This paper brings up various 'branching-COMP' phenomena, and shows that the version of c-command assumed in Chomsky 1981 to treat one branching-COMP phenomenon, i.e. that-trace effects, is inadequate to deal with other branching-COMP phenomena. To overcome the inadequacy, the claim is made that the notion c-command be revised so that any element under COMP will c-command others in embedded Ss, and that a rule be provided which reduces the maximal projection S' to S unless there is any element with features between it and another S-type node. For each of the claims, a good deal of evidence is provided independently of the main topic of this paper, branching-COMP phenomena. It is argued, furthermore, that the treatment of that-trace effects based on these claims is more natural than Chomsky's in the light of the generalization of environments where variables may appear.
Then, Chomsky (1981: 166) has revised it as in 2:

(2) $\gamma$ c-commands $\beta$ if and only if
(i) $\gamma$ does not contain $\beta$

Suppose that $\gamma_1, ..., \gamma_n$ is the maximal sequence such that
(a) $\gamma_n = \alpha$
(b) $\gamma_i = \alpha_j$
(c) $\gamma_i$ immediately dominates $\gamma_{i+1}$

Then if $\delta$ dominates $\alpha$, then either (I) $\delta$ dominates $\beta$, or
(II) $\delta = \gamma_i$ and $\gamma_i$ dominates $\beta$

In Chomsky 1981, though, there are cases where he does not seem to assume his own definition 2, but rather 1, most of which are concerned with phenomena involving 'branching-COMP'. A clear case is 'that-trace' effects, illustrated in 3-4:

(3) a. Who do you think $[S, t, that [S John saw $t'$]]$?
b. Who do you think $[S, t [S John saw $t'$]]$?

(4) a. *Who do you think $[S, t, that [S t' saw Mary]]$?
b. Who do you think $[S, t [S t' saw Mary]]$?

Chomsky attempts to account for the contrast among these four sentences in terms of the Empty Category Principle (ECP) in 5, a basic principle governing the distribution of empty categories:

(5) ECP: $[\alpha e]$ must be properly governed (Chomsky 1981: 250) The notion proper government referred to here means government by a lexical category, i.e. $V(erb), N(oun)$, or $A(djective)$,\(^1\) or by a coindexed NP in COMP. The notion government is defined as follows:

(6) $\alpha$ governs $\beta$ if $\alpha = X^0$ or a coindexed NP in COMP, $\alpha$ c-commands $\beta$, and $\beta$ is not protected by a maximal projection.

(cf. Chomsky 1982: 19)

The account of the grammaticality of 3a and 3b is straightforward, and does not necessitate a precise definition of c-command: the object trace $t'$ is always properly governed by the $V$; thus, it always satisfies the ECP regardless of whether its antecedent $t$ is in a branching or a non-branching COMP. The account of the contrast between 4a and 4b, on the other hand, requires a precise definition of the notion c-command. The subject trace $t'$ is not governed by any lexical category in the embedded $S$, and the only candidate for its governor is its antecedent in COMP, i.e. $t$. Then, it must be supposed, relative to the ECP, that in the grammatical

\(^1\) However, I take, in Nakajima 1984, the view that the other lexical category definable in terms of the combination of the features $[\pm V]$ and $[\pm N]$, i.e. $P(osition)$, is also a proper governor.
sentence 4b, \( t' \) is c-commanded, and consequently properly governed, by \( t \), while in the ungrammatical sentence 4a, \( t' \) is not c-commanded, and accordingly not properly governed, by \( t \). The sole structural difference between 4a and 4b is that in the former, the COMP is branching, whereas in the latter, it is non-branching. Thus, this structural difference may be associated with the difference in the c-command of \( t' \) by \( t \), in such a way that ‘if that is present, \( t \) does not c-command the subject trace in the embedded S (Chomsky 1981: 243)’. Given this, it immediately follows, as is expected, that the trace \( t' \) in 4a is not properly governed by \( t \), violating the ECP, while the one in 4b is properly governed by \( t \), satisfying it. This is Chomsky’s account of the that-trace effect in 3-4.

Notice that Chomsky’s cited idea of c-command is based on the definition in 1; if his own definition in 2 were assumed, the cited statement would not follow. To see this, let us first apply the definition in 2 to the sentences in 4a and 4b. Of the two parts in 2, part (II) is to apply when the node dominating \( \delta \), i.e. \( \gamma \), is a projection of \( \alpha \), and part (I), when \( \gamma \) is not a projection of \( \alpha \). Since, in both 4a and 4b, the node dominating \( \alpha \), COMP, is not a projection of the category of \( \alpha \), i.e. NP, (I) comes into play: in both sentences, with \( \alpha = t \) in COMP, \( \beta \) = the subject trace \( t' \), and \( \delta = S' \), \( \delta \) dominates \( \alpha \), and it also dominates \( \beta \); \( t \) c-commands \( t' \) in both sentences. Therefore, under the assumption of 2, no relevant difference arises between 4a and 4b as to the c-command of \( t' \) by \( t \).

Let us now apply to these sentences the other definition of c-command, 1. In 4a, that is present, and the first branching node dominating \( t \) is COMP, which does not dominate \( t' \); thus, \( t \) does not c-command \( t' \). In 4b, on the other hand, that is absent, and the first branching node dominating \( t \) is \( S' \), which dominates \( t' \); thus, \( t \) c-commands \( t' \). This consequence is in accord with Chomsky’s statement above, and from this, it follows, as shown, that the subject trace in 4b, but

---

2 An anonymous reviewer for this journal has commented that 2I ought to be interpreted in such a way that for any \( \delta \) that dominates \( \alpha \), \( \delta \) dominates \( \beta \), and that from this interpretation, if follows that \( t \) in both 4a and 4b does not c-command \( t' \). This interpretation, however, is hard to accept, since the embedded subject trace would never satisfy the ECP, and the extraction from the embedded subject position would be always prohibited, irrespective of whether the embedded clause lacks that or not. Even if it proves that this interpretation is viable, though, the point in this paper remains unchanged: insofar as 2 is assumed, it is impossible to distinguish the ungrammatical sentence 4a from the grammatical one 4b.
not in 4a, satisfies the ECP. In this way, the assumption of 1 as the definition of c-command is mandatory to account for the contrast between 4a and 4b along the line of Chomsky 1981 (and Stowell 1981, Kayne 1981, etc.).

