illustrated by passives in sentences and NPs. Agent Postposing will convert (1) into (2), where the traces are not properly bound by their NPs. Subsequent application of NP Preposing will yield (3), obliterating the traces left by Agent Postposing. Alternatively, a “spelling rule” might map (2b) into (4); there is no corresponding spelling rule applicable to (2a). Failure to obliterate the trace in (2) by either NP Preposing or a spelling rule gives an ungrammatical result, whereas the trace in (3) is in a position such that it is properly bound by its NP. This simple notion has many interesting implications.

\[
\begin{align*}
(1) & \quad \text{a. } [\text{the enemy destroyed the city}]_s \\
& \quad \text{b. } [\text{the enemy destruction the city}]_{NP} \\
(2) & \quad \text{a. } [t \text{ destroyed the city by the enemy}]_s \\
& \quad \text{b. } [t \text{ destruction the city by the enemy}]_{NP} \\
(3) & \quad \text{a. } [\text{the city destroyed } t \text{ by the enemy}]_s \\
& \quad \text{b. } [\text{the city destruction } t \text{ by the enemy}]_{NP} \\
(4) & \quad [\text{the destruction the city by the enemy}]_{NP}
\end{align*}
\]

* For useful comments on an earlier version of this article, I thank Mark Aronoff, Noam Chomsky, John Goldsmith, Norbert Hornstein, Elliott Macklovitch, and Tom Wasow. This work was conducted under the auspices of a Government of Quebec FCAC grant.

1 Throughout this article abstract structures will be given only in simplified form, omitting such details as tense markers, possessives, etc., and containing only information that is crucial to the discussion at hand.
The earliest motivation for introducing the notion of a trace was the desire to employ the Specified Subject Condition (henceforth SSC) to block the application of Each Movement in structures like (5a). It was argued that a moved NP, Mary, left behind a trace, which served as a specified subject and thereby activated the SSC. Chomsky (1973) also appeals to traces for the interpretation of wh structures.

(5) a. Mary seemed to each of the men [t to like the others]

b. *Mary seemed to the men to like each other.

But much of the subsequent appeal of the theory seems to lie in the claim that it yields exactly the right information to support semantic interpretation at the level of surface structure. In fact, the claim has been made (Dougherty (1975)) that the real motivation for trace theory is the desire to do all semantic interpretation off surface structures. Thus we may distinguish between two views of trace theory: the ‘‘pluralist’’ view mirrors the historical development of the theory and says that traces (a) play a crucial role in the syntax and (b) turn out to yield exactly the right information at surface structure to support semantic interpretation; the ‘‘exclusively semantic’’ view says that the early syntactic evidence for trace theory is not crucial and that the theory is motivated only by the requirement of surface structure semantic interpretation. Tied in with these two views of trace theory is the following question: does a moved NP always leave a trace or is a trace left only on the first movement, i.e. only in the original, deep structure position? It is sometimes assumed that a trace is left only on the first movement. If one adopts the exclusively semantic view of trace theory, there

---

2 If surface structure traces specify the deep structure position of NPs, then, for example, the thematic function of those NPs will be recoverable at the surface. Recall that in Jackendoff (1972) the role of deep structure in semantic interpretation is largely related to the role of thematic functions. The possibility of doing all interpretation at surface structure arises not from trace theory alone (because certain traces are obliterated in the course of derivations and thus some deep structure positions are not recoverable by traces), but from a conjunction of trace theory and a claim by John Goldsmith that whenever a trace is obliterated the thematic function of its NP turns out to be recoverable on independent grounds. So in (3) the thematic function of the enemy, Agent, is recoverable not by its trace, which has been obliterated by NP Preposing, but by the by phrase that characteristically marks Agents.

3 Insofar as the syntactic evidence was based on ad hoc ‘‘spelling rules’’ it looked quite weak. Arguably there is no need to appeal to spelling rules to explain the obligatoriness of NP Preposing in (2a). There seems to be a requirement that generally there must be an NP in sentence-subject position in surface structure. There is no such requirement for NPs and no corresponding phrase structure rule.

However, this is not to say that there is not syntactic evidence for trace theory (see note 12). Chomsky (1976) develops three independent lines of argument: (a) that it facilitates semantic interpretation, (b) that it permits a more restrictive typology of rules, and (c) that it allows a simplification of particular grammars (e.g. more ungrammatical sentences will be blocked by the SSC than would otherwise be the case, thereby avoiding the need for an ad hoc constraint to rule out sentences like (5b); Fiengo (1974) speculates that trace theory may eliminate the need for optional/obligatory conditions on syntactic rules (see p. 559). Others might subscribe to a version of the ‘‘exclusively semantic’’ view, claiming that while traces may not play any role in the workings of the syntactic component, they may affect the theory of syntax. For example, traces may facilitate semantic interpretation in such a way that one could eliminate a class of syntactic rules, permitting a more restrictive definition of transformational rules.
is no reason to have an NP leave a trace at intermediate stages of the derivation, because presumably such positions never play a role in semantic interpretation. I shall argue in this article that a trace must be left for every movement of a given NP, and hence that an NP may bind more than one trace (by transitivity). If this is correct, it will constitute support for the “pluralist” view of trace theory, claiming that it has a syntactic raison d’être as well as yielding good consequences for surface structure semantic interpretation.

Throughout the article I shall assume the correctness of trace theory and ask only whether, within that theory, an NP leaves a trace on each movement. My argument, which has two interesting consequences, will be based on a reworking of Bresnan’s (1971) analysis of *to* contraction.

2. G. Lakoff (1970) argued for a global rule to account for the contraction of *want to* → *wanna*, *used to* → *usta*, *supposed to* → *sposta*, *have to* → *hafta*, *got to* → *gotta*, *going to* → *gonna*, etc. His description used a notion of deletion sites. Larry Horn had pointed out the following minimal pair:

\[(6) \quad a. \text{ Teddy is the man I want to succeed.} \\
    b. \text{ Teddy is the man I wanna succeed.} \]

Here [(6a)] is ambiguous, and can be understood as either of the following:

\[(7) \quad a. \text{ I want Teddy to succeed.} \\
    b. \text{ I want to succeed Teddy.} \]

But [(6b)] can only be understood in the sense of [(7b)] since *want to* cannot contract to *wanna* if there is an intervening NP between *want* and *to* at an earlier point in the derivation. (Lakoff (1970, 632)).

This analysis is glaringly inadequate in that Lakoff ignores the fact that there was also an intervening NP in the (7b) “sense,” namely *I*, which is deleted by Equi NP Deletion.5

4 One might assume that under this view some metagrammatical convention will specify when a trace is to be left behind. For if transformations are purely “local”, no individual rule will “know” whether the NP it is moving is undergoing its first movement. Alternatively, traces may be treated in the syntax as nonterminal elements or coindexical markers; they will be spelled out as terminal elements only in logical form and then only in positions that are “semantically relevant”, i.e. in the original, deep structure position (Chomsky, class lectures). We shall return to this in section 6. Contrast this with the view taken by Fiengo (1974) and Chomsky (1975; 1976), who propose that movement rules be formulated so that *t* is inserted at the site from which an NP is moved, assuming, presumably, that *t* is left after every application of a movement rule (Fiengo (1974, 29)).

Even under the “exclusively semantic” view, traces might be required in intermediate positions. Lydia White has pointed out *John is believed to have cleverly been arrested*. *Cleverly* is a subject-oriented adverb in the sense of Jackendoff (1972) and cleverness is attributed to the cyclic subject *John*. If all semantic interpretation is to be done at the level of surface structure, it is not clear how the rules could identify *John* as the cyclic subject of *be arrested* unless a trace is left in that intermediate position.

