

Since the hypothesis in (5) does not provide an alternative explanation for the unacceptability of (i), I would rate this particular argument for the Freezing Principle as the strongest that Culicover and Wexler offer.

4. As Stan Peters has pointed out to me, the adoption of restrictions like those suggested in (13) would necessitate changes in the learning procedure outlined in Hamburger and Wexler (1975). The problem is that cases can arise in which a tentative transformational component may fail to include two transformations that are necessary in the derivation of a certain sentence *s* from a base phrase marker *b*. If hypothesized transformational rules are drawn only from the set defined by the more restricted framework, then the transformational component in the following trivial grammar is unattainable by the Hamburger-Wexler procedure:

(i) a. *Base component:*

$S \rightarrow abccd$

*Transformational Component:*

(1)  $a - b - X - d$   
       1    2    3    4  $\Rightarrow \phi, 2, 3, 4 + 1$

(2)  $b - X - d - a$   
       1    2    3    4  $\Rightarrow 3 + 1, 2, \phi, 4$

This grammar generates the language consisting of the single string *dbcca*. The Hamburger-Wexler procedure could not "learn" this transformational component, since the procedure requires that only one transformational rule be added at a time. Given the initial tentative transformational component (the empty one), the presentation of the datum ( $S[abccd]_S$ , *dbcca*) would not lead to any change in the component, since there is no single rule of the form given in (14) that would suffice to carry out the necessary permutation of the terminal elements. Several possibilities exist for modifying the procedure to avoid this problem.

5. This is not to say that obligatory rules create no serious learnability problems. Given a descriptive framework that allows both optional and obligatory rules without any restrictions, the most serious problem is precisely that of learning whether a given rule is obligatory or optional. Positive evidence would be of no avail here, since ungrammatical sentences characteristically provide the crucial evidence in favor of assigning obligatory status to a rule. This is thus another area in which the possibilities for greater restrictiveness deserve to be investigated.

P158. M15 1976

4755046  
 1977. New York: Academic Press.  
Formal Syntax, pp. 71-132.

eds: Peter Culicover, Thomas Wasow, Adrian  
 Akmajian

ON WH-MOVEMENT

Noam Chomsky

Department of Linguistics  
 Massachusetts Institute of Technology  
 Cambridge, Massachusetts

I will presuppose, in this paper, the general framework of the extended standard theory (EST), as outlined, for example, in Chomsky (1972, 1975b) and references cited there; and more specifically, the assumptions explored in Chomsky (1971, 1973, 1974, 1975b, c) and related work cited in these references. I want to examine some proposals put forth tentatively in the work cited and in so doing, to revise and extend some of the particular analyses and principles investigated. I will first review and somewhat reformulate some of the background assumptions drawn from earlier work and then apply them to several questions in English syntax.

I assume that a grammar is a theory of competence and that universal grammar (UG) is in essence a system of principles specifying the nature of linguistic representations and the rules that apply to them, and the manner in which these rules apply. A grammar (strongly) generates a set of structural descriptions and (weakly) generates a language, assigning one or more structural descriptions to each sentence of the language (and, in principle, to all potential sentences). A structural description of a sentence consists of a representation of the sentence on each linguistic level (cf. Chomsky, 1955). I assume that two of these levels are the levels of phonetic representation (PR) and what I will call "logical form" (LF), meaning by the latter the level that expresses whatever aspects of semantic representation are determined by properties of sentence-grammar. Cf. Chomsky (1975a, b, c) for discussion. Thus a grammar assigns to each sentence, in particular, a pair of representations (*pr*, *lf*), where *pr* is drawn from PR and *lf* from LF.

In accordance with EST, I assume here that a grammar consists of base rules, transformational rules, phonological rules and (semantic) interpretive rules. The base consists of a categorical component and a lexicon, the former satisfying the principles of some version of the X-bar theory (for recent discussion see Hornstein, 1975; Selkirk, 1975; Halitsky (1975); Emonds (1976); Bresnan, 1976a; Jackendoff, forthcoming), and the latter of the general character developed in Aronoff (1976). The base generates an infinite class of deep structures (initial phrase markers). I assume that thematic relations in the sense of Jackendoff (1972) and related work are determined by interaction of lexical properties and configurations of deep

structures. The transformational component of the grammar generates derivations  $D = (K_1, \dots, K_n)$ , where  $K_1$  is a base-generated deep structure,  $K_{i+1}$  is formed from  $K_i$  by a transformation, and no obligatory transformation is applicable to  $K_n$ .

The derivation  $D$  must be related to PR and LF. I will have little to say here about the relation to PR. As for LF, I assume that it is determined by interpretive rules applying to  $K_n$ . Under this assumption, it must be that thematic relations are properly expressed in  $K_n$ , though determined at  $K_1$ . I will assume that this is the case, in accordance with trace theory, as outlined in the references cited above. If so, then interpretive rules extend the derivation  $D$ , carrying  $K_n$  to a representation in LF. These interpretive rules are the rules SI-1 of Chomsky (1975b,c). It is in fact misleading to call these "rules of semantic interpretation," as in these references and elsewhere; they are more properly described as rules concerned with the syntax of LF. Note that  $K_n$  will not be surface structure in the familiar sense. It is more "abstract," by virtue of trace theory, and may be subject to nontransformational rules (e.g., "scrambling"). Some crucial aspects of PR may be determined by the extended derivation from  $K_n$  to LF. Thus, as noted first by Lees (1960), deletion seems sensitive to some aspect of semantic representation, and under the present theory that means that the possibilities of deletion are in part fixed by properties of representations at LF or between  $K_n$  and LF. Cf. Sag (forthcoming, 1976) for an analysis of such rules as VP-deletion and gapping along these lines.

This outline is extremely sketchy, and the analyses cited are not even mutually compatible in detail. I present it only so as to locate the following discussion within a familiar general framework.

I will be concerned now with a kind of "core grammar" for English consisting of a few general rules and some general conditions governing the operation of these rules. The rules in question include two transformational rules (1) and three interpretive rules (2):

- (1)
  - a. Move NP
  - b. Move *wh*-phrase
- (2)
  - a. Reciprocal rule: assign to *each other* the feature [+anaphoric to *i*] in a structure containing  $NP_i$
  - b. Bound anaphora: assign to a pronoun the feature [+anaphoric to *i*] in a structure containing  $NP_i$ , in the context [ $NP$ —Possessive— $N_x$ ]
  - c. Disjoint reference: assign to a pronoun the feature [-anaphoric to *i*] in a structure containing  $NP_i$

The rules of (2) are among those that Kenneth Hale has called "rules of construal" (cf. Hale, 1976). An informal explanation of their meaning will do for now. Let us assume that there is some standard method for indexing nonterminal symbols in deep structures, in particular, NPs; transformations will preserve the property that all nonterminals are indexed, in ways to be discussed. If *each other* is assigned the feature [+anaphoric to *i*], then the structure  $\dots NP_i \dots$  *each other*  $\dots$  (or  $\dots$  *each other*  $\dots NP_i \dots$ ) is assigned the appropriate reciprocal interpretation, whatever

this may be (for discussion, see Fiengo and Lasnik, 1973; Dougherty, 1974). A pronoun marked [+anaphoric to *i*] will be interpreted in LF as anaphoric to  $NP_i$ ; the relevant choice of  $N_x$  will be essentially as discussed in Helke (1970), including, for English, *self*, so that English (nonemphatic) reflexive is understood as bound anaphora. A pronoun marked [-anaphoric to *i*] will be understood as disjoint in reference to  $NP_i$ ; cf. Chomsky (1973); Lasnik (forthcoming). I assume that this rule falls under a more general rule of disjoint reference applying (in somewhat different ways) to all NPs. To make these vague remarks explicit, it is necessary to explain what is meant by the term "anaphoric." I assume that there is a procedure for introducing variables for NPs in LF, including pronouns, and that the notions "anaphoric," "nonanaphoric" will be understood as determining the choice of variables as the same or different. For present purposes, nothing much depends on how rules (2) are implemented, so I will not pursue the matter; as far as I can see, nontrivial questions arise in the case of (2a) and plural pronouns, the latter, a special case of problems concerning the semantics of plurality. I will assume that the rules (2) and others ultimately give representations in LF in a rather conventional form, with quantifiers and variables, for some empirical arguments, cf. Chomsky (1975c).

I assume that the rules (1) and (2) meet the following conditions:

- (3) Cycle: transformational rules, e.g., (1), meet the condition of the (strict) cycle; the subjacency condition is a property of cyclic rules, i.e., part of the definition of the cycle.
- (4) Propositional-island condition (PIC)
- (5) Specified subject condition (SSC)

I understand the notion of the cycle here in the sense of Chomsky (1973, (51)), with the qualifications given there. Assuming that transformational rules are either cyclic or postcyclic, it follows from this formulation that the rules (1), specifically (1b), are cyclic, since they apply in embedded structures.<sup>1</sup> I will understand the subjacency condition as holding that a cyclic rule cannot move a phrase from position *Y* to position *X* (or conversely) in (6):

- (6)  $\dots X \dots [\alpha \dots [\beta \dots Y \dots] \dots] \dots X \dots$ , where  $\alpha$  and  $\beta$  are cyclic nodes

For the present, I will take the cyclic nodes to be  $\bar{S}$  and  $NP_i$ ; on the effect of other choices, see below.

The subjacency condition applies to cyclic rules only; hence to cyclic transformational rules but not to interpretive rules or to postcyclic transformational rules. Thus for many people (myself included), such examples as (7) and (8) are fully acceptable:

- (7) *we want very much* [ $\bar{S}$  for [ $NP$  pictures of each other] to be on sale]
- (8) *the men expected* [ $\bar{S}$  that [ $NP$  pictures of each other] would be on sale]

Similarly, a postcyclic rule such as the major case of French clitic movement (cf. Kayne, 1975) need not, on these assumptions, meet the condition of subjacency.

It follows that rightward-movement rules are "upward bounded" (cf. Ross, 1967; Akmajian, 1975). But I am assuming that the same is true of "lowering rules" such as quantifier movement, and leftward-movement "raising" rules. It is easy enough to find phenomena that appear to violate the subjacency condition. Consider, e.g., the sentences (9), (10), where there is a relation between the phrase in bold face and the position marked by *t*, "violating" subjacency under the assumption that the rule in question is a movement rule:

(9) **John** seems [ $\bar{S}$  *to be certain* [ $\bar{S}$  *t to win*]]

(10) **who did Mary hope** [ $\bar{S}$  *that Tom would tell Bill* [ $\bar{S}$  *that he should visit t*]]

Putting the matter more carefully, a proposed condition on rules, such as subjacency, cannot be confirmed or refuted directly by phenomena of this (or any other) sort. A condition on rules can be confirmed or refuted only by rules, which observe or violate it, respectively. If the rule of NP-movement that yields (9) applies successive cyclically, as often assumed, then the rule will observe subjacency. If, as I have argued in the references cited, the rule of *wh*-movement applies successive cyclically, then it too will observe subjacency, giving (10). To find evidence to support or to refute a proposed condition on rules, it does not suffice to list unexplained phenomena; rather, it is necessary to present rules, i.e., to present a fragment of a grammar. The confirmation or refutation will be as convincing as the fragment of grammar presented. This is a simple point of logic, occasionally overlooked in the literature. The status of conditions on rules is empirical, but evidence can only be indirect and the argument, one way or another, is necessarily rather abstract and "theory bound."