However, the definition in 1 poses problems for other branching-COMP phenomena than the that-effect, as will be discussed in the next section. Before entering into them, confirmation of some assumptions upon which the structures in 3–4 are represented is in order. First, the rule of Move-\(\alpha\) is assumed to apply successive-cyclically, leaving a trace in COMP; if it were not for a trace in COMP, a trace left in the embedded subject would never have its proper governor, and always violate the ECP. Secondly, the node COMP is assumed, as represented in 7, to consist of two subparts, one dominating a complementizer, and the other used as an escape hatch:

\[
(7) \begin{array}{c}
\text{COMP} \ x \ \text{COMP} \ [\text{COMP} [\pm \text{WH}\ for]\]
\end{array} \quad \text{(Chomsky 1981: 53)}
\]

For the consistency of argument, we will maintain these two assumptions in the following arguments.

1.2. OTHER BRANCHING-COMP PHENOMENA. As suggested above, the definition of c-command in 1 raises difficulties for other branching-COMP phenomena than the that-trace effect.

We start with the case which Chomsky acknowledges must be left in an ‘unsatisfactory state’. Sentence 8a is an example of sentences with a purposive clause, and is considered to have the underlying structure in 8b:

\[
(8) \begin{array}{l}
a. \quad \text{John bought a book for Mary to read.} \\
b. \quad \text{John bought a book} \ [\text{S, [COMP PRO, for]} \ [\text{S Mary to read}} \ t]\] \end{array}
\]

The PRO has been moved from the object position into the embedded COMP, an ungoverned position. The moved PRO in COMP, under the definition of c-command in 1, ‘does not c-command the variable \(t\) because of the presence of for; yet it binds \(t\) (Chomsky 1981: 228)’. Since the adoption of 1, as Chomsky states in the same place, and as we saw in the preceding section, is essential to handle the that-trace effect, the example in 8 cannot but be left in an ‘unsatisfactory state’.

Secondly, consider an interrogative cleft sentence like 9.

3 This example was pointed out to me by Charles DeWolf.
Who do you think it was that saw Mary?

9 is similar to 4a in that the embedded subject is extracted from the finite clause with the complementizer that; yet, 9 is grammatical. The difference lies in the fact that the finite clause is embedded under the cleft clause, rather than immediately under the main clause do you think. As to the derivation of a cleft sentence, we follow Chomsky 1977, assuming that a focus phrase is base-generated in the position between a be-verb and that, and the place where a gap appears in S-structure is provided with an abstract wh-phrase, which is moved into COMP. Sentence 9 is considered to have the underlying structure in 10:

Who do you think \([S_1 \text{ it is } [S_{1'2} [\text{TOP } t_i] [S_{2'} [\text{COMP } wh_j, \text{ that}] [S_2 \text{ } t_j \text{ saw } \text{Mary}]]]]\)

10 contains two traces; one is \(t_i\), bound to \(\text{who}_1\), and the other is \(t_j\), bound to \(\text{wh}_j\). Note that \(\text{wh}_j\) is in the branching COMP. On the assumption of 1, \(\text{wh}_j\) does not c-command, and does not govern, \(t_i\). Therefore, \(t_j\) cannot be bound by \(\text{wh}_j\), and, moreover, fails to satisfy the ECP.

Thirdly, see the following two sentences, both of which have infinitive clauses with the complementizer for:

- a. John would prefer \([S_{1'} \text{ for } [S_1 \text{ Bill to have seen Mary}]]\).
- b. Who would John prefer \([S_{1'} t_i, \text{ for } [S_1 \text{ Bill to have seen } t_i]]\)?

The grammaticality of both sentences indicates that the embedded subject Bill passes the (Extended) Case Filter in 12:

*\([\text{NP } \alpha] \text{ if } \alpha \text{ has no Case and } \alpha \text{ contains a phonetic matrix or a variable.} \quad \text{(Chomsky 1981: 175)}\)

In 11a, the COMP of \(S'_1\) is non-branching; the for under it c-commands and governs the embedded subject Bill, assigning Case to it. Thus, Bill goes through the Case Filter. In 11b, on the other hand, the successive-cyclic application of \(wh\)-movement leaves a trace in the COMP of \(S'_1\), rendering the COMP branching. Under the definition of c-command in 1, the for in the branching-COMP does not c-command, and correspondingly does not assign Case to, Bill. Thus, the lexical NP Bill would fail to pass the Case Filter in 12.5

4 Chomsky 1981 substitutes, in some cases, PRO for the abstract wh-trace which he stipulated in Chomsky 1977. Cf. the case of the purposive clause in 8. In the light of the characterization of PRO, i.e., that PRO is an empty category bound to an element in A-position, however, the phrase moved into the COMP of \(S'_1\) in 10 is not taken as PRO, because it is bound to the element in TOP, an A-position. Thus, we assume, as in Chomsky 1977, that the moved phrase is an abstract wh-phrase.