5 This is also pointed out by Baker and Brame (1972). Lakoff (1972, 79) replies mystifyingly that “This is easily remedied by the addition of a condition limiting the rule to cases where the moved or deleted
Bresnan (1971) showed that the facts could be handled elegantly by a simple "local" transformation, which I shall call To Adjunction.\(^6\)

To Adjunction

\[
X \ V \ to \ V \ Y \quad \xrightarrow{\text{optional}} \quad X \ [V + to] \ V \ Y
\]

The rule adjoins the to to the preceding verb and then, as a "syntactic dependent", it is liable to phonological reduction. Bresnan argued that the rule must apply as the last of the known cyclic rules. So then, (8a) is ambiguous in that who may be interpreted as the underlying subject or object of succeed, whereas in (8b) who can only be interpreted as the object.

(8) a. Who do you want to succeed?
b. Who do you wanna succeed?

Bresnan's analysis, which she called the Ordering Hypothesis, captures this by postulating two deep structures (9a) and (9b), each of which may surface as (8a). On the other hand, (8b) can be derived only from (9b), i.e. the object interpretation for who.

(9) a. \([s + WH [s you want [s for [s who to succeed]]]]\)
b. \([s + WH [s you want [s for [s you to succeed who]]]]\)

(8a) is derived from (9a) by applying Complementizer Deletion; then, on the final cycle, who is moved into +WH. Only at this point is the structural description met for To Adjunction, though here it is blocked from applying by the principle of the cycle (Bresnan assumed S and \(\hat{S}\) to be cyclic nodes). However, in (9b) on the second S cycle, after Complementizer Deletion and Equi NP Deletion, the structural description for To Adjunction is met and the rule is free to apply to yield (8b) (with subsequent application of Wh Movement, Inversion, and Do Support on the final \(\hat{S}\) cycle). This analysis will also account for the unambiguous object interpretation for Teddy and who in (10), as compared to the ambiguous interpretations in (11).

(10) a. Teddy, I wanna succeed.
b. Teddy is the man (who) I wanna succeed.
c. The one (who) I wanna succeed is Teddy.

(11) a. Teddy, I want to succeed.
b. Teddy is the man (who) I want to succeed.
c. The one (who) I want to succeed is Teddy.

Bresnan's analysis makes two questionable assumptions: (a) There is a necessary constituent is a clause-mate of the auxiliary in question". I suppose that Lakoff assumes the existence of a rule of Subject Raising that will raise the subject of all infinitive clauses unless it has been deleted by an earlier application of Equi NP Deletion; this makes the required distinction.

\(^6\) I shall not follow Bresnan strictly, nor shall I argue for the very minor differences in formulation of her rule and deep structures.
ordering function that Equi NP Deletion precedes To Adjunction, which in turn precedes Wh Movement, Relativization (these may be the same rule), Topicalization, and Pseudocleft (arguably a special case of Relativization). The second relationship follows from the cyclic principle if one assumes (with Bresnan) a base rule $S \rightarrow \text{COMP}$ S, where COMP is the final resting place of wh elements and topoalized NPs; but the mechanism for achieving the ordering is irrelevant here, as we shall see. (b) There exists a transformational rule of Equi NP Deletion. The above analysis was incompatible with an interpretive version of Equi because at all stages of the syntactic derivation of, say, (8b) a $[\text{PRO}]_{\text{NP}}$ would intervene between want—to, presumably blocking application of To Adjunction. Furthermore, if the usual version of the SSC had been assumed, the PRO, controlled by the subject of want, would be defined as a specified subject, thereby activating the SSC and blocking To Adjunction. In section 5 I shall argue that the Equi hypothesis should be rejected and that an interpretive analysis should be amended so that the PRO will not “count” and will be ignored by the structural description of To Adjunction.

3. If we continue to assume, with Bresnan, that there is a transformational rule of To Adjunction, then clearly trace theory (which I assume to be motivated independently) will be applicable here. We shall see that the application yields some interesting conclusions. First, it removes the need for the first of Bresnan’s questionable assumptions. Trace theory requires that NPs moved by Wh Movement, Relativization, Topicalization, and Pseudocleft leave behind a trace. Hence (8a), (11a), (11b), and (11c) will each have two possible surface structures. (In this section I shall assume that wh elements do not move successive cyclically and therefore that $t$ is not also left in the COMP position; we shall return to this in section 6.)

(8)  a’. who do you want $[\text{PRO} \text{to succeed} \, t]$ (object interpretation)
     who do you want $[t \text{ to succeed}]$ (subject interpretation)

(11) a’. Teddy, I want $[\text{PRO} \text{to succeed} \, t]$
    Teddy, I want $[t \text{ to succeed}]

     b’. Teddy is the man (who) I want $[\text{PRO} \text{to succeed} \, t]$
    Teddy is the man (who) I want $[t \text{ to succeed}]

     c’. the one (who) I want $[\text{PRO} \text{to succeed} \, t]$ is Teddy
    the one (who) I want $[t \text{ to succeed}]$ is Teddy

Selkirk (1972, 121–129) also revised Bresnan’s treatment of To Adjunction by using a notion of “traces”. However, her analysis differs from the one adopted here in that the To Adjunction data were not used as crucial evidence to support her “traces convention”; rather, it was motivated on other grounds. It proved to be applicable to this area and to be an alternative to Bresnan’s description. Furthermore, Selkirk had a notion of “word boundary traces” that she claimed (1972, 73) to be quite different from “pronominal traces which are motivated by the syntax” (i.e. traces of the type used here and by Chomsky, Fiengo, and Wasow). The most important differences are that word boundary traces mark “removal sites” (i.e. gaps resulting from the application of either a deletion or movement rule), and that they were taken to be “invisible” to all transformations, to be unable to figure in the proper analysis of a string (1972, 128). I shall argue below that to handle the to contraction data it is crucial that the grammar distinguish the residue of a movement rule from what may be viewed as the residue of a deletion rule or a semantically interpreted but phonologically null element such as PRO, and that a trace be “visible”.

This content downloaded from 141.161.13.77 on Mon, 15 Apr 2013 11:54:50 AM
All use subject to JSTOR Terms and Conditions
Recall that Bresnan assumed a rule of Equi, so there would be no PRO in any of the above structures and To Adjunction would be unimpeded in the first of each of the above pairs. For the moment we simply assume that the intervening PRO on the object interpretations will not block To Adjunction, returning to this question in more detail in section 5. However, in the second of each of the above pairs the intervening trace will block application of To Adjunction—a correct result. Not only will it not be possible to factor the trees in such a way as to meet the structural description of To Adjunction, but recall also that in the earliest discussion of traces (Chomsky (1973)), they were defined as being specified subjects; they were abstract elements controlled by something other than the elements affected by the rule in question. Hence the SSC will block To Adjunction in the second case of each of the above pairs. Notice that Wh Movement in (8a') will not be blocked by the SSC, since it is a rule that moves material into a COMP slot, which is not a possible controller (Chomsky (1973)). However, the important thing to observe is that now the ordering is irrelevant, because To Adjunction will never be able to apply at any stage of the derivation, being blocked either by the intervening who, Teddy, etc., or (if it applies after Wh Movement, Topicalization, etc.) by the intervening trace. Hence, given trace theory, all the facts can be handled without recourse to ordering arguments.

Not only can the facts be handled without recourse to rule ordering; they must be. Assuming $\overset{\_}{S} \rightarrow \text{COMP S}$, then Wh Movement, Relativization, Topicalization, and Pseudocleft will each apply on the $\overset{\_}{S}$ cycle. Therefore, To Adjunction will be blocked from applying on the internal cycle by the intervening NP, which will still be present. Thus, a crucial test to choose between the ordering and trace accounts will be one in which no appeal can be made to this extra cycle.