The conditions (4) and (5) (PIC and SSC) refer to structures of the form (11), where  $\alpha$  is a cyclic node:

(11) ...*X*... [ $\alpha$ ... *Y*... ] ...*X*...

As in the case of subjacency, I will take  $\bar{S}$  and NP to be the cyclic nodes, delaying the discussion of other choices until later. PIC (the "tensed-S condition" of the references cited) asserts that no rule can "involve" *X* and *Y* where  $\alpha$  is a finite clause (tensed-S). SSC asserts that no rule can "involve" *X* and *Y* where  $\alpha$  contains a specified subject, i.e., a subject not containing *Y* and not controlled by *X* (I modify an earlier formulation here; I assume that *Y* contains *Y*). If  $\alpha$  contains a subject, then only the subject is accessible to rule, if the subject is specified in the defined sense.

The term "involved in" was left deliberately vague in the exploratory studies cited above, as was the category of rules to which the conditions are relevant. We may sharpen the formulation somewhat to include the desired cases and exclude unwanted ones. Let us restrict attention to rules specified in terms of a structural condition and a structural change, in the usual sense of transformational grammar (cf. Chomsky, 1955, 1961; Chomsky and Miller, 1963; Peters and Ritchie, 1973). We furthermore restrict attention to structural conditions of the elementary form (12), where  $\alpha_i$  is a constant or  $\alpha_i = vbl$ , and each constant may be either a single element of the X-bar system or a terminal string (perhaps only a single symbol):

(12) ( $\alpha_1, \dots, \alpha_n$ )

A terminal string with the successive factors  $x_1, \dots, x_n$  and the phrase marker *K* is subject to the structural change, with these factors, just in case  $(x_1, \dots, x_n)$  is analyzable as (12) with respect to *K*; i.e.,  $x_i$  is an  $\alpha_i$  with respect to *K*, where an arbitrary string is a *vbl*. Cf. references cited, and Chomsky (1975c).

We now say that a transformational rule *involves X* and *Y* when it moves a phrase from position *X* to position *Y* and a rule of construal *involves X* and *Y* when it assigns *Y* the feature [ $\pm$ anaphoric to *i*], where *X* has the index *i* (or conversely, in both cases). The two cases will be unified below.

Following a suggestion of Jean-Roger Vergnaud, we modify the definition of PIC, stipulating that  $\alpha$  is the cyclic node immediately dominating the category of *Y*. Then rule (2b), giving (8), will not violate PIC. For discussion of the effect of PIC and SSC on postulated rules of grammar, see Chomsky (1971, 1973, 1974, 1975b,c); Lasnik and Fiengo (1974); Kayne (1975); Fiengo and Lasnik (1976); Quicoli (forthcoming a, b, c); Pollock (1976).

Plainly, rules can vary from language to language within the constraints imposed by UG, but it is often assumed that conditions on rules must be invariant. This assumption is somewhat arbitrary; cf. Ross (1967); Bresnan (1972); Chomsky (1973). There is no a priori reason not to assume the opposite, and in fact, a very high level of explanatory adequacy might well be attained by a theory of UG that permitted either rules or conditions to vary, within fixed limits. To consider a case in point, Kim (1976) observes that rules of anaphora in Korean meet a condition rather like PIC, but with a somewhat different condition on  $\alpha$  of (11). There is no formal distinction in Korean between tensed and nontensed clauses, but there is a category of embedded clauses that are not islands, much like the infinitival clauses of English and the Romance languages: namely, the complements of a certain class of "assertive" verbs. It is interesting that these verbs are very close in meaning to the verbs that in English take infinitives. Thus we can formulate a variant of PIC for Korean, with the condition on  $\alpha$  modified, and we can suggest a somewhat more abstract formulation of PIC of which English and Korean are special cases. In the absence of more extensive work on rule systems in other languages, I am reluctant to suggest anything further. Note again that evidence bearing on questions of this degree of abstractness requires a fairly credible grammatical analysis, since only rules, not phenomena, have bearing on the validity of conditions on rules.

Similarly, application of SSC in a language depends on the characterization of the notion "subject" in this language. The work cited on English and Romance seems to require a formal definition of "subject" in much the sense of Chomsky (1955, 1965). For some case languages, one might want to characterize "subject" in terms of such notions as ergative, absolutive, or nonoblique. Hale (1976) proposes certain conditions on what can be taken as subject in the syntactically "unmarked" situation; in accordance with his approach, a language might characterize the notion "subject" differently, but at a cost in the grammar, in accordance with the logic of markedness. One would expect that current work in "relational grammar" will shed much light on these questions. For the moment, I would prefer to think of the

conditions cited as instances of condition-schemata, part of the core grammar of English, pending further relevant work on rule systems that may provide evidence bearing on their viability and the more general formulation of the relevant schemata.

In Chomsky (1973), two approaches to interpretation of conditions on rules are contrasted, an absolute and a relative interpretation; and the relative interpretation is proposed for conditions of the sort discussed there, including (4) and (5). Under this interpretation, a condition does not impose an absolute restriction against rules of a certain type (e.g., in the case of (4), rules not subject to PIC); rather a rule must be interpreted in accordance with the condition unless otherwise specified. Thus, one might construct a rule to "violate" the A-over-A condition, but only at a cost: the rule would have to make explicit the relevant structures so that it can apply without failing under the condition. "The logic of this approach," as noted, "is essentially that of the theory of markedness." That is, the conditions become an integral part of an evaluation measure, rather than imposing absolute prohibitions. I will continue to pursue this assumption here.

Let me now state the point somewhat more exactly. Assuming transformations and rules of construal to be defined as indicated above, in terms of (12), let us say that  $\alpha_i, \alpha_{i+j}$  are *adjacent* in (12) if each is constant (i.e.,  $\neq vbl$ ) and any term intervening between them is  $=vbl$  (i.e.,  $j=1$ , or  $j=2$  and  $\alpha_{i+1} = vbl$ ; these are the only cases we need consider in this rudimentary, but perhaps adequate theory of rules of transformation and construal).

Suppose now that we limit attention to rules of construal. Each such rule relates two categories of the phrase marker, assigning to one (the anaphor) the feature [ $\pm$ anaphoric to  $i$ ], where  $i$  is the index of the other (the antecedent). Let us say that the antecedent and the anaphor are *involved in* the rule if they are adjacent; otherwise not. Specification of constant terms intervening between antecedent and anaphor will then make the conditions inapplicable, at a cost, in accordance with the logic of markedness.

Consider now transformational rules, specifically, movement rules, which we assume leave trace. It is natural to regard the relation between a moved phrase and its trace as essentially bound anaphora. Furthermore, by pursuing this suggestion we can derive, in an interesting class of cases, a principled explanation for the fact that certain rules and rule sequences are permissible while others are not; cf. Fiengo (1974), Chomsky (1974, 1975b). But now observe that we can extend the notion "involved in" defined for rules of construal to movement rules by permitting the latter to apply freely, then applying the conditions to the moved phrase (the antecedent) and its trace (the anaphor). We can then formulate a somewhat stronger condition of autonomy of syntax (cf. Lightfoot, 1976c); namely, the semantic conditions that enter into SSC are restricted to the interpretive rules. Taking this approach, the movement rule reflected in the surface structure (13a) is blocked for the same reason that the cases of bound anaphora in (13b), (13c) are blocked:

- (13) a. \*Bill seems [ John to like  $t$  ] ( $t$  = trace of Bill)  
 b. \*Bill expected [ Mary to like himself ]  
 c. \*Bill expected [ Mary to find his way home ]

Restricting conditions (4) and (5), now, to rules of construal, we interpret them as applying to transformational rules as filters, in effect; the result of applying a transformational movement rule may or may not yield an appropriate case of "bound anaphora." It might be appropriate to give a similar interpretation to the subjacency condition for movement rules.

Under this interpretation of the application of conditions, we have the relative interpretation referred to earlier. That is, just as a language can have a rule that does not observe the A-over-A condition—at a cost, under the "logic of markedness"—so it can have a rule that does not observe, e.g., PIC—again at a cost, following the same logic. As an example, consider the "peripheral *Tous*-Movement phenomena" of Kayne (1975, pp. 63-64). Kayne argues for a general rule L-*Tous* moving quantifiers to the left; generally speaking, it observes the conditions on rules cited (cf. Quicoli, Pollock, for recent discussion). Unexplained in this or any other analysis is the appearance of the quantifier in such sentences as (14), accepted by many but not all speakers:

- (14) a. *il faut toutes* [ *qu'elles s'en aillent* ]  
 b. *il faut tous* [ *qu'on se tire* ]

In (14), the quantifier is construed with a pronoun that is within a tensed sentence. Kayne does not formulate a rule for these examples. He notes that it is doubtful that the L-*Tous* rule can be modified to apply, for one reason, because L-*Tous* applies only when the quantifier is not part of a larger NP, which would be false in these cases, and for another, because *tous* does not appear with *on*.<sup>2</sup> It seems that the phenomena can be described by a rule such as (15):

- (15) ( $vbl, V^*, Q, que, \alpha, PRO, vbl$ )

construing Q with PRO, where  $V^*$  is a certain class of verbs including *falloir, vouloir*, Q is a quantifier, and  $\alpha$  is either null or is a "sufficiently short" NP; apparently, informant judgments, which are at best conflicting, strongly prefer pronouns or simply proper nouns, with acceptability rapidly declining as  $\alpha$  becomes more complex. Suppose that (15) is the rule, more or less. Then, we do not have a violation of PIC, under the relative interpretation of conditions just outlined, the cost being the complexity of the rule (which does not strictly fall within the framework (12), incidentally). That is, PRO (or trace, if we regard the rule in question as a movement rule) is assigned the feature [ $\pm$ anaphoric to Q], but Q and PRO are not adjacent. As to whether this approach is general enough to deal with all such cases and no more, I would not hazard a guess, at this point. Note again that the question only arises when we can make a fair guess as to the relevant rule. Phenomena may be suggestive, but strictly speaking, they tell us nothing.

As formulated, conditions will apply to a construal rule when antecedent and anaphor are either (a) separated by *vbl* or (b) separated by nothing, i.e., successive. Case (a) is the general one; it is the familiar case of rules applying "over a variable." An example is *wh*-movement within a clause. Examples of (b) are few, and perhaps this case should be eliminated. One possible example is quantifier movement (or

construal; it is irrelevant for present purposes whether the quantifier is moved or generated in place and interpreted), as described in Fiengo and Lasnik (1976), with the structural description (16) for the associated surface filter.