5 Koster and May (1982: 121fn) also notice the necessity of supposing that for governs the embedded subject 'even when COMP contains a trace'.
At least two ways suggest themselves to avoid the problem. One way is to abandon the stipulation of the successive-cyclic application of Move-\(a\), admitting that it need not always apply in the successive-cyclic manner, and, in addition, to assume, as to bounding nodes, that the node \(S'\) and \(S\) of an infinitive clause are not bounding nodes (Chomsky 1981: 305). Given these two assumptions, the rule Move-\(a\) may move, in 11b, \(who\) in a single step from the original position to the place where it is. This movement is not in violation of the Subjacency Condition, nor does it produce a branching COMP. (I am thankful to Kunihiro Iwakura for calling my attention to this possibility.) However, this approach is problematic empirically as well as conceptually. For instance, consider, in connection with 11b, the following sentence:

(13) *Who is \([S_1 \text{ for } S_1 \text{ Bill to have seen } t]\) preferred by John? Here too, the single-step movement would move \(who\) without violating the Subjacency Condition,\(^6\) and 13 would be wrongly marked as grammatical. That is, the constraint relevant to the ungrammaticality of 13, the Sentential Subject Constraint of Ross 1967, could not follow from the Subjacency Condition, contrary to Chomsky's (1977: 131) claim.\(^7\) We will touch, in §3.1, on another piece of evidence against the approach. See, furthermore, Battistella 1983 for empirical and conceptual inadequacies of this approach. We continue to assume in the rest of this paper that Move-\(a\) always applies successive-cyclically.

\(^6\) Even if the sentential subject is supposed to be dominated by NP, the \(wh\)-movement in 13 would not violate the Subjacency in this approach. The bounding node over which \(who\) crosses would be only this NP, for the node \(S\) after a COMP into which a \(wh\)-phrase is ultimately moved must generally be assumed not to be a bounding node. Such an assumption is needed to account for the grammaticality of sentences like those in i:

(i) a. Which book \([S \text{ does John seem } S \text{ to have read } t]\)?
   b. Who \([S \text{ do you think } S \text{, that } S \text{ Mary saw } t]\)?

Without the assumption, the \(wh\)-phrase would cross, in ia, the governed \(S\) and the \(S\) after [+WH], and in ib, the \(S'\) before [-WH] and the \(S\) after [+WH]; both ia and ib would be incorrectly marked as being in violation of the Subjacency.

\(^7\) To deduce the Sentential Subject Constraint (SSC) from the Subjacency, on the one hand, and to differentiate the one-step movement and the successive-cyclic one, on the other, it is necessary to stipulate \(S'\), \(S\), and NP as bounding nodes. Since the extraction from a sentential subject, if dominated by NP, involves at least two bounding nodes, NP and embedded \(S'\), the SSC directly follows from the Subjacency. The extraction from an object clause, which is permissible, involves only one bounding node, i.e. embedded \(S'\), on the stipulation of the successive-cyclic movement, but two, i.e. embedded \(S'\) and \(S\), on the stipulation of the one-step movement. Cf. note 6.
Another way to avoid the problem is to make recourse to the assumption that a trace in COMP is deleted optionally (cf. Stowell 1981). If the trace in 11b is deleted, the COMP is rendered non-branching, and for seems able to assign Case to the embedded subject Bill. But, since a trace in COMP is supposed to receive Case from a matrix predicate (Kayne 1980), and the deletion in COMP is subject to the condition that only non-Case-marked trace is deletable (Chomsky 1981: 245), a trace in COMP will be deleted after the Case assignment has been conducted. Thus, the trace in 11b is supposed to be present when for tries to assign Case to Bill, which will result in not being assigned Case.

Fourthly, consider the derivation of 14b or 14c from 14a:

(14) a. *John would prefer [for [PRO to have seen Mary]].
b. *John would prefer [who, for [PRO to have seen t]].
c. *Who would John prefer [t', for [PRO to have seen t]]? 

The pronominal anaphor PRO is subject to the PRO theorem, i.e., that PRO must be ungoverned, which is derived from the two principles of binding theory (Chomsky 1981 and 1982). To account for the ungrammaticality of all the sentences in 14, the PROs in them must be supposed to be governed, and violate the PRO theorem. In 14a, the PRO is actually c-commanded and governed by for, violating the theorem. In 14b and 14c, however, it does not follow, under the definition in 1, that the PROs are governed: the successive-cyclic application of wh-movement moves who into the embedded COMP, as in 14b, or leaves its trace in it, as in 14c, rendering, in both cases, the COMP branching; the complementizer for under the branching COMP does not c-command, and therefore, does not govern the PRO; thus, the PROs are ungoverned, and would not violate the PRO theorem.

The fifth, and final, problem with the definition of c-command in 1 comes from LF-representations of ‘wh-in-situ’, i.e. wh-phrase in A-position. A wh-in-situ is also assumed to be moved into COMP by the application of Move-α, in this case, in LF. Cf. Pesetsky 1982 for the arguments for this. 15a and 16a are supposed to have the corresponding LF-representations in 15b and 16b, where the embedded COMPs have been rendered branching by the movement of the wh-in-situ:

(15) a. It is unclear who t saw what.
b. It is unclear [s, whatj, whoi [s t_i saw tj]]

(16) a. *It is unclear what who saw t.
b. *It is unclear [s, whoi, whati [s t_j saw ti]]

The contrast in grammaticality between 15 and 16 may probably be
accounted for anyway by the ECP (e.g. Chomsky 1981 and Huang 1983) or by some other devices (e.g. Pesetsky 1982). Our concern here does not lie in that, but rather in the binding of the traces by the wh-phrases under the branching COMPs. Aoun 1981 convincingly argues that the binding theory applies not only at S-structure, but also at LF, and Chomsky 1982 also suggests that it is operative at LF as well as other syntactic levels. If this is the case, the traces in the LF-representations of 15b and 16b must be assumed to be $\overline{A}$-bound, in order to count as variables, i.e. $\overline{A}$-bound empty categories. A category $a$ is $\overline{A}$-bound by $\beta$ if and only if $a$ and $\beta$ are coindexed, $\beta$ c-commands $a$, and $\beta$ is in an $\overline{A}$-position (Chomsky 1981: 184). On the assumption of 1, however, the wh-phrases in 15b and 16b, as Chomsky (1981: 279) notes, do not c-command the traces, which therefore, would fail to be identified as variables.