The contraction of use(d) to $\rightarrow$ usta works just like that of want to $\rightarrow$ wanna, in that it will be blocked just in the event that an NP has been extracted (i.e. moved) from the used—to position. Hence contraction is possible in (12), where under either the transformational or interpretive Equi analysis no extraction is involved, but is not possible in (13), where an NP has been extracted.

(12) Tom used ____ to meet Harry for lunch.
Tom used ____ to be considered smart.
Tom used ____ to seem smart.

(13) $\overset{\_}{[s}\text{what }[s\text{Tom used ____ to take the picture was a Nikon}]$
this is the Nikon $[s\text{(which) }[s\text{Tom used ____ to take the picture}]$
$[s\text{what }[s\text{did Tom use ____ to take the picture?] ]$
$[s\text{the Nikon, }[s\text{Tom used ____ to take the picture}] ]$

In (13), Bresnan would argue that the NP is extracted only on the uppermost $\overset{\_}{S}$ cycle and is still present on the cycle where To Adjunction might apply, hence blocking the rule. However, consider (14a), where contraction is also impossible. Here a Nikon is moved out on the same cycle as that on which To Adjunction might apply.
(14)  a.  A Nikon was used ____ to take the picture.
     b.  *A Nikon was usta take the picture.
     c.  A used a Nikon to take the picture.
     d.  A Nikon was used t to take the picture.

If we adopt the trace theory, then either a Nikon (14c) or a trace (14d) will intervene between used — to and the structural description for To Adjunction will never be met; this is a good result, and again no ordering requirements are needed. But if we adopt the ordering hypothesis, we shall need a statement that To Adjunction must precede Passive (otherwise Adjunction might apply to (14a)). This would contradict Bresnan’s claim that To Adjunction applies as the last of the known cyclic rules and would yield an ordering paradox: the required ordering would be Equi – To Adjunction – Passive, which conflicts with a solid argument for ordering Passive before To Adjunction. We proceed now to that argument.

I shall assume that in infinitive constructions want and all verbs of desiring take a for complementizer. This is essentially the proposal of Bresnan (1972). It can be justified by noting that for may surface optionally with many desiderative verbs (15), and must surface when lexical material separates the verb from its complement (16), unlike epistemic verbs, which never tolerate for (17). For surfaces as well in corresponding derived nominals (18), but not in derived nominals related to epistemic verbs (19).

(15)  I would prefer (for) John to do it.
       I would hate (for) John to do it.

(16)  I want very much for John to come.
       *I want very much John to come.
       I would like very much for John to come.
       *I would like very much John to come.

* An anonymous reader for Linguistic Inquiry suggests a possible way for Bresnan to avoid the ordering paradox. (ii) and (iii) seem to indicate that the italicized portion of (i) is an island, and therefore adjunction of to used would be blocked by island constraints.

(i)  A Nikon was used to take a picture of the man.
(ii) *The man who a Nikon was used to take a picture of ____ was tall.
(iii) *Fred, a Nikon was used to take a picture of ____.

Such an analysis assumes the validity of “island constraints”, which I dispute. Following Chomsky’s reanalysis (1975) of the putative rules of Topicalization, Clefting, Tough Movement, etc., as special cases of Wh Movement, then the usual “island” phenomena (cf. Ross (1967)) can be subsumed under the Subjacency Condition (see Chomsky (forthcoming) for details; for Subjacency, see Chomsky (1973, 196)). Application of To Adjunction in a structure corresponding to (i) would not be blocked by Subjacency. This analysis is controversial, but attractive from many points of view (see section 6 for more discussion). However, even assuming “islands”, (i) and (ii) indicate only that the complex NP a picture of the man is an island, as argued by Horn (1975); there is no reason to assume that to take is part of the island.

* The difference is that Bresnan claims that epistemic verbs take no complementizer, whereas I follow a slight revision made by Chomsky (class lectures, 1974) and claim that they have an empty complementizer. This is necessary in order to be consistent with the proposal, to be discussed later, that COMP provides an “escape hatch” for wh elements. The empty COMP will be a preterminal element and therefore “invisible” to rules such as Passive (or Adjective Formation; see note 11).
(17) *I believe (very much/sincerely) for John to be popular.
   *I expected (very much/sincerely) for John to have been elected.

(18) her desire for Mary to come [cf. her desire for fame]
    her preference for gin to be available [cf. her preference for gin]

(19) *her belief for John to be popular [cf. *her belief for God]
    *her expectation for John to have been elected [cf. *her expectation for John]

This *for complementizer, so justified, will intervene between want and to and therefore will have to be deleted before To Adjunction can apply to yield wanna. Hence To Adjunction must be ordered after Complementizer Deletion. But Passive must be ordered before Complementizer Deletion, in order to block (20). If Passive precedes Complementizer Deletion, then it will fail to apply in (20), because its structural description will not be met. Passive requires an NP V NP sequence, and so is blocked in the structures underlying (20) by the *for intervening between the V and the following NP—a correct result. Compare (21), where there is no *for at any stage of the derivation to block application of Passive.10

(20) *John was
    \[
    \begin{align*}
    \text{wanted} & \\
    \text{preferred} & \\
    \text{desired} & \\
    \text{hated} & \\
    \end{align*}
    \]
    to do it.

(21) John was
    \[
    \begin{align*}
    \text{believed} & \\
    \text{expected} & \\
    \text{known} & \\
    \end{align*}
    \]
    to be popular.

In this paragraph we have demonstrated by an indirect argument (via the interaction with Complementizer Deletion) that Passive must precede To Adjunction.11 Hence the ordering paradox.

This ordering paradox, of course, does not arise in a trace theory account of the contraction data along the lines of (8a'), (11a'), (11b'), and (11c'). We showed that if one adopts that account, then no ordering relationship needs to be specified. That alone

10 See Lightfoot (1976a) for more details of this analysis, which dispenses with the usual Raising rule.
11 This argument assumes (counterfactually, but in agreement with Bresnan (1971)) that there is a rule of Passive. If, instead, passive sentences are generated by rules of Agent Postposing and NP Preposing, then we might establish the ordering paradox by arguing that NP Preposing must precede Complementizer Deletion and therefore To Adjunction. Chomsky (personal communication) points out that this will follow only if either (a) NP Preposing applies only to immediately postverbal NPs, or (b) Bresnan's Fixed Subject Constraint is correct, i.e. the *for blocks the preposing rule. (a) is probably wrong if the rule is to be generalized to other construction types (see note 13), and Chomsky argues that also (b) cannot be right; in which case the constraint that in passives the moved NP must originate in postverbal position could be expressed by an Adjective Formation rule, that -en can make an adjective out of V + t. Thus any effort to "passivize" after *for will yield a structure from which no adjective can be formed, V for t. If that analysis is correct, then Adjective Formation would have to precede Complementizer Deletion, and NP Preposing, of course, would precede Adjective Formation: the required ordering would be NP Preposing – Adjective Formation – Complementizer Deletion – To Adjunction. A non-trace account of (14a), on the other hand, would require To Adjunction to precede NP Preposing. Hence there will be an ordering paradox whether one thinks in terms of a Passive rule or in terms of a rule of NP Preposing.
is a reasonable argument for using trace theory, assuming trace theory to have independent motivation, since it yields a simpler grammar. The argument gains force when it is shown that trace theory provides a way to avoid what would otherwise be an ordering paradox. We have therefore shown that trace theory gives a superior account of the contraction data, providing further syntactic motivation for trace theory and lending support to the "pluralist" view over the "exclusively semantic".12

4. After this demonstration, we are in a position to ask what happens to NPs that undergo two successive movements. It emerges that precisely parallel facts obtain. (22) and (23) are ambiguous between the subject or object reading for the moved NP. But if contraction to wanna takes place, who and what are unambiguously objects.