(16)  $\langle \nu bl, t, NP, Q, X^n, \nu bl \rangle$

where we take  $X^n$  to be an element of the X-bar system standing for the categories NP, VP, AP, and  $t$  to be the trace left by movement of the quantifier Q. The rule will permit (17) but not (18):

- (17) a. *I gave the men all presents*  
 b. *I persuaded the men all to leave*  
 c. *I painted the houses all reddish-yellow*

(18) *I saw the men all*

But as noted by Postal (1976), although (17b) is acceptable, (19) is not:

(19) *I promised the men all to leave*

Assuming these judgments, Fiengo and Lasnik observe that we can explain the facts on the basis of a version of SSC that they formulate. Making slightly different assumptions than they do, suppose we assume the structures of (17b) and (19) to be essentially (20), where  $\nu$  is either *persuade* or *promise*:

(20)  $I - \nu - t - the\ men - all - [ PRO - to\ leave ]$

Suppose we take PRO in (20) to be nonterminal—in effect, a feature on the subject NP; reasons will be given below. Then (20) is subject to the analysis (16), and the rule relating *all* and  $t$  should apply. Suppose now we were to extend our notion of “involvement” to relate also adjacent constant terms, one of which is either antecedent or anaphor and the other a constant category of the X-bar system. Then the pair (*all*, *to leave*) is involved in the rule. Suppose that we modify the notion “specified subject,” in a not unnatural way, revising SSC so that given (11), no rule can apply if  $X$  and  $Y$  are involved in the rule and  $\alpha$  contains a subject not containing  $Y$  and not controlled by the category containing  $X$  or its trace (a slightly different formulation is needed if we take the rule to be one of construal). This modification leaves other cases unchanged, but now we will derive (17b) and not (19) by virtue of familiar properties of control. The case is interesting in that the constant terms “involved” are Q and VP, although the application of the rule related NP and Q. Judgments are unfortunately somewhat variable in the relevant cases and there are other possible analyses, but perhaps we can take this example at least as an illustration of the logic of the problem, and perhaps an actual illustration of the operative principles, though I am rather skeptical.

Assuming this framework, with or without the modification just discussed, we have such examples of application of conditions as the following:

(21) *Reciprocal rule:*

- a. PIC: (i) *they want [ each other to win ]*  
 (ii) *\*they prefer [ that each other win ]*<sup>3</sup>

- b. SSC: (i) *they seem to me [ t to like each other ]*  
 (ii) *\*I seem to them [ t to like each other ]*  
 (iii) *what books do they expect [ to read t to each other ]*  
 (iv) *\*what books do they expect [ t to be read to each other ]*  
 (v) *\*what books do they expect [ Bill to read t to each other ]*

*Disjoint reference:*

- a. PIC: (i) *they want [ them to win ] (they  $\neq$  them)*  
 (ii) *they prefer [ that they win ]*<sup>4</sup>
- b. SSC: (i) *they seem to me [ t to like them ] (they  $\neq$  them)*  
 (ii) *I seem to them [ t to like them ]*  
 (iii) *what books do they expect [ to read t to them ] (they  $\neq$  them)*  
 (iv) *what books do they expect [ t to be read to them ]*  
 (v) *what books do they expect [ Bill to read t to them ]*<sup>5</sup>

*NP-movement*

- a. PIC: (i) *Bill is believed [ t to be a fool ]*  
 (ii) *\*Bill is believed [ t is a fool ]*
- b. SSC: (i) *John seems [ t to like Bill ]*  
 (ii) *\*Bill seems [ John to like t ]*<sup>6</sup>

*Clitic movement*<sup>7</sup>

- a. PIC: From infinitives, but not tensed clauses, by PIC<sup>8</sup>
- b. SSC: (i) *cela le [ fera téléphoner t à ses parents ]*  
 (compare *ce garçon* in place of *le* in base position)
- (ii) *\*cela leur fera [ téléphoner ce garçon t ]* (compare *à ses parents* in place of *leur* in base position)
- (iii) *elle lui fera [ boire du vin t ]* (compare *à son enfant* in place of *lui* in base position)
- (iv) *\*qui cette nouvelle m'a-t-elle fait [ téléphoner t(qui) t(me) ]*  
 (compare *à Jean* in place of *moi* in base position)

*Quantifier movement*<sup>9</sup>

- a. PIC: (i) *J'ai tout voulu lui laisser [ manger t(tout) t(lui) ]*  
 (ii) *\*J'ai tout voulu [ que Marie mange t ]*
- b. SSC: (i) *J'ai tout laissé [ manger t a Jean ]*  
*\*j'ai tout laissé [ Jean manger t ]*  
*\*Pierre m'a tous semblé [ t (Pierre) les avoir t(tous) lus ]*  
*I ordered the boys [ to have each finished the work by noon ]*  
*\*I promised the boys [ to have each finished the work by noon ]*

## Extraposition from NP

- SSC: (i) [a review of John's book] came out yesterday  
 (ii) a review came out yesterday of John's book  
 (iii) [Bill's review of John's book] came out yesterday  
 (iv) \*Bill's review came out yesterday of John's book

These are typical illustrative examples.

Note that the subjacency condition implies the complex noun phrase constraint (CNPC) and also the *wh*-island constraints, when taken in conjunction with SSC and an independently motivated condition to block \*"I remember what who saw" while permitting "I remember who saw what"; cf. Chomsky (1973, 1975b), for discussion. Thus any rule subject to subjacency must meet the CNPC and the *wh*-island constraint, which are independent (cf., e.g., \*"what do you wonder who saw"; cf. Chomsky, 1973, for discussion of some problematic cases).<sup>10</sup> On the other hand, interpretive rules, which do not observe subjacency, do not, on these assumptions observe these constraints. Thus on these assumptions we should have such sentences as (22):

- (22) a. they heard [some funny stories about [pictures of each other]]<sup>11</sup>  
 b. they developed [some strange attitudes about [each other's books]]

We return to some examples involving rules of construal and *wh*-islands below.

When we consider interpretive rules that do not, I believe, fall within the range of rules of construal as considered here, the situation seems reasonably clear. For example, in languages where relativization involves no movement rule at all but simply interprets a base-generated pronoun in the relative clause,<sup>12</sup> relativization can violate the usual constraints fairly freely, as noted by Ross (1967) and many others since. In Hebrew, for example, there are two processes of relativization, one involving a movement rule (with optional deletion of the moved pronoun if it is a direct object, and, I assume, obligatory deletion if it is the subject) and the other involving just interpretation of a base-generated pronoun in the relative clause. The movement rule observes the usual constraints; the interpretive rule violates them fairly freely. For example, we have (23):

- (23) i. ze ha-iš še (oto) ra'iti etmol]  
 (this-is the-man [that (him) I-saw yesterday])  
 ii. ra'iti et ha-iš [še natata li et ha-sefer [še hu katav oto]]  
 (I saw the-man [that you gave me the-book [that he wrote it]]

The same is true in the (rather artificial) English *such that* construction, which, though not part of normal English, can be used readily by English speakers without instruction, suggesting that they are drawing from resources of UG. Similarly, left-dislocation in English (using the term in a sense extended beyond Ross, 1967) allows such structures as (24):

- (24) as far as John is concerned, I will never believe the claims that have been made about him

In (24), *him* is understood to refer to John, "violating" CNPC, the *wh*-island constraint, and subjacency. If our approach is correct, then, no movement rule applies in this case. Nor can a rule of construal apply, on the assumption that these rules are subject to PIC and SSC. A natural approach, I think, is to assume that pronouns are base-generated and permitted to refer freely (Dougherty's "anaphoric relation"; cf. note 12). Thus, the base rules could have introduced arbitrary NPs in the italicized positions of (23ii), (24). In some cases, rules of bound anaphora (e.g., (2b)) limit the choice of NP to bound pronouns, in effect. In the present case, however, it is not a rule of construal that is involved but rather a rule of a different category that we may call "rules of predication" (cf. Faraci, 1974). The rule of interpretation for relatives requires that the relative be taken as an open sentence satisfied by the entity referred to by the NP in which it appears; hence there must be an NP in the relative that is interpreted as having no independent reference—i.e., a pronoun with the appropriate inflections that can be given the "anaphoric" interpretation. The requirement is met automatically where relativization is by a movement rule, under the trace-theoretic assumptions of the references cited. Left-dislocation might be assumed to have a similar rule. The proposition must be "about" the item focused in the left-dislocated phrase. How close the relation of "aboutness" must be is unclear; some speakers seem to permit a rather loose connection, roughly as in the somewhat comparable Japanese *wa*- constructions that are said to permit, e.g., (25):

- (25) as for the circus (circus-wa), elephants are funny

In the narrower case, where the left-dislocated phrase is an NP, the situation is comparable to relatives. So interpreted, the rules in question fall completely outside the framework I have so far discussed and are not subject to any of the conditions cited, as seems to be the case. The same is true of rules that are not rules of sentence grammar at all, e.g., VP-deletion, which, as observed by Sag and Hankamer (1976), can apply across speakers in discourses and, correspondingly, is not subject to principles or sentence grammar; cf. (26):

- (26) a. Speaker 1: John didn't hit a home run  
 Speaker 2: I know a woman who did  
 b. John didn't hit a home run, but I know [a woman who did—]  
 c. that John didn't hit a home run is not surprising, but that Bill knows that John didn't—is a real shock.

Cf. Sag (1976, forthcoming).

Before turning to *wh*-movement, I want to say a word about "trace theory."

Let us continue to assume, as before, that categories introduced in a base derivation are indexed. Thus rules of construal can be given in the form described and derivations can simply be extended to LF; the properties of deep structure relevant to LF, and only these, are represented in surface (or shallow) structure. The question then arises, what happens to the indexing of phrases under a movement rule? For sake of illustration, suppose that English contains a rule of NP-postposing, one component of the passive rule, as often assumed. What does the theory of transformations tell us about the derived constituent structure given by application of this rule?

Suppose that the structure to which the rule applies is (27):

- (27) [S [NP<sub>i</sub> John] [VP be+en kill [NP<sub>j</sub> Bill] by [NP<sub>k</sub> e]]]

The rule of NP postposing moves NP<sub>i</sub>, replacing the terminal identity element *e*, in NP<sub>k</sub>. It is natural to assume that the moved NP, *John*, retains its index, so that in place of NP<sub>k</sub>, we have NP<sub>i</sub> of (27). It is generally assumed—and if we accept the framework of Emonds (1976), must be assumed—that the NP subject position remains after application of the rule, but that it is not filled by a terminal string. The position will later be filled by a structure-preserving rule of NP-preposing. Thus we do not assume that after NP-postposing (27) is just a VP. Following these assumptions, the output of NP-postposing is (28):

- (28) [S [NP<sub>i</sub> e] [VP be+en kill [NP<sub>j</sub> Bill] by [NP<sub>i</sub> John]]]

On the same assumptions, after NP-preposing we will have (29):

- (29) [S [NP<sub>j</sub> Bill] [VP be+en kill [NP<sub>j</sub> e] by [NP<sub>i</sub> John]]]

We may now define the substructure [NP<sub>i</sub> e] of (28) as the “trace” of NP<sub>i</sub> (= [NP<sub>i</sub> John]), and represent it by convention as *t(i)* (read: “trace of NP<sub>i</sub>”). Similarly, the substructure [NP<sub>j</sub> e] of (29) is the trace of NP<sub>j</sub>, represented as *t(j)*.<sup>14</sup> We may think of “trace,” then, as an indexed NP, with null terminal. The notion “trace,” taken (as it must be) as a function, falls naturally out of some reasonable assumptions about derived constituent structure.