**Solution**

2.1. **Modification of c-command.** The phenomena discussed in §1.2 all indicate that it is necessary to assume that an element in COMP does c-command any elements in the embedded clause, irrespective of whether the dominating COMP is branching or not. To meet this necessity, we need to revise the notion c-command so that there will be no distinction between branching and non-branching. In Nakajima 1984, I have proposed, independently of the problems discussed in §1, to modify the notion as 17:

(17) $\alpha$ c-commands $\beta$ if neither $\alpha$ nor $\beta$ dominates the other, and the first node $\gamma$ which dominates either (A) $\alpha$, or (B) a COMP dominating $\alpha$ also dominates $\beta$.

17 eliminates thoroughly the ‘branching’ condition mentioned in 1. Part A has to do with cases where $\alpha$ is dominated by nodes other than COMP, and part B, cases where $\alpha$ is dominated by COMP. The latter treats the node COMP exceptionally in the sense that when $\alpha$ is dominated by COMP, it c-commands $\beta$ even if it is not directly dominated by $\gamma$. The node COMP serves to, say, ‘dissolve’, including itself, all intervening nodes between $\alpha$ and $\gamma$, so that $\alpha$ will be treated as if it were immediately dominated by $\gamma$. Of the two parts in 17, part B is relevant to the present argument. For the advantage of part A over 1 and 2, see Nakajima 1984.

17B partially incorporates the idea of 21, in that $\alpha$ need not be immediately dominated by $\gamma$; there may be intervening nodes between
BRANCHING-COMP PHENOMENA AND THE DEFINITION OF C-COMMAND

them. However, 17B and 2I differ in a non-trivial point: in 2I, the indirect domination of $\alpha$ by $\gamma$ is not limited to particular cases, whereas, in 17B, it is restricted to the special position, i.e. the COMP position. Thus, 2I and 17B have almost the same effect with regard to elements in COMP, but have different effects with regard to elements in other positions. For example, 17B does not take V in S or N in NP to c-command a subject phrase in them, while 2I does, because the node S or NP, which dominates the lexical category V or N, also dominates the subject phrase. This difference will be crucial when one attempts to account for the ungrammaticality of such sentences as *Whose did he meet [t brother] in terms of the ECP.

Before seeing how 17B deals with the branching-COMP phenomena in §1, let us digress a little and consider some facts that are supposed to give support to 17B, and which are not directly related with branching-COMP phenomena, the main topic of this paper.

First of all, consider the contrast between 18 and 19, where the italicized phrases are intended to indicate the coreferentiality:

(18) a. *Which of the girls who saw him, did Susan say that Mary talked to t about John?  
    b. *Which picture of him, do you think that Mary gave t to John?

(19) a. Which of the girls who saw John, did Susan say that Mary talked to t about him?  
    b. Which picture of John, do you think that Mary gave t to him?

In 18, the pronoun NP$_i$ under COMP is hard to take to be coreferential with the full-NP, NP$_j$,$^8$ whereas in 19, the full-NP, NP$_i$, under COMP is interpretable as coreferential with the pronoun NP$_j$. A similar contrast holds for ‘weakly crossed’ sentences, as is observed by Wasow 1972:9

(20) a. *How many of them, do Bostonians who know the police,
believe t are on the take?

b. *The sheriff whose picture of them, those people who had encountered the gang, always recognized t was ultimately gunned down.

(21) a. How many of the demonstrators, did the police who arrested them, beat up t?

b. Which of John’s teachers, do the people who know them, all respect t most?

Notice that both in 18 and 19, NP_i does not crossover NP_j, and both in 20 and 21, NP_i does crossover NP_j; thus, neither the contrast between 18 and 19 nor that between 20 and 21 can be accounted for by any version of the so-called crossover constraint. In all the sentences of 18–21, under the definition of 17B, NP_i proves to c-command NP_j, since NP_i is under COMP, and the node dominating the COMP, S’, also dominates NP_j. To use a term of Reinhart’s 1983, NP_j is in the ‘c-command domain’ of NP_i. She proposes, based on this notion, a condition on the interpretation of non-coreference, to the effect that a given NP must be interpreted as non-coreferential with any non-pronoun in its c-command domain (Reinhart 1983: 30). In 18 and 20, NP_j is in the c-command domain of the pronoun NP_i, and is a non-pronoun; thus, NP_i is interpreted as non-coreferential with the non-pronoun NP_j. The intended reading (i.e. NP_i=NP_j) is ruled out. In 19 and 20, on the other hand, NP_j is a pronoun; thus, NP_i may be coreferential with the pronoun NP_j. The intended interpretation is possible in these sentences.

Moreover, if the larger phrases containing the pronoun NP_i in 18 and 20 are ‘moved back’ to the non-COMP positions, then NP_i can be coreferential with the non-pronoun NP_j. Compare 18 with 22, and 20 with 23, respectively:10

(22) a. Susan said that Mary talked to one of the girls who saw him, about John.

b. I think that Mary gave a picture of him to John.

(23) a. Bostonians who know the police believe that many of them are on the take.

b. Those people who had encountered the gang always recognized the sheriff’s picture of them.

The sole relevant difference between 18 and 20 on the one hand and 22 and 23 on the other is that in the former group, NP_i is under COMP, while

10 The sentences in 23 are cited from Wasow (1972: 148).
in the latter group, it is not. To NP\textsubscript{i} in the latter group of the sentences, part A, rather than part B, of the definition of c-command in 17 applies: the first node dominating NP\textsubscript{i} does not dominate NP\textsubscript{j}; NP\textsubscript{i} does not c-command NP\textsubscript{j}; thus, NP\textsubscript{i} may be coreferential with the non-pronoun NP\textsubscript{j}, in accordance with Reinhart’s principle. These two contrasts, i.e. the one between 18, 20 vs. 19, 21, and the one between 18, 20 vs. 22–23, are both in favor of 17B, since they can be accounted for uniformly, namely, in terms of Reinhart’s principle, once we assume it.

Next, consider the following sentences, which involve the movement of NPs embedding \textit{wh}-phrases:

(24) a. We respect the pianist [the neighbours of \textit{whom} hate him].
    b. The businessman [gossip about \textit{whom} harmed \textit{his} career] will be fired soon.