(22) Who do you want to be certain to succeed?
   a. [who do you want [t to be certain [t to succeed]]]
   b. [who do you want [PRO to be certain [PRO to succeed t]]]

(23) What do you want to melt?
   a. what do you want [t to melt t]
   b. what do you want [PRO to melt t]

I assume that who and what have each undergone two successive movements in (22a) and (23a). Thus who originates as the deep structure subject of succeed, moves by Subject-to-Subject Raising to become the subject of be certain to succeed, and then moves again by Wh Movement. Similarly, in (23a) the deep structure subject of melt is an indefinite NP and what is the object: [Δ to melt what].13 What moves into subject

12 Another piece of syntactic evidence is that trace theory automatically blocks the reapplication of NP Preposing in passive constructions, something that in other theories might require an ad hoc constraint. Given a deep structure (i), the application of Dative Movement, Agent Postposing and Deletion, and NP Preposing will yield (ii) in a non-trace analysis and (iii) under trace theory. In (ii) the structural description is still met for another application of the NP Preposing component of passives, which would yield garbage (I ignore the morphology of the passive verb). In (iii) the trace intervening between the V and following NP blocks application, and so the problem does not arise.

(i) Δ gave the book to Mary.
   (ii) Mary was given the book.
   (iii) Mary was given t the book.

These facts might also be explained by a version of the structure-preserving hypothesis in conjunction with the recoverability-of-deletion principle, but this would not suffice to block the book was given being derived from Δ gave the book to Δ; Dative Movement and then NP Preposing would yield Δ was given the book, and now nothing would block a further application of NP Preposing to produce the book was given.

13 Under this analysis (Chomsky (class lectures 1974)), if the subject NP is lexically filled in deep structure, it will receive an agentive interpretation, yielding a "causative" sense: Peter melted the ice. Inchoatives, such as The ice melted, are transformationally derived from a structure in which the subject NP is empty; hence Δ melted the ice → the ice melted t. This derivation is effected by a rule of NP Preposing, a rule of great generality.

NP Preposing
X NP Y NP Z 1 4 3 t 5

This rule is also involved in the derivation of structures corresponding to passive sentences (the ice was
position by a rule of Middle, giving [what to melt t], and then to the front of the sentence by Wh Movement. Now, if NPs leave traces on every movement, then the surface structures will be as in (22a) and (23a), and Adjunction cannot apply. But if a trace is left only on the first movement, then the surface structures will be (22c) and (23c).

(22) c. who do you want [to be certain [t to succeed]]
(23) c. what do you want [to melt t]\(^{14}\)

There is nothing to prevent To Adjunction from applying to (22c) and (23c), yielding incorrect results (i.e. correlating contraction with the subject readings). Some independent means will have to be found to block Adjunction in these cases. But if the surface structures are as in the (a) examples, then Adjunction is blocked by the intervening trace, as in (8a'), (11a'), (11b'), and (11c'). In the absence of any independent means to block Adjunction in (22c) and (23c), we have an argument that NPs leave traces on their second movement.\(^{15}\) Q.E.D.

The same facts obtain for NPs whose second movement results from (a) Relativization, (b) Topicalization, and (c) Pseudocleft. In (24) and (25), contraction is possible only where the moved NP is interpreted as the deep object of the lowest verb.

(24) a. Teddy is the man (who) I want to be certain to succeed.
   b. Teddy, I want to be certain to succeed.
   c. The one (who) I want to be certain to succeed is Teddy.

(25) a. The big popsicle is the one (which) I want to melt.
   b. The big popsicle, I want to melt.
   c. What I want to melt is the big popsicle.

As further confirmation of this analysis, notice passive sentences like (26) and (27), where only (26) will allow contraction.

(26) Who do you want to be arrested by?
   a. who\(_t\) do you\(_t\) want [PRO\(_t\) to be arrested \(t_j\) by \(t_i\)]
   b. you want [who to arrest PRO]

(27) Who do you want to be arrested?
   a. who\(_t\) do you want \([t_i] to be arrested t_j\]
   b. you want [\(\Delta\) to arrest who]

\(^{14}\) Another reason to rule out the possibility of such surface structures is that they would not be semantically interpretable in any clear way. In particular, it is not clear how one could specify any kind of subject for be certain to succeed and melt.

\(^{15}\) A further argument for the same conclusion can probably be constructed from a sentence pointed out by Howard Lasnik, the books seem to the men [t to have been given t to each other]. The books is moved first to the subject position of to have been given; but a trace must be left in that position on the second movement in order to block Reciprocal Interpretation (or Each Movement) relating the men to each other.
If the deep structures are as in (b), then trace theory will specify (a) as the surface structures. In (26), for example, who begins as the deep subject of arrest and is moved by Agent Postposing: \[t \text{ to be arrested } \text{PRO} \text{ by who}.\] NP Preposing then gives \[\text{PRO} \text{ to be arrested } t \text{ by who}.\] Wh Movement on a later cycle gives (26a), where the leftmost trace is bound by PRO and the rightmost trace by who. No trace intervenes between want and to and so Adjunction is free to apply. In (27), Agent Postposing and subsequent Deletion yield \[t \text{ to be arrested who}.\] NP Preposing then produces \[\text{who to be arrested } t\], obscuring the trace left by the original subject. If a trace is left on the second movement, as we have argued here, Wh Movement (and Inversion and Do Support) then give (27a). Here a trace intervenes between want and to, hence blocking the application of Adjunction. Identical facts obtain when the second movement is by Relativization, Topicalization, and Pseudocleft, where again contraction is possible only if by is present, i.e. where no trace intervenes between want and to.16

(28) a. Teddy is the one (who) I want to be arrested (by).
   b. Teddy, I want to be arrested (by).
   c. The one (who) I want to be arrested (by) is Teddy.

Here we have an argument that a trace is left whenever an NP is moved, and we have shown, therefore, that trace theory has consequences for syntax and is not motivated solely by semantic considerations.17 It is worth calling attention to two points. First, it is usually claimed (e.g. by Chomsky (1976, ch. 3)) that it is a happy consequence that a view of syntax incorporating trace theory provides exactly the right amount of information at surface structure to support semantic interpretation. It is by no means self-evident that traces marking the intermediate positions of NPs are at all relevant to semantic interpretation, and it is therefore possible that they are semantically redundant. This seems to me to be an open question. Second, for those who are (rightly) disturbed by the proliferating types of dummy symbols for NPs that are lexically empty at surface level, we have here an argument that there is a crucial distinction between a trace, which marks an “extraction site”, and a PRO, which marks a coreferential NP position where no extraction has been involved. Traces, but not PROs, block Adjunction. This indicates that PROs do not “count” when trees are analyzed to match structural descriptions of transformations, as we shall discuss in section 5. Under this account we do not have to appeal to the SSC to block Adjunction. It will be blocked by “visible” material intervening between want and to.

---

16 Potentially there should also be an argument based on using Tough Movement for the first movement, but sentences such as Who do you want to be sure to please? seem bizarre on both the subject and object interpretations and do not give clear facts. Similarly bizarre are corresponding sentences that have not undergone Tough Movement: Who do you want it to be easy to please? Compare Who do you want to be sure to please?, in which the wanna contraction correlates with interpreting who only as the object of please, not as the subject.

17 It is tempting to take a stronger position and assert that no linguistic entity may be postulated that does not play some role in the syntactic, semantic, and phonological components. At this stage such a position would only reflect a theoretical prejudice, albeit a healthy one.
either lexical NPs or traces; PROs will be "invisible". This requires further examination.