Consider now the status of the item often written as PRO, which appears in such structures as (20). We may take PRO to be just base-generated *t(x)*, *x* a variable; i.e., as base generated NP<sub>x</sub>, an NP without a fixed index. The index is then assigned by a rule of control. E.g., in (20), if *v* = *persuade* and *the man* is NP<sub>i</sub>, then PRO will become NP<sub>i</sub>; and if *v* = *promise* and *I* is NP<sub>j</sub>, then PRO will become NP<sub>j</sub>. In the former case, PRO = *t(i)*; in the latter, PRO = *t(j)*.

It follows, then, that trace and PRO are the same element; they differ only in the way the index is assigned—as a residue of a movement rule in one case, and by a rule of control in the other. We would expect, then, that trace and PRO have the same effect on rule application. This seems to be the case; cf. Chomsky (1975c) for some discussion, following Quicoli (forthcoming, a). Note also that PRO is non-terminal, as required in the discussion of (16)–(20).

So conceived, trace theory (incorporating the theory of PRO), is a trivial modification of the conventional theory of transformations, making explicit assumptions about derived constituent structure that are fairly conventional, taken together with a theory of indexing that is rather natural within the framework of EST. But there are substantial empirical consequences that result from making explicit these assumptions.

This completes the review and restatement of the general framework I want to assume. Let us now turn to the rule of *wh*-movement. In this section too I will reformulate some assumptions of the work already cited.

According to the conceptions just outlined, *wh*-movement leaves a nonterminal trace, just as all movement rules do. That is, the position from which the *wh*-phrase

moved remains in the derived constituent structure with its index, identical to the index of the *wh*-phrase, now in COMP. It seems clear that words such as *who*, *what*, etc., should be regarded (at least in questions) as quantifiers of some sort. Thus at the level LF, the sentence (30) will be represented essentially as (31):

- (30) *who did John see?*

- (31) *for which x, x a person, John saw x*<sup>15</sup>

There is good reason to suppose that the rules extending a derivation to LF form such expressions as (31), and that variables are introduced in other ways as well, in particular, by the expansion of NP quantifiers such as *every* and by a rule of FOCUS. Cf. Chomsky (1975b,c), where it is shown that a variety of “cross-over phenomena” can be explained on this assumption, modifying an approach proposed by Culicover and developed by Wasow (1972) to a set of problems discussed first by Postal (1971). The variable introduced by the rules giving the meaning of quantifiers (*who*, *every*, etc) is a terminal symbol of LF. Therefore, although the structure resulting directly from *wh*-movement does not have a terminal symbol in the position of trace, the structure resulting from the interpretive rule expanding the quantifier does have a terminal symbol in this position.

In Chomsky (1975c), I referred to trace as a terminal symbol. That was an error. It is not trace that is a terminal symbol but rather the variable introduced in the position of trace by the rules giving the meaning of such quantifiers as *every* and *who* (and also by the rule of FOCUS). Difficulties in the assumption that trace is terminal were shown by Lightfoot (1976a) and Pollock (1976). Furthermore, the assumption is incompatible with the analysis of quantifier-movement (or interpretation) given above, following (essentially) Fiengo and Lasnik. The error of identifying trace itself as the variable within the scope of the *wh*-quantifier, which is overcome in the much more natural theory just outlined, resulted from concentration on too narrow a class of *wh*-phrases. Thus when we consider only such sentences as (32), the trace can be virtually identified with the variable:

- (32) *who did Mary say that John kissed t*

But the distinction becomes obvious when we consider more complex cases, such as (33), (34):

- (33) *whose book did Mary read t*

- (34) *pictures of whom did Mary see t*

Here, trace marks the position from which the *wh*-phrase was moved, but the rule expanding the quantifier *wh* will have to yield the LFs (35), (36), respectively:

- (35) *for which x, x a person, Mary read [x's book]*

- (36) *for which x, x a person, Mary saw [pictures of x]*

Correspondingly, the correct LF for (32) should be (37):

- (37) *for which x, x a person, Mary said that John kissed [x]*

The LF (37) has a terminal symbol,  $x$ , in the position of the NP source of *who*, but (32) has only a trace, i.e., only the structure  $[NP_i e]$ , where  $i$  is the index of *who*.

The rule of interpretation for *wh*-phrases must introduce the expressions given in brackets in (35)-(37) in the position of trace. We may take the rule to be essentially as follows:<sup>16</sup>

- (38) Given an  $\bar{S}$  of the form:  
 $[_{COMP} \text{--[}wh\text{-}\bar{N}\text{]--[+WH]}] [_{S} \dots t \dots ]$   
 where  $t$  is the trace of  $[wh\text{-}\bar{N}]$ , rewrite it as:  
 $[_{COMP} \text{for which } x, x \text{ an } \bar{N}] [_{S} \dots [-x\text{-}] \dots ]$

The framework assumed here is that of Chomsky (1973), and the analysis can be extended to the other cases discussed there; cf. Vergnaud (1974), for extension to relatives.

Note that on this theory, the phonetic consequences of presence of trace are limited to the terminal symbols (variables) introduced by the rule (38). We can then maintain the analysis of such examples as (39) as outlined in Chomsky (1975c), but without the complications noted by Lightfoot (1976c):

- (39) \**who do you wanna see Bill*

Similarly, consider the case of French liaison discussed by Selkirk (1972). She observes that in one style, there is no liaison across the site of *wh*-movement, though there is liaison across the site of raising of NP to subject (and, it seems, clitic movement, though she states that the facts are obscure in this case). According to the present theory, NP-raising and clitic movement cannot have phonetic effects, but *wh*-movement may, depending on the ordering of the rule (38) and the rule of liaison. In fact, it seems that speakers of French agree that there is liaison across the raising site, but there is much variation and uncertainty about the *wh*-movement cases. Perhaps this means that the ordering of rules is rather uncertain in this (somewhat artificial) style. Unfortunately, the relevant data are much less clear than one might hope, and since the style in question is not conversational but rather taught, it is not so clear how seriously one can take the facts. Some educated speakers regard them as quite dubious.

To summarize, we assume that when a phrase moves by a transformation, its category remains as an "unfilled node," and that the moved phrase and the original position have the same index. The unfilled node labelled  $i$  is  $t(i)$ , the trace of  $P_i$ , the phrase moved from position  $i$ . The trace will invoke SSC and is available for assignment of thematic relations. PRO and trace are identified; they differ only with respect to the origin of the index. The position of trace may be filled by a phrase containing a variable, by expansion of a quantifier. There may be phonetic effects of trace in the latter case.

The rules and conditions given so far permit *wh*-movement within a clause, giving such sentences as (40), but not extraction of *wh*-phrases from a clause,<sup>17</sup> as in (41):

- (40) *who did Mary meet t*

- (41) *who did you tell Mary that she should meet t*

The two cases are in fact quite different in character. Many languages permit the first but not the second (e.g., Russian, German). Furthermore, whereas *wh*-movement within a clause is unconstrained, extraction from a clause is lexically governed, as has frequently been remarked. Thus we have such examples as (42):<sup>18</sup>

- (42) a. \**what did John complain that he had to do this evening*  
 b. \**what did John quip that Mary wore*  
 c. ?*who did he murmur that John saw*

Just what property of the matrix VP permits it to be a "bridge" (in the sense of Erteschik, 1973), permitting escape of the *wh*-phrase from the  $\bar{S}$  "island," is unclear. Some proviso is necessary, however.

Suppose that we formulate the basic rule of *wh*-movement essentially as (43):

- (43) move *wh*-phrase into COMP

The rule will apply freely clause-internally, but will not yet move the *wh*-phrase over a bridge. We may then formulate a language-specific COMP-COMP movement rule (44):

- (44) move *wh*-phrase from COMP to a higher COMP over a bridge

The structural description of this rule (subject to modifications about placement in COMP to be discussed) will be approximately (45):

- (45) (COMP,  $X$ , *wh*-phrase, *wh*), where  $X$  contains a VP with certain special properties

If we incorporate the "bridge" properties in (45), then the rule will not fall strictly within the format we have proposed for transformational rules. Moreover, under the relative interpretation of conditions discussed before, it might be argued that the conditions are inapplicable; more precisely, it is easy to see how "involved in" can be sharpened so as to make them inapplicable, along the lines discussed earlier. Suppose, alternatively, that we dispense with (45) and interpret the "bridge" conditions as conditions on rules of interpretation. Then COMP-COMP movement by (43) will be blocked by the conditions. We must therefore introduce a language-specific proviso in (11), for English, namely, (46):

- (46) where  $Y$  is not in COMP

Which of these approaches is preferable is unclear. I will assume the latter, without much reason. Thus we add the language-specific proviso (46) to (11), permitting COMP-COMP movement, and we assume that the "bridge" conditions fall within the interpretive rules, either SI-1 or SI-2 (cf. Chomsky, 1975b, c; Erteschik, 1973).

Sentence (41) will be formed, as in the references cited, by successive-cyclic application of *wh*-movement, now understood to be reapplication of (43). The rule is subject to all of the conditions on movement rules, so that we have the consequences already noted.<sup>19</sup>

Continuing to adopt the framework of the references cited, as modified above, I will assume that the rule (43) places a *wh*-phrase within the COMP node to the left



of [ $\pm$ WH], which is realized phonetically as *that*, *for*, or null. There are a number of apparently rather idiosyncratic rules that determine the phonetic realization of the items in COMP. A formulation given in Chomsky (1973) can be considerably improved and extended, but I will not go into the matter here. One general rule for Modern English is that sequences of the form *wh*-phrase + complementizer are not permitted, as they were in earlier stages of the language. Thus we will have rules such as (47), (48):

- (47) *wh*-phrase becomes null  
 (48) a. *that* becomes null  
       b. *for* becomes null

One of the three must apply. By general conditions on recoverability of deletion, which we may assume to exist though they are not understood in detail, (47) will be inapplicable when the *wh*-phrase contains actual lexical content (e.g., prepositions, possessives, etc.). The rules (48) apply more broadly; e.g., *that* can be deleted under certain circumstances in nonrelatives, *for* is deleted immediately following verbs of the *want* category and under certain circumstances before *to*, etc.