It is widely assumed that a \textit{wh}-phrase is a sort of quantified phrase (e.g. Chomsky 1977, 1980, and 1981). Then, the sentences in 24 are subject to a condition on coreference between a quantified phrase and a pronoun. As such a condition, Reinhart 1983 proposes that ‘quantified NPs can have anaphoric relations only with pronouns in their c-command syntactic domain’. This condition is responsible for the ungrammaticality of the sentences in 25:\textsuperscript{11}

(25) a. *The neighbours of each of the pianists hate him.
    b. *Gossip about every businessman harmed \textit{his} career.

Notice that the bracketed parts of the grammatical sentences in 24 and the ungrammatical sentences in 25 are completely the same except that in 24, the quantified phrases are \textit{wh}-phrases, and therefore the NPs embedding them are in COMP, while in 25, the quantified phrases are not \textit{wh}-phrases, and hence the NPs containing them are not in COMP. On the assumption of 17B, this structural difference directly leads to the difference in grammaticality: in 24, the quantified NPs, i.e. \textit{wh}-phrases, c-command the pronouns, and can have an anaphoric relation with them; in 25, on the other hand, the quantified NPs do not c-command the pronouns, and cannot have an anaphoric relation with them. The intended interpretation is possible in 24, but not in 25.

Finally, consider the following sentences:

(26) a. \textit{Whom} did you talk to \textit{t} about \textit{himself}?
    b. To \textit{whom} did you talk \textit{t} about \textit{himself}?

\textsuperscript{11} Reinhart (1983: 124).
The two sentences are grammatical on a par,\(^\text{12}\) except for the difference in formality commonly observed between cases with only the object of PP preposed and ones with an entire PP preposed, as in *Whom did you talk to about it*? vs. *To whom did you talk about it*? (cf. Quirk et al. 1972: 394). Reflexive pronouns are, according to Reinhart 1983, subject to the same condition as quantified phrases, i.e., that (ordinary) reflexives can be coreferential only with c-commanding NPs. In 26a, the antecedent of the reflexive, *whom*, obviously c-commands the reflexive, satisfying the condition. In 26b, the antecedent is dominated by the branching node PP under COMP. Only on the assumption of 17B, does it turn out that the antecedent c-commands the reflexive, and therefore satisfies the condition.\(^\text{13}\) Incidentally, sentence 26b, unlike such sentences as *You talked to Bill about himself*, has no possibility of undergoing the operation of reanalysis, which is supposed to apply under the condition that its trigger c-commands, and/or is adjacent to, target categories (cf. Hornstein and Weinberg 1981, and Chomsky 1981). In 26b, the candidate for the trigger, i.e. V, does not c-command nor is adjacent to the preposition *to*; the V and the preposition cannot be fused to free the object of the preposition, *whom*, from the domination by PP, so *whom* will never, under the definition in 1, come to c-command *himself*. In this respect, 26b is basically different from sentences where a PP containing the antecedent of a reflexive occupies a non-COMP position.

All of the facts above strongly suggest that it is necessary to treat elements under COMP exceptionally, and allow them to c-command others even if they are not immediately dominated by the node immediately above COMP. This idea is formulated as 17B in the definition of c-command.

\(^\text{12}\) The grammaticality of these sentences poses a problem for Chomsky’s binding principle (A): an anaphor must be bound in its governing category. The governing category for the reflexive is S; yet, the reflexive is not bound in this S.

\(^\text{13}\) This line of argument would lead one to expect that i is also grammatical, but in reality it is not:

(i) *Which picture of the boy, did you give t to himself,*?

We speculate that reflexives are subject to conditions other than the c-command condition, say, the A-over-A principle. i violates it.

Interestingly, some speakers find ii worse than i:

(ii) *I gave the picture of the boy, to himself,*

Both i and ii violate the A-over-A principle, but, under the definition of 17B, the antecedent in i meets the c-command condition, while the one in ii does not. This is why ii is worse than i. If this reasoning is correct, the subtle contrast between i and ii also provides a piece of evidence for 17B.
2.2. Accounts of the problems in §1.2. Now that 17B is provided as part of the definition of c-command, all the facts discussed in §1.2 are not problematic any more. On the assumption of 17B, any element under COMP, whether branching or non-branching, c-commands another element in clauses below it. Thus, in 8, an example of a purposive clause, the PRO under the branching COMP c-commands and governs its trace in the embedded object position; the object trace is bound by the PRO. In 10, an interrogative cleft sentence, the wh-phrase under the branching COMP c-commands and governs the subject trace in the embedded S; the subject trace is bound by its antecedent in COMP, and satisfies the ECP. In 11a and 11b, where the lexical subject ought to be taken as assigned Case, the complementizer for under COMP, whether branching or non-branching, c-commands and governs the embedded subject, assigning Case to it; thus, the lexical subject passes the Case Filter. Next, in all the sentences of 14, where the subject PRO ought to be regarded as governed, the complementizer for in COMP, branching or non-branching, again c-commands and governs the PRO in the embedded subject position; the PRO does not satisfy the requirement of the PRO theorem that PRO must be ungoverned. Finally, in the LF-representations of 15b and 16b, the two wh-phrases under the branching COMP both c-command, and accordingly A-bind, the traces left by Move-α; thus, the traces are identified as A-bound empty categories, namely, as variables.

2.3. Change-to-S rule. The modification of c-command as in 17B is desirable, and mandatory, to account for all the facts in §1.2. But, 17B will make it impossible to account for the that-trace effect in 4 along the lines discussed in §1.1, namely, in terms of whether the subject trace t' satisfies the ECP. Since any element under COMP, whether branching or non-branching, c-commands an element in embedded Ss below the COMP, the antecedent t under the branching COMP in 4a as well as under the non-branching COMP in 4b is taken as c-commanding, and consequently properly governing, t'. The subject trace t' in both 4a and 4b equally satisfies the ECP. Thus, the grammatical sentence 4b and the ungrammatical one 4a cannot be distinguished along the lines of §1.1.