5. While PROs may not affect the structural requirements of transformational rules, it has been argued that they may affect the application of rules in another way, and this should block To Adjunction. The SSC (Chomsky (1973, 254)) specifies that an improperly controlled PRO can block the application of rules.

 Specifications Subject Condition
 No rule can involve X, Y (X superior to Y) in the structure
 where Z is the specified subject of WYV.

 A specified subject is one that is either (a) lexically filled (see Chomsky (1973, note 39)) or (b) controlled by a category not containing X (p. 262). We shall refer to (a) as the syntactic definition; we shall call (b) semantic, since it refers to the property of "control". This constraint predicts that Adjunction will be blocked, say, (23b).

 (23) b. what do you want [PRO to melt \]

The rule affects want (="X" of the definition) and to (="Y"); PRO is controlled not by want but by the subject you, and therefore is a specified subject. Therefore, this analysis is an exception to the SSC, since Adjunction can in fact apply here. This leads me to suggest a revision to the SSC.

The SSC is a constraint on syntactic (and other) rules and is sensitive to control properties, which in turn are determined in the semantic component. This raises a question about the hypothesis of the autonomy of syntax. It is consistent, of course, with the notion of a parameterized autonomy thesis (Chomsky (1974)), according to which there are certain prescribed areas where syntactic and semantic information may be confounded, such as the lexicon or conditions on rules. It is also consistent with the notion that conditions on rules may be part of the theory of grammar, and not part of specific grammars. Nonetheless, we could construct a plausibility argument for some version of the autonomy thesis, if we could show that syntactic rules are subject only to the syntactic aspect of the SSC and that only semantic rules are sensitive to the notion of control. To do this, we would need to show that syntactic rules may be blocked by an intervening lexically specified subject but not by an intervening PRO controlled by something other than "X" in the definition. It seems to me that this is a plausible position and, if sustained, it would allow To Adjunction to take place in (23b) and all other structures in which only PRO intervenes between the verb and its potential clitic to. In fact, under this proposal we may dispense with the symbol PRO, dealing instead with an empty node that may be "filled" by the rules of semantic interpretation. Assuming that empty nodes are "invisible" to the structural descriptions of transformations, then nothing will intervene between want and to, and To Adjunction will not be blocked. Alternatively, we might regard PRO as a preterminal element, as Chomsky (class lectures, 1975) treats trace (see section 6).

This content downloaded from 141.161.13.77 on Mon, 15 Apr 2013 11:54:50 AM
All use subject to JSTOR Terms and Conditions
This position derives its plausibility from the fact that several syntactic rules are not sensitive to the controlled PRO aspect of the definition of a specified subject, and I know of no clear case of a syntactic rule that is. Relevant cases will be syntactic rules that operate across S boundaries and “dip down” into infinitival complements. To Adjunction is an example of such a rule, and it is blocked only by intervening lexical material or a trace and not by an improperly controlled PRO. The same situation obtains for Tough Movement, as shown by Chomsky himself. Chomsky (1973) notes the ungrammaticality of (29), where Tough Movement has taken place over a lexically specified subject, and argues that (30a) and (30b) have underlying structures of the form of (30c,d), where for us and for the rich are part of the matrix sentence and control the lower PRO. The structures of (30c,d) allow Tough Movement although the PRO subject is controlled by something other than it (which is “X” in the definition) and therefore counts as a specified subject. Chomsky (1973, 263ff) suggests various more or less ad hoc ways around this problem, “reaching no firm conclusion”.

(29)  a. *Latin is a waste of time for us for them to teach us.
    b. *The hard work is pleasant for the rich for poor immigrants to do.

(30)  a. Latin is \{easy a waste of time\} for us to learn.
    b. The hard work is pleasant for the rich to do.
    c. It is \{easy a waste of time\} for us [PRO to learn Latin]
    d. it is pleasant for the rich [PRO to do the hard work]

Wh Movement is a third rule that can dip down into an infinitival complement, but it obeys neither component of the SSC in sentences and can move elements both over a lexically specified subject (31) and over an improperly controlled PRO (32). However, it is subject to the syntactic version of the SSC in NPs (33a,b); the semantic version will never be applicable since PROs do not occur as the “subject” of NPs (or rather, if they do, they cannot be interpreted by the usual coreference rules, and the uninterpreted PRO (or, under our proposal above, the unfilled node) ensures that the derivation is blocked).

(31) who do you want [John to visit ____]?
(32)  a. what is easy for us [PRO to learn ____]?
    b. who did you persuade Bill [PRO to see ____]?

(33)  a. who did you see [pictures of ____]?
    b. *who did you see [John’s pictures of ____]?

Chomsky accounts for these facts by moving wh elements into COMP on the

---

18 I ignore here the irrelevant question of whether this is better formulated as a deletion rule, as argued by Lasnik and Fieno (1974). If one assumes the trace theory of movement rules (as they did not), Jackendoff’s (1975) refutation succeeds with pleasant consequences and avoids the need for end-of-cycle interpretation. In any case, more recent work by Chomsky shows that the data handled by a rule of Tough Movement/Object Deletion can be subsumed under Wh Movement, as we shall discuss in a moment.
innermost cycle and then specifying that COMP acts as an "escape hatch", whereby wh may leave its S if it moves directly to another COMP. In this way Wh Movement is rendered subject to the SSC and we have an explanation for (33). Of course, NPs have no COMP and therefore no escape hatch; hence, if we assume the Subjacency Condition, extraction of wh elements from within an NP embedded in an NP will be impossible. A possible argument that Wh Movement is sensitive to an improperly controlled PRO lies in Chomsky's recent explanation for *What sonata is this violin easy to play on by his generalized Wh Movement rule (see section 6). The structure after the first application of Wh Movement will be COMP this violin is easy [which for PRO to play what sonata on t]; what sonata now cannot move to the highest COMP, it might be claimed, by virtue of the SSC and the requirement that COMP cannot be doubly filled. But this is false since although PRO is not controlled by COMP, i.e. "X" for the purposes of the SSC, COMP is not a possible controller and so PRO will not count as a specified subject. However, such movement would also be blocked by Subjacency on the crucial assumption that either S or AdjP is a cyclic node (for the question of the cyclicity of S, see section 6), and therefore the SSC is not crucially involved. In the example cited, PRO is interpreted as indefinite, but it would be controlled in the parallel example *what sonata is this violin easy for you [for PRO to play on]. Notice that the (strict) cyclic principle will block successive cyclic movement of what sonata after deletion of which on the higher cycle. So we have a solution to the old violin–sonata problem, but we are still maintaining our claim that Wh Movement is not sensitive to an inappropriately controlled PRO.

In fact, the only syntactic rule that seems to be subject to the semantic aspect of the SSC is Each Movement, and this is by no means a clear case of a syntactic rule. If we claim that each is generated in its surface position and is subject to an interpretive rule to determine the NP it binds (the position advocated by Jackendoff (1972), although not discussed or argued for in detail there), then we lose the earliest motivation for making syntactic rules sensitive to improperly controlled PROs. This seems to me to be a reasonable view, if only because of the high cost of a syntactic Each Movement rule in terms of special requirements needed for Chomsky's (1973) conditions on rules.