I will assume that the *wh*-phrase moved by the rule is as determined by Bresnan's relativized A-over-A principle (cf. Bresnan, 1976a; Woisetschlager, 1976, Sag, 1976, for somewhat different versions).

The rule of *wh*-movement has the following general characteristics:

- (49) a. it leaves a gap  
       b. where there is a bridge, there is an apparent violation of subjacency, PIC, and SSC  
       c. it observes CNPC  
       d. it observes *wh*-island constraints

The properties (49) follow, on the theory outlined, from the assumption that *wh*-movement moves a phrase (implying (a)), observes SSC, PIC, and subjacency (implying (c) and (d)<sup>20</sup>), and is permitted from COMP-to-COMP under "bridge" conditions (implying (b)).

So far, I have been recapitulating and somewhat revising earlier work. Now I want to turn to the main question of this paper, namely, (50):

- (50) Where we find the configuration (49) in some system of data, can we explain it on the assumption that the configuration results from *wh*-movement?

In other words, does the configuration (49) serve as a kind of "diagnostic" for *wh*-movement. That it may have been suggested, quite tentatively and without elaboration, in earlier work. I now want to investigate the plausibility of the contention. The following remarks, then, have a narrower and a broader aim. The narrower aim is to provide evidence that certain examples with the configuration (49) may in fact plausibly be understood as cases of *wh*-movement. The stronger aim is to suggest that this may be true in general. By the logic of the question, the stronger proposal cannot be demonstrated but only suggested.

I will assume, following the analysis in the references cited, that *wh*-movement is what underlies restrictive and nonrestrictive relatives and direct and indirect questions. There are, of course, some distinctions among these cases. Some of them can be accounted for by considering the contexts in which the *wh*-movement rule applies. E.g., questions but not relatives can have *wh*-movement of adjective phrases, but this distinction will obviously follow from the rule of relativization, whether it is a raising rule (cf. Vergnaud, 1974) or an interpretive rule. In other cases, stipulation may be necessary to distinguish some types from others (though this is not obvious), but if so, there seems no compelling reason to suppose that the stipulation is a condition on the *wh*-movement rule itself, though even if it were, it would not materially affect the point at issue.

Apart from these cases, the best-studied relevant example is the case of comparatives. It has been frequently noted (first, I believe, by David Vetter) that comparatives essentially have the properties (49), and it was therefore proposed in Chomsky (1973) and Vergnaud (1974) that "comparative deletion" is in reality a case of *wh*-movement. The contrary position is argued by Bresnan in an important article (Bresnan, 1975), which, together with Bresnan (1972, 1973), constitutes the most extensive and illuminating study of comparatives available. The issue is complex. Let me try to sort it out.

First, is there evidence for a *wh*-movement rule underlying comparatives? For some dialects of English, there is direct evidence for such a rule, as noted in Bresnan (1972). Thus many dialects of American English normally have such comparatives as (51):

- (51) a. *John is taller than what Mary is*  
       b. *John is taller than what Mary told us that Bill is*

For such dialects, the comparative rule is virtually identical to the general rule of *wh*-movement. Subject to the qualifications given above, it seems that the rule postulated for relatives and questions can simply extend to comparatives, with essentially no change. The properties (49) will then follow directly.

But there is evidence (Richard Kayne, personal communication) in support of a *wh*-movement analysis for other dialects of English as well. Consider the sentence (52), where brackets bound internal cyclic nodes:

- (52) a. *Mary isn't the same as [ she was five years ago ]*  
       b. *Mary isn't the same as [ John believes [ that Bill claimed [ that she was five years ago ] ] ]*  
       c. *\*Mary isn't the same as [ John believes [ Bill's claim [ that she was five years ago ] ] ]*  
       d. *\*Mary isn't the same as [ I wonder [ whether she was five years ago ] ]*

This construction has the properties (49). The "gap" is an adjective phrase, just as in comparatives; we can replace "the same as" by "taller than" throughout. There are similar constructions in which even the phrase *the same* does not appear, as in (53), etc.:

- (53) a. *Mary is (more or less) as she was five years ago*  
 b. *Mary is rather like John thought she was* [in colloquial English]  
 c. *Mary isn't as John believes that Bill claimed that she was five years ago*

In these cases, a deletion analysis, if possible at all, seems rather artificial, since in contrast with comparatives, there is no overt matrix phrase that can trigger and control the deletion. We can easily account for (52-3) by a *wh*-movement rule of the sort postulated for the dialects that permit (51). The rule will give (54a), just as it gives (54b) in the dialects that have an overt *wh*-form in comparatives:

- (54) a. *Mary isn't (the same) as* [what she was five years ago]  
 b. *Mary isn't taller than* [what she was five years ago]

Sentence (54b), for dialects that do not permit it, can be regarded as the structure underlying (55) by a rule of *wh*-phrase deletion, falling under (47):

- (55) *Mary isn't taller than she was five years ago.*

The same rule will give (52-3). The dialects differ, then, in obligatoriness of *wh*-phrase deletion; as noted, this and related rules are subject to a variety of apparently rather idiosyncratic conditions.

According to this analysis, the sentences of (52)-(53) are regarded as analogous to those of (56):

- (56) a. *Mary isn't different than* [ what she was five years ago ]  
 b. *Mary isn't different than* [ what John believes [ that Bill claimed [ that she was five years ago ] ] ]  
 c. *\*Mary isn't different than* [ what John believes [ Bill's claim [ that she was five years ago ] ] ]  
 d. *\*Mary isn't different than* [ what I wonder [ whether she was five years ago ] ]

Examples (56c,d) are ruled out by subadjacency, PIC, and SSC. Under the analysis that presupposes (54a) underlying (52a), (53a), the same is true of (52c, d), etc.

Proceeding, we may treat *as*, *than* as prepositions, analogous to *than* in (56). This seems reasonable anyway; it means that such sentences as (57) will be analyzed as having final prepositional phrases of the form P NP, rather than being derived by deletion of *be* from (58):

- (57) *John is taller than Bill*

- (58) *John is taller than Bill is*

Cf. Hankamer (1973) for arguments supporting this analysis of (57).

The analysis of (52-3) along these lines seems natural and perhaps compelling. If it is correct, then all dialects that permit (52-3) have a rule of *wh*-movement forming comparatives. Therefore, there is no need for a new rule of comparative deletion.

If this is correct, we might propose further that there do not exist rules of "deletion over a variable." Thus the category of permissible rules is reduced, always a welcome step. Furthermore, we have some support for a positive answer to the

question (50). Correspondingly, we have some evidence that the island constraints of (50iii, iv) can be explained in terms of general and quite reasonable "computational" properties of formal grammar (i.e., subadjacency, a property of cyclic rules that states, in effect, that transformational rules have a restricted domain of potential application; SSC, which states that only the most "prominent" phrase in an embedded structure is accessible to rules relating it to phrases outside; PIC, which stipulates that clauses are islands, subject to the language specific "escape hatch" (46)<sup>21</sup>). If this conclusion can be sustained, it will be a significant result, since such conditions as CNPC and the independent *wh*-island constraint seem very curious and difficult to explain on other grounds.<sup>22</sup> Whether or not these further consequences prove tenable, it seems clear that a strong argument would be required to show that English has a second rule of comparative deletion that gives exactly the same forms as the independently motivated and quite general *wh*-movement rule (subject, again, to the qualification on p. 87). It would be rather paradoxical for a language to contain a general rule of *wh*-movement forming all comparatives (and much else), along with a second rule (comparative deletion) that is extensionally identical (as a mapping) with the first over the subdomain of structures such as (58).

Bresnan (1975) argues that the rule of comparative formation falls together with her rule of comparative subdeletion, which gives such sentences as (59):

- (59) *they have many more enemies than we have—friends*

She argues further that comparative subdeletion is a rule of deletion over a variable. Let us put aside the second contention for the moment and ask whether there is strong evidence that comparatives fall under a rule that gives comparative subdeletion as a special case. I am not convinced. In fact, Bresnan cites differences that seem to me significant (cf. pp. 58-9, particularly note 10), and that raise a serious question as to whether these rules are subcases of a single process. A rule to provide the cases of comparative subdeletion is no doubt needed, in some form, but I see no compelling reason to suppose that a rule of comparative deletion will fall out as a special case. If not, then there is no reason on these grounds for postulating a rule of comparative deletion, essentially duplicating the effects of the rule of *wh*-movement and *wh*-phrase deletion (independently motivated for (51), (52), and far more general in extension) over the subdomain of comparatives. I will tentatively conclude, then, that English does not have a rule of comparative deletion.

It remains to discuss Bresnan's argument that comparative subdeletion is a rule of deletion over a variable meeting such conditions as (50iii, iv), and other arguments that she puts forth to show that island constraints cannot be explained in the terms suggested here. I will return to these questions below. Note that these considerations relate to the query (50) and the broader aim sketched above, but they do not bear on the question as to whether English has a rule of comparative deletion in addition to *wh*-movement and *wh*-phrase deletion.

Bresnan notes that comparatives have the cross-over properties discussed by Postal, Wasow and others. She then argues that cross-over properties are not a

diagnostic for movement rules, on her assumption that comparatives are formed by a deletion rule. If she is correct, it would follow that the explanation for cross-over suggested in Wasow (1972) and in another form in Chomsky (1975b,c) is incorrect or at least incomplete, since it would seem that this explanation could not be extended to deletion rules. But if comparatives are formed by *wh*-movement, as suggested above, it follows at once that they should have exactly the cross-over properties of relatives and questions; the proposed explanations would directly cover the cases that Bresnan cites, with no changes. It seems to me fair to take this as an indirect but significant additional argument in favor of the hypothesis that comparatives are formed by *wh*-movement. The argument is, in this case, that under this hypothesis we retain a fairly general, and, I believe, rather convincing explanation for cross-over phenomena.

The cross-over cases that Bresnan cites are (essentially) the following:

- (60) a. *more students flunked than—thought they would (flunk)*  
 b. *more students flunked than they thought—would (flunk)*

*Students* is the understood subject of *think* in (a) and *flunk* in (b). But in (a), *they* can refer to the students, whereas in (b) it cannot.

According to a *wh*-movement analysis, the structure of (a) and (b) after *wh*-movement will be approximately (61a), (61b), respectively:

- (61) a. *more students flunked than* [[ *wh-many (students)* ] [ *t thought [they would flunk]* ] ]  
 b. *more students flunked than* [[ *wh-many (students)* ] [ *they thought [t would flunk]* ] ]

The structures of (61) are analogous in relevant respects to the direct questions (62a), (62b):

- (62) a. *how many (students) [t thought [they would flunk]* ]  
 b. *how many (students) [they thought (did they think) [t would flunk]* ]

The analysis proposed in the references cited accounts for all of these cases, in what seems to me a very natural way, on the basis of fairly general principles. It remains to be determined whether all cases of cross-over in comparatives fall so readily under the analysis developed for *wh*-movement.