A trace left in COMP is to be A-bound. In the light of the characterization of empty category that an A-bound trace is variable (cf. Chomsky 1982: 21), it is reasonable to assume that a trace in COMP is also
If so, in 4a and 4b, the trace $t$ in COMP, as well as the subject trace $t'$, ought to be subject to the ECP. There is no possibility for $t$ to be properly governed by any lexical category in the embedded S, since it is not c-commanded by any element in the embedded S. The remaining possibility for $t$ to meet the ECP is that it proves that $t$ is c-commanded, and properly governed, by the verb of the matrix S. But, in both 4a and 4b, a maximal projection, $S'$, intervenes between $t$ and the matrix V. Thus, the trace $t$ in 4a and 4b is equally not c-commanded by the matrix V, and fails to meet the ECP. The ungrammatical sentence 4a and the grammatical one 4b are not distinguished in this respect, either.

The assumption that a trace in COMP is also subject to the ECP, however, opens the way to make a distinction between 4a and 4b. In Nakajima 1984, I have proposed, independently of the treatment of that-trace effects, a rule called ‘Change-to-S rule’:

\[(27) \text{Change-to-S rule}\]
\[S_i \rightarrow \text{S} \alpha / \cdots \text{[s}_{i-1} \cdots \text{, where } i \geq 1, \text{and } \alpha \text{ dominates no element with features} \]

This rule has the effect of reducing the maximal projection $S'$ (S' and S'') to the non-maximal one S unless there is any element with features between the $S_i$ and $S_{i-1}$. Informally put, 27 states that if an $S_i$ stands before another S-type node without being interrupted by any element with features, it should be treated, at some level of syntactic representation, substantially the same as S. The application of 27 to 28a, for example, changes it to 28b, where $\phi$ means that the node COMP dominates nothing, or nothing except a trace:

\[(28) \begin{align*}
\text{a.} & \quad \text{COMP} \quad \downarrow S' \quad \Rightarrow \quad \text{b.} \\
& \quad \phi \\
\text{S} \quad \phi \\
\text{S} \quad \phi
\end{align*} \]

The Change-to-S rule seems similar to the S'-deletion proposed in Chomsky 1981, though they are distinct rules. The S'-deletion is supposed to eliminate the node COMP under S' simultaneously with the deletion of the node S', while the Change-to-S rule merely changes the label of $S_i$ to S, and does not affect the node COMP under it. And, the applicability of the S'-deletion depends on the lexical idiosyncrasy of matrix

---

14 Kayne 1980 argues that a trace in COMP is assigned Case. If so, it should necessarily be taken as a variable, since only a variable, among empty categories, is assigned Case (cf. Case Filter in 12).

15 I revise the version in Nakajima 1984 slightly so as to incorporate PRO.
predicates, whereas that of the Change-to-S rule is determined by the structural configuration of the post-S' position.\textsuperscript{16}

Turning to the contrast between 4a and 4b, the Change-to-S rule does not apply to the embedded S' in 4a, where the complementizer \textit{that}, an element with features, is present, and S' is kept intact. In 4b, on the other hand, the rule applies to the embedded S' because no complementizer is present after S'; thus, the S' is reduced to S. The structures of 4a and 4b are now represented schematically as in 29a and 29b, respectively:

\begin{equation}
\begin{align*}
(29) & \quad a. \quad \ast \quad \ldots \quad V [S', t, \text{that} [S t' \ldots] \\
      & \quad b. \quad \ldots \quad V [S t [S t' \ldots]
\end{align*}
\end{equation}

In 29a, the maximal projection S' intervenes, as before, between the trace t and the matrix V; thus, t is not properly governed, violating the ECP. In 29b, no maximal projection intervenes between t and the matrix V; t is properly governed by the matrix V, satisfying the ECP. The ungrammatical structure 29a and the grammatical one 29b are now distinguished from each other in terms of the ECP, in relation to the trace in COMP, rather than to the subject trace in the embedded S.\textsuperscript{17}

The Change-to-S rule plays a part in accounting for the grammaticality of 9 as well. Sentence 9 and its underlying structure are repeated below for ease of reference:

\begin{equation}
\begin{align*}
(30) & \quad a. \quad \text{Who do you think it was } t \text{ that } t' \text{ saw Mary?} \\
      & \quad b. \quad \text{Who do you think } [S_{[S_{1} \text{ it was } [S_{2} [\text{TOP } t] [S_{2} [\text{COMP } \text{wh, that} ] [S_{2} t' \text{ saw Mary} ]]]]}
\end{align*}
\end{equation}

As argued in §2.2, t' is properly governed by the abstract wh-phrase in COMP, satisfying the ECP. The abstract wh-phrase, not being an empty category, is exempt from the ECP. The other trace, t, is subject to the

\textsuperscript{16} The Change-to-S rule is also similar to the S'-to-S rule in Chomsky 1981. But, the applicability of the latter seems to rely on the lexical idiosyncrasy of matrix predicates. Moreover, Chomsky (1981: 303) says that it subsumes the S'-deletion. Thus, the S'-to-S rule is supposed to achieve the same effect as the S'-deletion, deleting the node COMP together with the deletion of S'. These points differentiate the Change-to-S rule from the S'-to-S rule as well.

\textsuperscript{17} An anonymous reader has pointed out that if a trace in COMP is treated as a variable, it is to be assigned Case, a sort of feature; since the trace in COMP comes to have the feature of Case, the Change-to-S rule should, according to definition 27, not be able to apply to the embedded S' even when the COMP under the S' dominates nothing but a trace. However, recall that Case-assignment is contingent on government. In 29b, the matrix V does not govern nor Case-assign to the trace in COMP until the Change-to-S rule applies to the embedded S'. That is to say, the trace in COMP has not had a Case feature yet at the stage where the Change-to-S rule is about to apply.
ECP, but is not properly governed as given, for $t$ is protected by the maximal projection $S''$. There is, after $S''_2$, nothing but the trace $t$, an element without feature. The Change-to-S rule applies to reduce the $S''$ to $S$, rendering $t$ in TOP properly governed by the V in $S_1$. Thus, both of the two traces in 30b satisfy the ECP. The grammaticality of 30 shows that the Change-to-S rule is needed not merely to deal with that-trace effects along the line here, but also to handle another case. The necessity of 27 is motivated apart from that-trace effects.