I know of no syntactic rules that are sensitive to the semantic definition of a specified subject, while three are sensitive to lexically filled subjects, i.e. the syntactic aspect of the definition: To Adjunction, Tough Movement, and Wh Movement. Hence I propose a revision of the SSC: that syntactic rules are sensitive only to the syntactic definition of a specified subject. This (a) solves our problem with to contraction in that Adjunction is no longer blocked by the SSC from applying to (23b) and our analysis becomes compatible with an interpretive treatment of Equi, and (b) yields another plausibility argument for some version of the autonomy thesis. Hence a syntactic rule will be blocked from dipping down into a tenseless structure (S or NP) only if there is lexical material (which includes t, which I take as a terminal element, but not PRO) in

This content downloaded from 141.161.13.77 on Mon, 15 Apr 2013 11:54:50 AM
All use subject to JSTOR Terms and Conditions
the subject position. Note in particular that Wh Movement behaves in accordance with this proposal. This takes on importance in the light of recent work by Chomsky, following up suggestions in Chomsky (1975). He argues that two rules, NP Preposing and Wh Movement, constitute the core of English syntax and that each has much “broader scope than has hitherto been imagined”. “Rules” such as Comparative Deletion (Bresnan 1973), Tough Movement, Infinitival Relative Formation, Topicalization, and Clefting are special cases of Wh Movement. Having so much subsumed under Wh Movement increases the plausibility of our proposal.

An alternative approach is to argue for a deletion rule erasing the subject of the complement of such verbs as want. Chomsky (1975) adopts this approach, while maintaining a PRO analysis for verbs such as promise and persuade; he calls this the Equi Hypothesis. Thus I promised Bill to come home, I persuaded Bill to come home, and I wanted to come home will have the deep structures shown in (34).

\[
\begin{align*}
(34) & \quad a. \text{ I promised Bill [PRO to come home]} \\
& \quad b. \text{ I persuaded Bill [PRO to come home]} \\
& \quad c. \text{ I wanted [for X-self to come home]}
\end{align*}
\]

The Equi Hypothesis will handle the contraction data in that after deletion of the designated element for X-self, nothing will intervene between want and to, and To Adjunction will be free to apply; also, there will be no inappropriately controlled PRO to activate the standard version of the SSC. Two independent arguments are adduced for this position. The first concerns clitics in Portuguese and Czech (based on work by Quicoli and Toman, respectively), where a pronominal object of an embedded verb may be cliticized to the matrix verb.

\[
(35) \quad \text{O medico nos quer examinar.}
\]

\[
\text{the doctor us \_wants [\_ to examine t]}
\]

However, such cliticization will be blocked if the embedded subject is lexically specified, a trace, or an inappropriately controlled PRO.

\[
\begin{align*}
(36) & \quad a. \text{ *Carlos us$_t$ saw [the doctor examine t$_i$]} \\
& \quad b. \text{ *the man [who$_i$ Maria us$_j$ saw [t$_t$ examine t$_j$]] disappeared} \\
& \quad c. \text{ *the doctor, us$_i$ promised [PRO$_j$ to examine t$_i$]} \\
& \text{ *the doctor it$_t$ persuaded us$_j$ [PRO$_j$ to examine t$_i$]}
\end{align*}
\]

Cliticization to the matrix verb from the embedded object is possible only with verbs such as querer ‘want’. Postulating a deleted subject here will make the appropriate distinctions. If we assume that the grammar of a given language L may be underdetermined by data from L, we may accept an argument of this form, i.e. that given clear evidence from Portuguese, the conclusion may carry over to English on some currently ill-understood general grounds. However, clitics in Old French behave in such a way
as to cast strong doubts: here sentences like (35) and (36c) are grammatical. In other words, an inappropriately controlled PRO will not block cliticization, and therefore postulating a deletion rule for want, etc., will not make the required distinctions. I cite the Old French, the modern equivalents, and a gloss.19

(37)  
   a. Deus le me doinst venger (Chanson de Roland)  
       Dieu me donne de le venger  
       ‘may God allow me [PRO to avenge him]’  
   b. qu’i la viegne chacier  
       qu’il y vienne la chasser  
       ‘that he come [PRO to chase her]’

This suggests that it can scarcely be argued that the Equi Hypothesis is motivated for English on universal grounds; it will have to be motivated on internal grounds.

A second argument is based on English. Discussing problems in determining the scope of only, Fodor (1975) argues against the proposals of McCawley (1970) and for the Equi Hypothesis on the basis of (38).

(38) Only Churchill remembers giving the speech about blood, sweat, and tears. The descriptive problem is to characterize exactly what Churchill is the only person to remember. Clearly Churchill is not the only person to remember Churchill giving the speech, since millions of others remember the same event. Rather, Churchill is the only one to remember that he himself gave the speech. This seems to be an argument against deleting Churchill, but it does not choose between the Equi Hypothesis above and the PRO analysis. Since Chomsky allows rules of interpretation to operate on partially determined logical forms (LF), the scope of only could be prescribed in LF: specifying that of the people remembering PRO giving the speech, Churchill was the only one. In fact, the data seem to be the same with promise: Only Churchill promised to give the

19 I assume that (37a,b) involve clear cases of a PRO and that an Equi analysis would not be applicable, but that assumption would need to be supported by a detailed analysis of Old French. Clitics in Italian seem to have some of the properties of those of Old French; again it appears that they can be promoted over an inappropriately controlled PRO (Radford (1976)).

(i) Lo insegn a fare a Paolo.  
    it he-teaches to do to Paolo  
    ‘he teaches Paolo [PRO to do it]’

(ii) Ti vengo a trovare.  
     you I-come to find  
     ‘I am coming [PRO to find you]’

There is much more to be said about the Italian data, since it seems possible also to promote a clitic over a lexically specified subject.

(iii) Ci ho visto Paolo entrare.  
     there I-have seen Paolo enter  
     ‘I have seen [Paolo enter there]’

Again, detailed analysis is called for, but the differences between the clitics of Portuguese and those of Old French and Italian should lead to caution in generalizing the Portuguese data as support for an analysis of parallel verbs in English. (The Old French data are from Morin (1975), who uses them to make different claims.)
speech about blood, sweat, and tears. If one assumes promise to be a clear case of a PRO verb, the argument for assigning remember an Equi analysis fails.

Having concluded that the independent motivation for the Equi Hypothesis is weak, we must choose between it and the PRO-doesn’t-count analysis. Both accounts handle the To Adjunction data. The PRO analysis, if it can be sustained, must be preferable: it permits a more restrictive version of the autonomy thesis (and provides another plausibility argument for it), while the Equi Hypothesis adds another phonologically null element to the proliferating inventory. The PRO analysis reflects a somewhat different view of trace theory: that the “residue” of movement rules has some kind of privileged status, as indicated by the contraction data, and in a sense the moved element is always present in its original position, in contrast to the residue of what have been viewed as deletion rules or interpretive rules specifying coreference for phonologically null elements. On the other hand, Chomsky views a trace and a PRO as essentially the same thing. He seems to regard a trace as a function, a co-indexing device: when NP_ moves, it leaves a trace t(NP_). We may then view PRO as a base-generated t(x), where x is a variable to be filled in by a control rule, and we reduce the proliferation of null elements. The contraction data pose a problem for this view and seem to necessitate adopting the Equi Hypothesis, thereby adding another type of null element.

6. The analysis adopted in sections 3 and 4 has a further consequence in that it is incompatible as it stands with what Postal (1972) calls the “successive cyclic treatment” of Wh Movement, advocated by Chomsky (1973). One may derive Who do you want to try to visit either by a single application of Wh Movement, as in (39), or by passing the wh element successively into each COMP node, as in (40).

(39) COMP you want [COMP PRO to try [COMP PRO to visit who]]

(40) COMP you want [COMP PRO to try [COMP PRO to visit who]]

For example, consider one reading of (22), which I repeat for convenience.

(22) Who do you want to be certain to succeed?
   (b) who do you want [PRO to be certain [PRO to succeed t]]

If one adopts the successive cyclic hypothesis, then the structure at the end of the second S cycle will be (22d).