I am not arguing that a language might not have two rules yielding a single structure such as comparatives, but rather that a substantial argument must be given to motivate a second rule, particularly, when it is extensionally equivalent to the first over a subdomain of the first. Cases of "double rules" exist, it seems. Recall the case of Hebrew relatives discussed above (cf. (23)). Here, however, the two processes do not cover the same domain for principled reasons, as noted.

Let us turn now to another example of a grammatical process that gives the configuration (49), namely, topicalization. To begin with, topicalization does yield this configuration. Thus we have (63):

- (63) a. *this book, I really like*  
 b. *this book, I asked Bill to get his students to read*  
 c. *\*this book, I accept the argument that John should read*  
 d. *\*this book, I wonder who read*

Before proposing an analysis of topicalization, let us consider again left-dislocation as in (64) (cf. (24)):

- (64) *as for this book, I think you should read it*

Plainly in this case, there can be no transformational analysis in our terms since no transformation can "create" the structure "as for this book" or even more complicated phrases that can appear in this position. Suppose, then, that we postulate the base rule R1 in addition to Bresnan's R2, already assumed:

- (65) R1:  $\bar{S} \rightarrow \text{TOP } \bar{S}$   
 R2:  $\bar{S} \rightarrow \text{COMP } S$

In addition, we assume the semantic rule of predication already discussed informally in connection with (24).

As Sag observes, structures such as (64) can be embedded, with varying degrees of acceptability, as in (66):

- (66) *I informed the students that as far as this book is concerned, they would definitely have to read it*

To accommodate such cases, let us revise rule R2 to (67):

- (67) R2:  $\bar{S} \rightarrow \text{COMP } \left\{ \begin{array}{l} \bar{S} \\ S \end{array} \right\}$

These rules will allow recursions, giving such sentences as (68):

- (68) *as for John, as far as this book is concerned, he will definitely have to read it*

If such structures are to be permitted, the rule of predication will have to be extended in an obvious way.

Let us now return to topicalization. Suppose that the analysis is just like left-dislocation, except that in the TOP  $\bar{S}$  structure,  $\bar{S}$  is a *wh*-clause—in effect, a kind of free relative, as in comparatives. Thus (63b) will derive from (69), which in turn derives from (70):

- (69)  $\left[ \bar{S} \left[ \text{TOP } \textit{this book} \right] \left[ \bar{S} \left[ \text{COMP } \textit{what} \right] \left[ \textit{I asked Bill to get his students to read t} \right] \right] \right]$

- (70) *this book, I asked Bill to get his students to read what*

To form (63b) from (69) we use the obligatory rule of *wh*-phrase deletion already motivated for comparatives.

On these assumptions, (63b) is analogous to such sentences as (71):

- (71) a. *this book is what I asked Bill to read*  
 b. *it is this book that I asked Bill to read*

From the point of view of the semantics as well as the syntax, the analogy seems appropriate.

In (69) the rules already discussed introduce a bound variable, giving (72):

- (72) [ $\bar{S}$  [<sub>TOP</sub> *this book*] [ $\bar{S}$  [<sub>COMP</sub> *what x*] [*I asked Bill to get his students to read x*]]]

Deletion of the *wh*-phrase leaves an open sentence,<sup>25</sup> which we may assume to be interpreted by the predication rule that applies in the case of left-dislocation and relatives.

It follows from these assumptions that topicalizations, like left-dislocation, should be possible with varying acceptability within embedded clauses, as in (73):

- (73) *I informed the students that this book, they would definitely have to read*

It seems to me that (73) is about on a par with the formally analogous (66).

It also follows that topicalization should have the properties of (49), as was illustrated in (63).

Before we leave this topic, let us consider some further consequences of the analysis. Notice that although topicalization is possible within *that*-clauses, as in (73), it is impossible within relatives or questions. Thus we cannot have (75) corresponding to (74):

- (74) *John gave away the books to some friends*  
 (75) a. *\*to whom the books did John give away (to whom did the books John give away)*  
 b. *\*whom the books did John give away to*  
 c. *\*the boy to whom the books John gave away*  
 d. *\*the boy whom the books John gave away to*

The structure underlying, e.g., (75c,d) would on our assumptions be (76):

- (76) *the boy* [ $\bar{S}$  COMP [ $\bar{S}$  [<sub>TOP</sub> *the books*] [ $\bar{S}$  COMP *John gave away which to whom*]]]

The structure (76) is generable by the base rules. Furthermore, *wh*-movement can apply to *which* in the embedded sentence, placing it in the internal COMP position and leaving a trace. If the dominating  $\bar{S}$  were within a *that*-clause instead of a relativized NP, we would then derive (77):

- (77) *I believe that the books, John gave away to some friends*

While (77) is not very elegant, it is surely far better than (75c,d), which would derive from (76) by still another application of *wh*-movement, namely to (*to*) *whom*, placing it in the position of the higher COMP.

The problem with (75) does not seem to be just a surface difficulty; compare the sentences (78), which seem much better than (75) and more or less on a par with (77):

- (78) a. *I believe that this book, you should read*  
 b. *I believe that this book, you should give away*  
 c. *I believe that his friends, John gave some books away to*

We can explain the impossibility of the sentences (75) by essentially the same line of argument that accounts for the *wh*-island constraint. Movement of (*to*) *whom* to the internal COMP is blocked, because the internal COMP is already filled by *which* under the *wh*-movement analysis of topicalization. Movement of (*to*) *whom* to the higher COMP node is impossible because it would violate SSC and PIC (and, if  $\bar{S}$  is a cyclic node, subjacency). Even if the already moved *which* could move by COMP-COMP movement to the higher COMP, freeing the lower one, subsequent movement of (*to*) *whom* to the lower COMP would be excluded by strict cyclicity. Since the trace left by movement of *which* is (when replaced by a variable) taken to be satisfied by *the books* under the predication rule, there is no possible interpretation of (76) or of any of the sentences of (75). Thus there are a number of reasons why (75) are ungrammatical, on the *wh*-movement analysis of topicalization. In effect, we can form (75) only by extraction from a *wh*-island.

There is some reason to suppose that  $\bar{S}$  is indeed a cyclic node. Thus consider the sentence (79):

- (79) *it is believed [ $\bar{S}$  that [ $\bar{S}$  [<sub>TOP</sub> *this book*] [ $\bar{S}$  *you should read*]]]*

As it stands, (79) is on a par with (78). But NP-movement cannot apply to (79) to yield (80):

- (80) *\*this book is believed you should read*

The explanation for this fact could be that  $\bar{S}$  is a cyclic node, so that the application of NP-movement to (79) would violate subjacency. Note that we cannot appeal to PIC in this case, because TOP is outside of the finite clause, presumably.

On the assumption that  $\bar{S}$  is cyclic, it follows that left-dislocation should also be impossible in relatives, just as topicalization is. Thus (81) should be as bad as (75):

- (81) *the boy to whom, as far as this book is concerned, John gave it away*

My intuitions collapse at this point. Some instances of these structures seem to me perhaps acceptable, e.g., (82):

- (82) *I want to find a corporation to which, (as far as) my new invention (is concerned), I can offer (it) with a feeling of security that it will be exploited for the good of mankind.*

Compare (82) with the parenthesized phrases deleted. If, indeed, these two sentences are significantly different in status, this may show that  $\bar{S}$  is not a cyclic node, since on the assumption that it is not, (82) should be grammatical but the corresponding topicalized form (with parenthesized phrases deleted) should not be. However, I do not think that any conclusion can rest on such data.

There is, I think, a clear difference between topicalization and left-dislocation in direct questions. Compare (83), (84) (and (75)):

- (83) a. *\*to whom, this book, should we give*  
 b. *\*this book, to whom should we give*  
 c. *\*John, who do you think saw*
- (84) a. *\*to whom, as for this book, should we give it*  
 b. *as for this book, to whom should we give it*  
 c. *(as for) John, who do you think saw him*

The sentences (83a,c) are ruled out by SSC and PIC (i.e., extraction from *wh*-island), as before. (83b) is ruled out because it has a doubly filled COMP node under the *wh*-movement analysis of topicalization. There is no barrier against (84b,c) however, since there is no *wh*-movement in left-dislocation, just as I assume that there is none in relativization where a pronoun appears in the open sentence. To block (84a) we must assume either that  $\bar{S}$  is cyclic or that TOP is not a bridge for COMP-COMP movement.

Indirect questions are apparently like relatives, requiring no special comment.

Over a considerable range, then, analysis of topicalization as *wh*-movement seems quite reasonable. The proposal is that in the TOPIC position there is a base-generated structure and that the associated proposition, which is an open sentence except for some cases of left-dislocation, says something about it. There are in principle two ways to derive an open sentence: by *wh*-movement (and *wh*-phrase deletion; but cf. note 25) or with an uninterpreted pronoun. Both of the available ways are used. The first gives topicalization; the second, left-dislocation.

I do not want to suggest that there are no remaining problems. There are—quite a few. Unfortunately, crucial examples seem often to involve ambiguous judgments. I will simply leave the matter here. As far as I can see, the *wh*-movement analysis of topicalization is reasonably successful, has some explanatory power, and does not, to my knowledge, face difficulties that do not arise in a comparable form on other approaches. It also has the advantage of extending the framework outlined to yet another class of cases, thus offering some further evidence in support of a positive answer to (50).

Consider next cleft sentences. In Chomsky (1974) I suggested that these be derived from a structure in which the focussed phrase is base-generated in the predicate position of the matrix sentence rather than by a movement rule. We can then take the associated proposition to be formed by *wh*-movement, in conformity with the analysis that we are now considering. As has often been noted, topicalization and cleft seem to share striking properties. The suggested analysis exploits this fact.

Actually, we can draw an even closer connection between topicalization and clefts by pursuing a slightly different path. Suppose that we take the underlying structure of cleft sentences to be as in (85):

- (85) *it - is -  $\bar{S}$*

Then any topicalized sentence can appear in (85) in the position of  $\bar{S}$ . Thus alongside of (63) we have (86):

- (86) a. *it is this book that I really like*  
 b. *it is this book that I asked Bill to get his students to read*  
 c. *\*it is this book that I accept the argument that John should read*  
 d. *\*it is this book that I wonder who read*

Two provisos are necessary. First, we must stipulate that left-dislocations cannot appear in (85); the  $\bar{S}$  within  $\bar{S}$  must be subject to *wh*-movement. Second, as in a number of other constructions, the COMP node cannot become terminally null under rules (47), (48).<sup>26</sup> As far as *that* is concerned, deletion in topicalization and left-dislocation is presumably a special case of the process that applies uniformly in matrix sentences.<sup>27</sup> Perhaps one can extend to (86) the restriction against deleting *that* in subjects and extraposed *that*-clauses.