The Change-to-S rule does not apply to $S'$ in 8b, which is repeated below as 31:

$$\text{(31) John bought a book \([S' \text{COMP PRO, for}] [S \text{Mary to read } t]\)}$$

The presence of the complementizer for in COMP prevents the application of the rule, and keeps the $S'$ node as it is. Thus, the PRO remains ungoverned. Incidentally note that the Change-to-S rule does not apply to $S'$ even when PRO is under a non-branching COMP, as in 32 below, since PRO, unlike trace, has features such as person, number, and gender (Chomsky 1981: 322):

$$\text{(32) John is too stubborn \([S', \text{PRO, } [S \text{PRO to talk to } t]\)]\)}$$

The PRO under COMP in 32, like in 31, is kept ungoverned, satisfying the PRO theorem.

**Some consequences of the proposals**

3.1. **Distribution of empty categories under want-class verbs.** The proposals made in §2, i.e. the revision in the definition of c-command and the postulation of the Change-to-S rule, yield some interesting consequences, some of which we will discuss below.

First, the proposals enable us to describe the overall distribution of empty categories (i.e. wh-trace, NP-trace, and PRO), in particular, their distribution in the subject position of infinitive clauses embedded under want-class verbs (e.g. want, desire, prefer, hate, etc.), to which no satisfactory account seems to have been offered in Chomsky 1981. Compare 33, a paradigm of want-class verbs, with 34, a paradigm of believe-class verbs:

$$18\text{ From the formulation in 27, it is correctly predicted that the Change-to-S rule does not apply to the } S' \text{ in i and the } S'' \text{ in ii:}

(i) *Who do you think \([S' \text{COMP to } t] [S \text{ John gave the book } t']\)\)?

(ii) *Who was it \([S'' \text{TOP to } t] [S' \text{ that John gave the book } t']\)\)?

The trace $t$ in i and ii is not properly governed, violating the ECP.
(33) a. Who does John want \([S, t' [S t to see Mary]]\)?
b. *Tom is wanted \([S, t to see Mary]\).
c. John wants \([S, [S PRO to see Mary]]\).

(34) a. Who does John believe \([S, t' [S t to be honest]]\)?
b. Tom is believed \([S, [S t to be honest]]\).
c. *John believes \([S, [S PRO to be honest]]\).

The paradigms clarify two contrasts, the one between 33b and 34b, and the one between 33c and 34c. These two contrasts can be accounted for by assuming, along with Chomsky (1981: 252 and 305), that the S'-deletion applies to the embedded S' of believe-verbs, but not to that of want-verbs. The application of the S'-deletion eliminates the node S' from all the embedded clauses in 34, with the embedded subject governed by the matrix V. Thus, in 34a, the variable t is properly governed, satisfying the ECP, and, at the same time, is assigned Case, passing the Case Filter; in 34b, the NP-trace t is properly governed, meeting the ECP; and in 34c, the PRO is governed, violating the PRO theorem. The non-application of the S'-deletion, on the other hand, keeps intact the S' node in the embedded clauses of 33, with the embedded subject not governed by the matrix V. Thus, in 33b, the NP-trace t is not properly governed, violating the ECP; in 33c, the PRO is ungoverned, satisfying the PRO theorem. But, since, in 33a too, the maximal projection S' still remains between the trace t' in COMP and the matrix V, t' is not properly governed, violating the ECP. Therefore, 33a would be wrongly marked as ungrammatical. This fact indicates that the simple option of application or non-application of the S'-deletion, as in Chomsky 1981, does not suffice to describe the overall contrasts between the want-class and the believe-class.

In §2.3, we have provided the Change-to-S rule, apart from the S'-deletion. The rule applies to reduce S' to S unless S' is immediately followed by elements with features. The embedded S' in 33a is immediately followed only by t', an element without features, and meets the condition under which the Change-to-S rule applies.\(^{19}\) The application of the rule reduces S' to S. Then, the matrix V want properly governs t', which in turn properly governs the other trace, t.

\(^{19}\) The embedded S under want-class verbs is assumed to have underlingly the complementizer for (cf. Chomsky 1981), which serves to distinguish 33a from i:

\( (i) *Who did John try \([S, t' [S t to meet Mary]]\)? \)

The subject trace is Case-assigned in 33a, but not in i.
Both of the two traces in 33a now satisfy the ECP, ensuring the grammaticality of the sentence. The application of the Change-to-S rule does not affect embedded subjects like \( t \) in 33b or PRO in 33c. The rule merely reduces \( S' \) to \( S \), but does not eliminate it. Between the embedded subject and the matrix \( V \) in 33b and 33c still lie two \( S \)s, the presence of which is supposed to be 'enough breadth' to prevent government (Kayne 1981).

Recall one alternative noted in §1.2 to avoid the problem connected with 11b, namely, the approach resorting to the long-distance extraction from infinitive clauses. This approach would allow the embedded subject \( \text{who} \) in 33a to move, in a single step, to the place where it is. Such a movement is not in violation of the Subjacency Condition, provided that the node \( S' \) and \( S \) of infinitive clauses are not bounding nodes. The trace left in the embedded subject position, however, would not be governed by the matrix \( V \), and necessarily violate the ECP, since it is separated from the matrix \( V \) by \( S' \) and \( S \), or by two \( S \)s if the Change-to-S rule applies to \( S' \).