(22) (d) you want [[who]COMP PRO to be certain [[t]COMP PRO to succeed t]_s]_s

Who will not be moved out of this position until the final S cycle and therefore would serve to block application of To Adjunction on the last S cycle, the only possible point of application—an incorrect result. Even if application of To Adjunction can somehow
be delayed until after who is moved out of this position, the rule will still be blocked, this time by the intervening trace. On the other hand, if one assumes the single movement hypothesis illustrated in (39), who will at no point intervene between want and (PRO) to, and therefore Adjunction will be free to apply—the correct result. Throughout the analysis of section 4 we assumed the single movement hypothesis, and that assumption was crucial. The reader may verify this in the case of (23b), (24a,c), (25a,c), (26), and (28a,c). Under the successive cyclic hypothesis, a wh element will intervene between want and (PRO) to, thus blocking To Adjunction in all cases. Notice that this problem arises under both the Equi and PRO analyses discussed in section 5.

This corroborates the results of Jenkins (1975). Chomsky (1973) gives an argument that Wh Movement should be postcyclic (based on its interaction with Each Movement) but shows that the force of this argument is avoided under trace theory. Jenkins argues to the same effect and shows that his argument (based on the elimination of the condition (123.c), that anything moved from a COMP must be moved only into another COMP position) cannot be avoided by appeal to traces. Hence, he argues, the postcyclic treatment of Wh Movement must be correct, since it gives the correct result in both cases. However, adopting a postcyclic Wh Movement also entails violations of the Subjacency Condition and the SSC. For example, a single, postcyclic application of Wh Movement in (41a) will cause the wh element to cross two cyclic nodes; the movement in (41b) contravenes the SSC. But if we make Wh Movement an exception to these constraints, we are left with no way to characterize the ungrammaticality of (41c).

\[
\text{(41) a. who do you want [PRO to hear [stories about t]_{NP}]s}\\
\text{b. who do you think [John saw t]}\\
\text{c. *who do you want [PRO to hear [John's stories about t]_{NP}]s}\\
\]

Making Wh Movement postcyclic and marking it as exceptional to Subjacency and the SSC is a high price to pay. Only if the rule is successive cyclic and subject to Subjacency can Chomsky’s proposals for subsuming so much under it stand, and successive cyclicity is essential if Ross’s Complex NP Constraint, Coordinate Structure Constraint, and Wh Island Constraints are to be just special cases of Subjacency. To illustrate, consider infinitival relatives such as (42a), which Chomsky proposes to derive by application of Wh Movement with subsequent deletion of the wh element, as in (42b).

\[
\text{(42) a. He is looking for a man to build a pyramid.}\\
\text{b. he is looking for a man [\{wh\}_{COMP} t to build a pyramid]}\\
\]

\[
\phi
\]
(43)  a. *What are you looking for a man to build?  
    b. [\[wh_i\]_{COMP} you are looking for a man to build \[wh_j\]_{S}]

If Wh Movement is successive cyclic and subject to Subjacency, (43a) will be blocked: \[wh_j\] cannot move to the highest COMP in (43b) without violating Subjacency (assuming \[\hat{S}\] and S to be cyclic nodes; if S is not cyclic, the SSC will block the movement). If Wh Movement is to be postcyclic, it will clearly not be subject to Subjacency or the SSC and therefore the movement of \[wh_j\] will have to be blocked by a not otherwise needed wh island constraint. Such illustrations can be multiplied by examining the so-called rules of Comparative Deletion, Topicalization, Relativization, Clefting, etc., and showing how the grammar would have to include not only the Subjacency Condition, but also Ross’s constraints above, stated separately—an undesirable complication.

Some alternatives suggest themselves, although none is wholly satisfactory:

(i) Under the Equi Hypothesis, let the designated element to be deleted include the \(t\) in COMP, \([\{t\} \text{ for } X\text{-self}]\), an effective but surely ad hoc move.

(ii) Postulate that Wh Movement, unlike NP Preposing as discussed in section 4, leaves behind a trace only on the first application of the rule. Again, stated in this form, this is quite ad hoc. However, Chomsky (class lectures, 1975) has outlined just such a proposal but formulated it in non-ad hoc fashion. The proposal views a trace as a preterminal element, essentially a coindexing device. A trace left by the first application of Wh Movement will be spelled out as a variable in LF, but not a trace left in an intermediate (COMP) position. Thus the surface structure (44a) will be mapped into the partial LF (44b), but only the original trace is realized as a variable (Chomsky (1975, 35)).

(44)  a. who_{i} did you tell Bill \[\{t_{i}\}_{COMP} PRO_{j} to visit t_{i}\]  
    b. for which person \(x\), you told Bill, Bill to visit \(x\)

So in LF the right distinctions are made: a \(t\) in COMP will not appear as a variable, which is a terminal element. Chomsky proposes that To Adjunction and other rules should have access to this distinction: the rule will be blocked by terminal material intervening between want and to, although some terminal elements may not become such until at least a partial LF has been determined. This raises two major problems: first, it weakens the autonomy thesis insofar as a syntactic rule will need to “know” what is going on in LF; second, it reopens questions about “globality” since To Adjunction will have to look ahead in the derivation to the results of semantic interpretation. The second point needs further discussion.

The SSC may be seen as a condition on the applicability of rules or as a wellformedness condition operating at surface structure. If it is a condition on rule applicability, then globality will be introduced, since the SSC will block To Adjunction only if there is intervening between want and to either lexical material or a variable that is to occur later in LF. The structures for the two readings of \textit{Who do you want to succeed} will be (45) at the point that To Adjunction might apply.
(45)  a.  who<sub>i</sub> you<sub>j</sub> want [[t)<sub>i</sub>]<sub>COMP</sub> PRO<sub>j</sub> to succeed<sub>t<sub>i</sub></sub>
   b.  who<sub>i</sub> you want [[t)<sub>i</sub>]<sub>COMP</sub> t<sub>i</sub> to succeed

Under this proposal, the traces and the PRO are preterminal elements and therefore
“invisible” to the structural description of To Adjunction. However, the rule can be
blocked from applying to (45b) if it is known that either of the traces will be spelled out
as a variable in LF, as in fact will happen with the subject trace: for which person x,
you want x to succeed. This globality problem can be avoided by viewing the SSC as a
well-formedness condition on derivations. At the end of the derivation, it will be clear
whether To Adjunction has taken place. It will also be known whether a subject of to
succeed has been spelled out in LF. Let us assume that if a logical variable does occur,
the SSC, a well-formedness condition, will be able to detect a violation and will
assign *. This avoids the globality problem and has a certain naturalness in reinforcing
the analogy between (NP, t) and bound anaphora, but it will fail to handle the
contraction data, since LF will not make the right distinctions. If the LF of (45b) is
(46b), then presumably the LF of (45a) will be (46a).

(46)  a.  for which person x, you want \{
        |                                                                |
        | self                                                          |
        |you                                                            |
                             } to succeed x
   b.  for which person x, you want x to succeed

Notice that this is a claim about LF, which will hold whether one adopts the Equi
Hypothesis or the PRO analysis (section 5). In LF, material intervenes between want
and to on both the subject and object readings, and therefore a derivation will be
blocked in either case if To Adjunction has applied.\(^{20}\) What is needed is a level of
representation at which (45b) is characterized as having an embedded subject while
(45a) does not. Given my revision of the SSC above, this will not be LF, but rather
surface structure. However, the same level of representation must also distinguish a
trace in COMP from traces elsewhere, where only the latter can be “visible”. LF
makes that distinction. The problem is that there seems to be no single level of
representation that makes both distinctions. There is yet another alternative, a slight
variant on the one above, which I mention briefly: Chomsky (personal communication)
suggests that one could allow all movement rules to be quite free and specify that the

\(^{20}\) This approach will also make the wrong predictions in the case of (23b), (26b), and (27). Under
the usual assumptions, the (partial) LF will presumably be as in (23'), (26'), and (27'). The problem in (23') and
(27') is the same as in (46a), but the LF (26') would predict that Adjunction could not apply to yield Who do
you wanna be arrested by? because in LF a variable (i.e. a terminal element) intervenes between want and
to.