Let us assume that these matters can be properly worked out. Then we should expect to find such sets as the following:

- (87) a. *the book is what I read; the book, I read; it was the book that I read*  
 b. *this book is what I asked Bill to read; this book, I asked Bill to read; it was this book that I asked Bill to read*  
 c. *John is who I want Bill to tell Mary to meet<sup>28</sup>; John, I want Bill to tell Mary to meet; it is John that (who) I want Bill to tell Mary to meet*  
 d. *in England is where I told Bill that I want to live; in England, I told Bill that I want to live; it was in England that I told Bill that I want to live*  
 e. *where he went to school is what I wish you would ask him to emphasize in his application; where he went to school, I wish you would ask him to emphasize in his application; it is where he went to school that I wish you would ask him to emphasize in his application*  
 f. *pea green is what he painted his boat; pea green, he painted his boat; it is pea green that he painted his boat.<sup>29</sup>*

The structures, in each case, are as in (88), respectively:

- (88) NP is  $\bar{S}$ ; [ $\bar{S}$  TOP  $\bar{S}$ ]; *it is* [ $\bar{S}$  TOP  $\bar{S}$ ]

In each case, *wh*-movement must take place within S. One would not expect the parallelism to be exact, since the surface rules of interpretation for the three structures, though similar, seem to be somewhat different. It seems to me a reasonable hypothesis, however, that it is just the interpretive rules that account for whatever differences there may be among the three structures. Of course, this hypothesis suggests a direction for research rather than a confirmed result.

There are other examples of clefts that cannot be analyzed in this way, however; e.g., the following, from Pinkham and Hankamer (1975):

- (89) a. *it's only when it rains that we have to sweep the court*  
 b. *it was (purely) out of spite that he assigned it that number*  
 c. *it was only reluctantly that he agreed to swim at all*

Note that in these cases we do not have parallel structures of the sort illustrated in (87). We do, however, have parallels with adverb preposing:

- (90) a. *only when it rains we have to sweep the court*<sup>30</sup>  
 b. *(purely) out of spite, he assigned it that number*  
 c. *only reluctantly he agreed to swim at all*<sup>30</sup>

Suppose we postulate that adverb preposing, in some cases at least, places the adverb in the position TOPIC. Then rule (85) already accommodates (89). If this is correct, we have in effect two sources for clefts but no separate rules; furthermore, we need not postulate a "structure-building" rule, adding the "it-be-Predicate" structure by transformation. The latter is a much-to-be-desired consequence for two reasons. Most importantly, it is a vast and otherwise (to my knowledge) unmotivated extension of the power of transformations to permit them to be "structure-building" in the required sense.<sup>31</sup> Furthermore, it would simply be an unexplained accident that the "structure-building" rule would yield an already existing structure, derived from another source under the two-rule analysis. This point is similar to Dougherty's observation with regard to the anaporn relation. Cf. note 12.

Following this analysis, we would expect clefts that derive from preposing to TOPIC to have the same sources as the noncleft analogues. Thus, just as in (91) the preposed constituent is naturally construed with the matrix rather than either embedded clause and presumably is extracted from the matrix clause, so in (92) we have the same interpretations:

- (91) a. *out of spite, I asked the students to refuse to hand in their assignments*  
 b. *only reluctantly did I order the students to refuse to hand in their assignments*  
 c. *only under highly unusual circumstances do I ask students to refuse to hand in assignments*
- (92) a. *it was out of spite that I asked the students to refuse to hand in their assignments*  
 b. *it was only reluctantly that I ordered the students to refuse to hand in their assignments*  
 c. *it is only under highly unusual circumstances that I ask students to refuse to hand in assignments*

In contrast, clefts that derive from topicalization, hence ultimately from *wh*-movement, permit construal with the embedded sentences, as in (87b-e). This difference of behavior is a consequence of the proposed analysis, and provides another reason to suppose that there is no independent rule (or rules) of cleft-formation.

A direct prediction of this analysis is that such pairs as (93a,b) should have the same interpretations:

- (93) a. *only rarely are the students believed to have handed in their assignments on time*  
 b. *it is only rarely that the students are believed to have handed in their assignments on time*

I am not sure that this is correct. It seems to me that (b) may permit construal with the most deeply embedded clause more readily than (a), but my judgments are

quite insecure. If there is a systematic distinction, contrary to the data of (91), (92), then either the analysis is incorrect or there is still another source for clefts or (more plausibly, in my opinion) such distinctions as there may be are to be attributed to the rules of interpretation for cleft and preposing.

Again there are unsolved problems, but it seems to me that it is reasonable to explain the class of cleft sentences that have the properties (49) (e.g., (87) but not (89); cf. (92)) in terms of a rule of *wh*-movement. If the proposal proves tenable, we have still further evidence in support of a positive answer to (50).

Consider next indirect questions. These have the general properties (49), and it seems that a rule of *wh*-movement is involved, analogous to direct questions. I will assume here the general analysis of Chomsky (1973). Thus we have (94):

- (94) a. *I wonder [who John saw]*  
 b. *I wonder [who John believed [that Mary would claim [that Bill would visit]]]*  
 c. *\*I wonder [who John believed [the claim [that Bill would visit]]]*  
 d. *\*who<sub>2</sub> did you wonder [who<sub>1</sub> t<sub>1</sub> saw t<sub>2</sub>]*

As is well known, in the contexts of (95) there can be no lexical NP:

- (95) a. *I wonder [who - to visit]*  
 b. *I wonder [where - to put the book]*  
 c. *I wonder [how - to get to Chicago]*  
 d. *it is unclear [what - to do]*

We might stipulate that in the base rules, NP is required to be *t*(*x*) (i.e., to be NP with variable index, not further specified lexically), our element PRO, in the context (96):

- (96)  $[[\text{COMP}+\text{WH}] [- \text{to VP}]]$

In this context, the value of *x* of *t*(*x*) is determined by a rule of control or NP<sub>*x*</sub> is given the sense: unspecified NP. Presence of PRO invokes the *wh*-constraint, under SSC; in contrast, SSC is inapplicable in the complement of *want*-type verbs (cf. note 4). Perhaps the base condition (96) falls together with other similar rules for "bare" infinitivals, e.g., the *promise-persuade* cases.

Given the stipulation (96), we can add infinitival indirect questions to our list of constructions based on *wh*-movement, with the properties (49), as illustrated in (97), analogous to (94):

- (97) a. *I wonder [who to see]*  
 b1. *I wonder [who to order<sup>32</sup> Mary [to promise [to visit]]]*  
 b2. *I wonder [who to persuade Mary [that she should promise [to visit]]]<sup>33</sup>*  
 c. *\*I wonder [who to insist on [the principle [that Bill should visit]]]*  
 d. *\*who<sub>2</sub> do you wonder [what<sub>1</sub> to give t<sub>2</sub> t<sub>1</sub>]; \*what<sub>2</sub> do you wonder [[to whom]<sub>1</sub> to give t<sub>2</sub> t<sub>1</sub>] (cf.: I wonder (don't remember) [what to give t to whom]; I wonder [to whom to give what t])*

Correspondingly, we have infinitival relatives alongside of the finite relatives, as in (98).<sup>34</sup>

- (98) a. *I found a book* [[ *which for*] *you to read t*] – *I found a book for you to read*  
 b. *I found a man* [[ *to whom for*] PRO *to give the book t*] – *I found a man to whom to give the book*

Infinitival relatives, under this analysis, differ from finite relatives in the rules specifying the surface form of the elements in COMP. Thus in a finite relative corresponding to (98a) we may delete either *which* or the complementizer *that*, giving either (99a) or (99b); or we can delete both, obtaining (99a):

- (99) a. *I found a book which you can read*  
 b. *I found a book that you can read*  
 c. *I found a book you can read*

But in the infinitival relative, the rule (47) deleting *wh-* is obligatory, as in other cases already discussed. Recoverability of deletion prevents it from applying in (98b), just as it cannot apply in (100):

- (100) *I found a man to whom you can give the book* (\**I found a man that you can give the book*)

Thus in (98b) the complementizer *for* must delete, as *that* must delete in (100); we have already remarked that there are rules deleting *for* before *to* (recall that PRO is not terminal).

A further difference between finite and infinitival relatives is that the latter cannot have a lexical NP subject when the complementizer is deleted. Thus we have (98b) but not (101):

- (101) *I found a man* [[ *to whom*] *you to give the book*]

This observation recalls the property of indirect questions captured in (96). Perhaps in place of the base rule (96) we should impose a surface condition excluding phrases of the form (102):

- (102) [COMP *wh*-phrase] NP to VP, where NP is lexical or trace ( $\neq$  PRO)

This will cover the cases excluded by (96) and will also block (101), while permitting (98). It also eliminates the need to make *wh*-phrase deletion obligatory in infinitival relatives (cf. (98), (99)). One might try to generalize (102) to include other phenomena, e.g., the obligatory PRO in infinitival complements of *persuade-promise* type verbs and the heavy restrictions on null complementizers in infinitives at the surface, the surface filters that exclude *for-to* structures, and the rules governing *that*-deletion. I will not pursue these questions here, however. Cf. Chomsky and Lasnik, forthcoming.

The suggested analysis for infinitival relatives seems to me reasonably satisfactory, though the status of (102) remains open along with other questions. Under this analysis, the rule of *wh*-movement extends to all relatives and to both direct and indirect questions, finite or infinitival.

Consider now the sentences (103):

- (103) a. *John found* [NP *a book* [ $\bar{S}$  *which for*] *him to read t*]  
 b. *we found* [NP *books* [ $\bar{S}$  [*which for*] *each other to read t*]]  
 c. \**who*<sub>2</sub> *did he find* [NP *a book* [ $\bar{S}$  [*which*<sub>1</sub> *for*] *t*<sub>2</sub> *to read t*<sub>1</sub>]]

In all three cases, *which* must delete, by the processes just discussed; in (103c), *for* will delete as well, before *to*.

Case (c) is excluded by our conditions, which make the relative clause an island.

But the position marked by *t*<sub>2</sub> in (c) should, on our assumptions, be accessible to interpretive rules, for which the subjacency principle does not hold. Thus in (103a), the rule of disjoint reference (2c) applies, compelling *him* to be distinct in reference from *John*; Similarly, (103b) should be subject to reciprocal interpretation under (2a).<sup>35</sup> On the assumptions we are investigating, bound anaphora (rule (2b)) may also apply in the position of the anaphor (*him*, *each other*, *t*<sub>2</sub>) in (103), giving, e.g., (104), which becomes (105) by EQUI (cf. note 4 and references cited there):

- (104) *John found* [NP *a book* [ $\bar{S}$  [*which for*] *himself to read t*]]  
 (105) *John found a book to read*

The examples (103)–(105), then, illustrate one primary difference between transformational rules and rules of construal, turning on cyclic application and subjacency. Cf. (7), (8), (22), and the discussion in Chomsky (1973).