Summing up the effect of the Change-to-S rule, it has almost the same effect as the \( S' \)-deletion on traces in COMP (e.g. \( t' \) in 33a): it thins the boundary between the embedded \( S \) and the matrix \( S \) to render a trace in the embedded COMP governed by a matrix \( V \). But, it does not have so strong an effect as the \( S' \)-deletion on embedded subjects (e.g. \( t \) in 33a and 33b, and PRO in 33c); it merely thins, but does not eliminate, the boundary, so an embedded subject is not governed by a matrix \( V \).

3.2. **Generalization into COMP-trace effects.** Another consequence of the present theory is that so-called *that*-trace effects can be generalized as 'complementizer-trace' effects. As with finite clauses in 3-4, the extraction from infinitive clauses depends on the two factors: the position from which a wh-phrase is moved (cf. 35 vs. 36), and the presence or absence of a complementizer (cf. 36a vs. 36b):

(35) a. Who do you prefer \([s, t, \text{for } [s \text{ John to have seen } t']]\)?
   b. Who do you prefer \([s, t \text{ for } [s \text{ John to have seen } t']]\)?

(36) a. *Who do you prefer \([s, t, \text{for } [s \text{ } t' \text{ to have seen Tom}]]\)?
   b. Who do you prefer \([s, t \text{ for } [s \text{ } t' \text{ to have seen Tom}]]\)?

It is reasonable, therefore, to attempt to treat these facts in the same way as *that*-trace effects.

The structures in 35–36 are those produced under the assumption
of the successive-cyclic \textit{wh}-movement. The configuration of the COMPs in these structures are essentially the same as those in 3–4: in 35a and 36a, as well as 3a and 4a, the COMP is branching, and in 35b and 36b, as in 3b and 4b, it is non-branching. Thus, a unified treatment is expected. As noted earlier in connection with 11b, however, on the assumption of 1, the presence of the branching COMP in 35a prevents the complementizer \textit{for} from c-commanding, and consequently assigning Case to, the lexical NP \textit{John}, which fails to pass the Case Filter. The grammaticality of 35a cannot be guaranteed insofar as 1 is assumed.

One way out of this difficulty with maintaining 1 was to assume the long-distance movement of a \textit{wh}-phrase instead of the successive-cyclic one. On this assumption, the embedded COMP in 35a does not become branching, and the above-mentioned problem does not arise. However, the \textit{wh}-movement in the same manner renders the embedded COMP in 36a also non-branching. Such a configuration of the COMP is not identical to the one in its finite-clause counterpart, i.e. 4a, where the COMP is branching. Hence, one must give up a unified treatment of 3–4 and 35–36. Actually, Chomsky (1981: 252) abandons an attempt to account for the ungrammaticality of 4a and 36a in some unified way. He instead takes recourse to the hypothesis that preposition as a class is not a proper governor. Since, though governing, the prepositional complementizer \textit{for} does not properly govern the embedded subject trace \textit{t'} in 36a, the last trace does not meet the ECP. In the light of the noted significant properties common to 3–4 and 35–36, however, it is reasonable to treat both groups in some unified way.

The two proposals made in §2 make the unification possible. In 35a, the complementizer \textit{for} under the branching COMP c-commands and governs the embedded subject \textit{John}, which is assigned Case, and goes through the Case Filter. No problem arises which takes place on the assumption of 1. Next, in both 36a and 36b, \textit{t'} is properly governed by \textit{t}; in 36b, \textit{t} is also properly governed by the matrix \textit{V}, because the Change-to-S rule applies to reduce the embedded \textit{S'} to \textit{S}; in 36a, on the other hand, \textit{t} fails to be properly governed, since the Change-to-S rule does not apply to the embedded \textit{S'}. The ungrammaticality of 36a is attributed to the failure of the trace in COMP to meet the ECP. Such an account of the contrast between 36a and 36b is completely parallel to that of the contrast between 4a and 4b which we saw in §2.3. Then, the \textit{that}-trace effect can be unified with the \textit{for}-trace effect, and correspondingly they are generalized as \textquoteleft complementizer-trace\textquoteright effects.
Conclusion

In this paper, we have been concerned with various branching-COMP phenomena, and claimed that they can be handled without contradiction if we accept the definition of c-command in 17, and the Change-to-S rule in 27, both of which have been proposed in Nakajima 1984.

The approach to complementizer-trace effects in this paper is distinct from Chomsky's in one crucial point. In the latter approach, the legitimacy of a subject trace depends on the trace itself; a subject trace is taken as legitimate only if it per se meets the principle to which a trace is subject. In the former approach, on the other hand, the legitimacy of a subject trace depends on the relation of it to a matrix V via a matchmaker, i.e. trace in COMP; a subject trace is regarded as legitimate only if a trace in COMP to which it is bound is governed by a matrix V. The present approach tries to relate an embedded subject trace to a matrix V via a trace in COMP, and treat it as if it were the object of the V. This difference of the two approaches becomes significant when we consider environments where variables may generally appear.

In non-subject positions, a variable appears immediately after V or P, or a Case-assigner. In an embedded subject position, it appears immediately after the matrix V, that is, it may occur if the embedded S does not have a complementizer, but otherwise not. As noted by Kayne 1980, furthermore, a variable is unlikely to appear immediately after A or N. These observations are put together as 37:

\[(37)\] A variable must immediately follow a Case-assigner if it follows it.

37 states the fact that a variable is legitimate only when it occurs, or is treated, as an 'object' of V or P.

Given this generalization, we should give preference to an approach which reflects the spirit of 37 over one which does not. The present approach to subject traces does reflect the spirit of 37; as noted above, it determines the legitimacy of a subject trace in relation to a matrix V, i.e., depending on whether the trace is treated as an 'object' of V. Chomsky's approach, on the other hand, does not reflect the spirit of 37; it determines the legitimacy based on a subject trace per se. Thus, the present approach to complementizer-trace effects may be concluded to

\[20\] The conditional clause is needed to exempt from 37 a subject trace in independent clauses, as in *Who t saw it?*
be more consonant with the generalization stated in 37, and therefore more natural, than Chomsky's.

REFERENCES


——. 1981. ECP extensions. LI 12.93–133.