(23') for what thing<sub>x</sub>, you want \{
        |                                                                |
        | you                                                          |
        |self                                                          |
                             } to melt x

(26') for what person<sub>x</sub>, you want x to arrest \{
        |                                                                |
        | you                                                          |
        |self                                                          |
                             }

(27') for what person<sub>x</sub>, you want \{
        |                                                                |
        | you                                                          |
        |self                                                          |
                             } to arrest x
various conditions (SSC, Subjacency, etc.) hold for rules relating surface structures to LF. Thus SSC applies to the pair (NP, t) in the same way as it applies to the pair (NP, each other). We could then argue that To Adjunction is really an interpretive rule applying (under the Equi Hypothesis) after the rules associating (NP, t) (or, more generally, the rules specifying bound anaphora) but before the rule filling in the “logical subject” where Equi has applied, assuming these to be distinct interpretive rules. This raises many questions, and the analysis of Lasnik and Fiengo (1976) raises a problem for this view of the conditions.

Two things emerge from this discussion: first, Chomsky’s proposal to make traces preterminal elements, as stated here, will not make the right distinctions; second, given the To Adjunction data (and likewise the clitics of Portuguese, Czech, Old French, and Italian), it is difficult to see how the SSC can be viewed as anything other than a condition on the applicability of rules. If it is to be a surface structure condition on the well-formedness of derivations and if the derivations are to include the rules determining LF, the relevant distinctions will not be made.

(iii) A third and more promising way to make this analysis compatible with a successive cyclic treatment of Wh Movement would be to provide no COMP to house the offending trace. One might adopt an analysis along the lines of what Faraci (1974) has proposed for various types of purpose clauses, as in (47) and (48), with Wh Movement applying as indicated.

\[
(47) \quad [\text{COMP} [I \text{ want} [\text{[for}]], \text{PRO to succeed} [\text{wh}]]_\text{PP}]]_S
\]

\[
(48) \quad [\text{COMP} [you \text{ want} [\text{[for}]], \text{PRO to seem} [\text{[wh}]]_\text{COMP} \text{ PRO to succeed} [t]_S]]_\text{PP}]]_S
\]

Faraci (ch. 2) derives He was hoping to find a good movie from he was hoping [for [to find a good movie]_PP] on the basis of parallels with for-phrases he was hoping for a good movie, and he extends this analysis to other infinitival expressions like John built a robot to entertain his guests (cf. John built a robot for entertainment). To reinforce this analysis, there is historical evidence that for must be a preposition in such constructions and not a COMP. The argument is that if one assumes for to be a preposition, one can give a natural account of a rash of changes taking place simultaneously in late Middle English by showing that they are in fact the surface realizations of a single change in the base component. Such an analysis does not seem possible on the assumption that for was a COMP, and in that case there is no way to relate the changes and their simultaneity must be viewed as accidental (for details see Lightfoot (1976b)). There are some difficulties with this analysis; I shall mention two. If all fors are prepositions, it is not clear how to derive those in surface subject complement position, For John to leave would distress me. Clearly one does not want to permit a base structure [PP VP]_S, so the PP would have to be moved into subject...
position by transformation, perhaps by Emonds's (1972) Subject Replacement or "Intraposition". Second, if the analysis is to be compatible with the Subjacency Condition, S must be treated as a noncyclic node. If S were a cyclic node, then the movement in (47) and (48) would violate Subjacency. It is clear that S and S (I assume phrase structure rules S \rightarrow \text{TOP} \ S, S \rightarrow \text{COMP} \ S) must be cyclic, but the evidence on the status of S is conflicting. For example, if S is cyclic, then all COMP-to-COMP movements of wh elements will violate Subjacency in its present form.

\[
(49) \quad [[\text{COMP} \ldots \{(wh)_{\text{COMP}} \ldots \} s]s]s
\]

On the other hand, it has been claimed that S needs to be cyclic if *John, I wonder who saw is to be blocked by Subjacency (assuming Chomsky's reanalysis of topological sentences as special cases of Wh Movement). If only S and S are cyclic nodes, Subjacency will not block the movement indicated in (50). However, the movement in (50) will also be blocked by the SSC, even by the revision proposed above, whereby a trace but not a PRO will count as a syntactically specified subject. In which case, if such movements will always be blocked by the SSC instead of by Subjacency, we may well be able to claim that S is not cyclic; this seems plausible.

\[
(50) \quad [[\text{John}]_{\text{TOP}} [[\text{COMP} \ldots \{(\text{who}_t)_{\text{COMP}} \text{[t} \text{saw} (wh)_{s}]s]s]s]s
\]

7. In this article, I have assumed with Lakoff, Bresnan, and Chomsky that there is a transformational rule of To Adjunction. I have shown that if we assume the trace theory of movement rules, we avoid the ordering paradox inherent in Bresnan's account. This, of course, is not to say that there are not other analyses and other ways of avoiding the ordering paradox.21 However, this analysis has interesting consequences: it shows that a trace must be left on every application of NP Preposing and that the traces left will play a crucial role in the syntax. In this respect traces differ from interpreted PROs, and this led me to propose a revision to the SSC in such a way as to strengthen the autonomy thesis. Finally, there are consequences for the mode of application of Wh Movement, since the analysis adopted is incompatible with the successive cyclic treatment of this rule under the usual assumptions; various ways were discussed of making the analysis compatible with a successive cyclic Wh Movement, but no firm conclusion was reached. The wanna, gotta, usta data continue to have many interesting implications for linguistic theories.

21 For example, one might argue that the facts be handled by subcategorization restrictions on complement types, whereby have, be supposed, used(d), etc. must be followed by a VP complement and want may be followed by either VP or S. This would avoid the need for both the Equi Hypothesis and making PRO a nonterminal element (with the concomitant revision of the SSC). There are problems with such an account, but it might be made to work.
References

Bresnan, J. W. (1972) Theory of Complementation in English Syntax, unpublished Doctoral
dissertation, MIT, Cambridge, Massachusetts.
Inquiry 4, 275–343.
Chomsky, N. (1973) “Conditions on Transformations,” in S. Anderson and P. Kiparsky, eds., A
Chomsky, N. (forthcoming) “On wh Movement,” in A. Akmajian, P. Culicover, and T. Wasow,
Forms,” paper read at the Canadian Linguistic Association meeting, Edmonton.
dissertation, MIT, Cambridge, Massachusetts.
Emonds, J. E. (1972) “A Reformulation of Certain Syntactic Transformations,” in S. Peters,
dissertation, MIT, Cambridge, Massachusetts.
Horn, G. (1975) The NP Constraint, unpublished Doctoral dissertation, University of Massachu-
setts, Amherst, Massachusetts.
4, McGill University, Montreal.
Jackendoff, R. S. (1972) Semantic Interpretation in Generative Grammar, MIT Press, Cam-
bridge, Massachusetts.
6, 437–446.
Frühlingsstaging für Linguistik, Narr. Tübingen.
Language 14.
Christie, ed., Proceedings of the Second International Conference on Historical Linguis-
tics, North Holland, Amsterdam.

Department of Linguistics
McGill University
Montreal, Canada