Infinitival relatives, under this analysis, should have the properties (49). Thus we should find the arrangement of data in (106):

- (106) a. *I found a book for you to read t*  
 b. *I found a book for you to arrange for Mary to tell Bill to give t to Tom*  
 c. *I found a book for you to insist that Bill should read t*  
 d. *I found a book for you to insist that Bill tell Mary that Tom should read t*  
 e. \**I found a book for you to insist on the principle that Tom should read t*  
 f. \**who did he find a book t to read* (= (103c)).<sup>36</sup>

Cases (106c,d) seem to me less acceptable than the comparable examples in the applications of *wh*-movement cited in finite clauses. If this judgment is correct, then the special COMP-COMP movement rule, which permits certain apparent violations of PIC,<sup>37</sup> is less readily available in the case of infinitival relatives.<sup>38</sup> I do not know why this should be so, and am unsure of the judgements. But if (106d) is not acceptable then we really have no argument that the CNPC is in force in (106e), since a demonstration that CNPC is operative requires that analogous cases of comparable complexity with S in place of NP be grammatical. The same question seems to me to arise in other cases of infinitival complements, including (97b2).

Again, it seems to me plausible to extend the rule of *wh*-movement to infinitival relatives as well.

Let us now turn to infinitival complements within the category of adjective phrases.<sup>39</sup> Consider first structures of the form (107), where I assume that  $\bar{S}$  is a complement of the adjective qualifier *enough*.

(107) *John is tall [enough [ $\bar{S}$  for us to see him]]*

Note that although we would normally take *him* in (107) to refer to John, it is not clear that this is necessary, and, in fact, we have such sentences as (108) in which, with the parenthesized material deleted, the complement of *enough* contains no term referring to John:

- (108) a. *John is tall enough for us to be able to see Bill (by standing on his (= John's) shoulders)*  
 b. *John is slow enough for us to win the race (against him (= John))*  
 c. *the car is fast enough for us to win the race (driving it (= the car))*

It seems that (107) can be interpreted as analogous to (108), with the reference of *him* free. If so, then structures such as (107) have essentially the properties of left-dislocation, as described above; that is, we have a focused NP and a proposition that we would normally take to be about this NP, the natural (though not necessary) method being to apply the rule of predication that takes the complement to contain an open proposition satisfied by the referent of the NP, the pronoun taken as a free variable. Assuming that this is the right tack, we may conclude that the base rules generate  $\bar{S}$  freely in such structures as (107).

Alongside of (107) we also have (109), which I assume to derive from (110):

(109) *John is tall [enough [ $\bar{S}$  for us to see ]]*

(110) *John is tall [enough [ $\bar{S}$  [who for us to see t]]]*

The *wh*-phrase in (110) deletes obligatorily, as in comparatives and topicalization. Thus we can have (111) but not (112):

- (111) *John is poor enough for us to give present to*  
 (112) *\*John is poor enough to whom to give presents*

Examples (111) and (112) are analogous, respectively, to (113), (114):

- (113) *I found a person for us to give presents to*  
 (114) *I found a person to whom to give presents*

Note that (114) (derived by EQUI, cf. (105)) is grammatical but not (112), the difference being that *wh*-phrase deletion is not obligatory in the headed relatives; cf. (100).

These examples suggest that the complement of *enough* has a structure analogous to the TOPIC and relative structures described earlier. The complement in this case is infinitival, but, as in the case of TOPIC (and in some languages, relative), it may be either a full sentence with a preference for interpretation as an open sentence, or a *wh*-derived sentence with a free variable in the position marked by trace, which must be interpreted as an open sentence. If so, we would expect to find that

alongside of such structures as (107) (analogous to left-dislocation), we also have *wh*-infinitivals with the properties of (49), except for the obligatory deletion of the *wh*-phrase, already noted; these structures, then, combine the properties of topicalization and those of infinitival relativization. Thus we have (115) analogous to (106):

- (115) a. (i) *John is tall enough for you to see t*  
 (ii) *the job is prestigious enough for us to offer t to John*  
 (iii) *the job is prestigious enough for us to advertise t*  
 b. (i) *John is tall enough for us to arrange for Bill to see t*  
 (ii) *John is famous enough for us to arrange for the committee to offer the job to t*  
 (iii) *the job is prestigious enough for us to arrange for the committee to offer t to John*  
 (iv) *the job is prestigious enough for us to arrange for the committee to advertise t*  
 c. (i) *John is tall enough for us to insist that John (should) pick t for the team*  
 (ii) *John is famous enough for us to insist that you (should) visit t*  
 (iii) *the job is important enough for us to insist that they (should) advertise t*  
 (iv) *the job is important enough for us to insist that they (should) offer t to John*  
 d. (i) *the job is important enough for us to order them to insist that the committee (should) advertise t*  
 (ii) *the job is important enough for us to order them to insist that the committee (should) offer t to John*  
 e. (i) *\*the job is important enough for us to insist on the principle that the committee should advertise t*  
 (ii) *\*the job is important enough for us to insist on the principle that they should offer t to John*  
 f. *\*who<sub>2</sub> was the job good enough for us to offer t<sub>1</sub> to t<sub>2</sub> (etc. as in note 36).*

There is no question that (e) and (f) are excluded, as in (106). Note that in all cases, there is an alternative form, with a pronoun in place of *t* (the analogue of left-dislocation). This alternative form is highly preferred for the (c), (d) cases. We have discussed the analogous observation in connection with infinitival relatives. That is, (106c,d) are also dubious or starred. The (c and d) cases of (115) seem to me still worse than those of (106), which may perhaps be attributed to the fact that in the case of (115), but not (106), there is an alternative form, namely, with a pronoun in place of *t*.

With these provisos, the case of infinitival complements seems to me to be essentially as predicted under the *wh*-movement analysis, namely, as having essentially



the intersection of properties of infinitival relatives (since *wh*-movement is involved) and topicalization (since there is a parallel form without *wh*-movement).

Before we leave this topic, let us consider further the relevant cases of the *wh*-island constraint. Consider the sentences (116), (117):

- (116) a. *the job was good enough* [*for us to offer it to John*]  
 b. *who was the job good enough* [*for us to offer it to t*]  
 c. *to whom was the job good enough* [*for us to offer it t*]
- (117) a. *the job was good enough* [(*which*) *for us to offer t to John*]  
 b. *who<sub>2</sub> was the job good enough* [(*which*<sub>1</sub>) *for us to offer t<sub>1</sub> to t<sub>2</sub>*]  
 c. [*to whom<sub>2</sub>*] *was the job good enough* [(*which*<sub>1</sub>) *for us to offer t<sub>1</sub> t<sub>2</sub>*]

On the assumptions of our analysis, the examples of (116) should all be grammatical (subject to dialect differences with regard to preposition stranding). Similarly, (117a). But (117b,c) should be ruled out by the *wh*-island constraint (ultimately, subjacency and SSC). I think that these conclusions are correct. Problems arise, however, when we try to question the direct rather than the indirect object in such cases as (116). Compare (118), (119):

- (118) a. *John was famous enough* [*for us to offer the job to him*]  
 b. *what job was John famous enough* [*for us to offer t to him*]
- (119) a. *John was famous enough* [(*who*) *for us to offer the job to t*]  
 b. *what<sub>2</sub> job was John famous enough* [(*who*<sub>1</sub>) *for us to offer t<sub>1</sub> to t<sub>2</sub>*]

As expected, (119b) is ungrammatical. But (118b) ought to be grammatical, under our assumptions. It does not seem to be, however. The status of (116b,c) is also unclear. One can imagine a formulation of bridge conditions that would rule out all of these examples, or assign them a marginal status, analogous to (42).

Summarizing, it seems to me that the *wh*-movement analysis gives a reasonably good first approximation in this case, though some problems concerning infinitival clauses remain. I know of no problems specific to this analysis.

Other complements of adjective qualifiers, as in (120), have about the same properties as the complements of *enough*, so far as I can see, so I will have nothing to say about these:

- (120) *Muhammad Ali is too good* [(*who*) *for Bill to arrange for John to fight t*]

The final case I would like to consider is that of the infinitival complements of *easy*, etc. The analysis proposed in Chomsky (1973) was unsatisfactory, as pointed out by Sterba (1972), Lasnik and Fiengo (1974), and Bach and Horn (1976). With regard to such structures as (121) there have been two widely studied proposals:

- (121) *John is easy (for us) [to please]*

One proposal assumes that the subject, *John*, is moved from the object position in the embedded complement phrase by a transformational movement rule. The other assumes that the subject is generated in place and that a rule of object-deletion (or interpretation) guarantees that *John* is interpreted as the object of *please* in (121).

I will not try to survey the arguments here. Rather, let us take a fresh look within the present framework.

I will assume that the phrase *for us* in (121) is, as indicated, generated in the matrix sentence. Cf. Bresnan (1971), Chomsky (1973), Lasnik and Fiengo (1974), and Brame (1975). If so, then according to our present assumptions, the underlying structure must contain an embedded  $\bar{S}$  as complement to *easy*, with an obligatory PRO subject, as in the case of the infinitival complements already mentioned.<sup>40</sup> In some similar structures the *for*-phrase appears in both the matrix and embedded sentence, as in (122):

- (122) a. *it is a waste of time for us* [*for them to teach us Latin*]  
 b. *it is pleasant for the rich* [*for the poor to do the hard work*]

And there are, of course, adjectival complements of various sorts that exhibit the full infinitival construction, e.g., (123):<sup>41</sup>

- (123) a. *John is eager* [*for Bill to leave*]  
 b. *John would be happy* [*for Bill to win*]  
 c. *the house is ready* [*for John to buy (it)*]

On the assumption that the complement clause in (121) is essentially the same as those in (122), (123), we may take the underlying structure for (121) to be essentially (124), though nothing much depends on the choice of complementizer, it seems:

- (124) *X is easy (for us) [ $\bar{S}$  for PRO to please Y]*

The complementizer *for* will then delete before *to*, as in cases discussed above, e.g., (125):<sup>42</sup>

- (125) a. *who does John want very much (for) to win*  
 b. *he is the man who John wants most of all (for) to win*

Assuming this much, we now face the question: what are *X* and *Y* in (124)?

Our assumptions lead us to suppose that each of the competing familiar analyses is in part correct: that is, *X* = *John*—the subject is generated in place—but there is a movement rule applying to *Y*, namely, *wh*-movement.<sup>43</sup> Thus we may take the structure directly underlying (121) to be (126):

- (126) *John is easy (for us) [ $\bar{S}$  [*who for*] PRO to please t]*

In (126), *wh*-movement has applied on the inner cycle and we have obligatory deletion of the *wh*-phrase, as in other cases already discussed. We are left, then, with an open embedded proposition; the now familiar predication rule will correctly interpret it as being about the subject *John*.

We then expect to have, again, the properties (49), as in the infinitival relatives and related constructions. Thus we have (127) corresponding to (106):

- (127) a. *John is easy (for us) to please t*  
 b. (i) *John is easy (for us) to convince Bill to do business with t*  
 (ii) *John is easy (for us) to convince Bill to arrange for Mary to meet t*