

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 ($\langle abccd \rangle_S$, *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.

76

In Formal Syntax, ed. by P. Culicover, A. Akmajian, and T. Wasno. Academic Press, N.Y. (1977), pp. 71-133.

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_c , 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.¹ 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 α and β are cyclic nodes

For the present, I will take the cyclic nodes to be \bar{S} and NP; 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 α is a cyclic node:

(11) ... X ... [α ... 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 α is a finite clause (tensed-S). SSC asserts that no rule can "involve" X and Y where α 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 α 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 α_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 α_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 α 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 α 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 α 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 α_i, α_{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 [+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 subadjacency 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*.² 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 α 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 α 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 [+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 ν 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 α 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]*³

- 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 to each other]*

Disjoint reference:

- a. PIC: (i) *they want [them to win] (they \neq them)*
 (ii) *they prefer [that they win]*⁴
- 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]*⁵

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]*⁶

Clitic movement :⁷

- a. PIC: From infinitives, but not tensed clauses, by PIC⁸
- 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 ⁹

- 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).¹⁰ 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]]¹¹
 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,¹² 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 in 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 of 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) $[_S [_{NP_i} \textit{John}] [_{VP} \textit{beten kill} [_{NP_j} \textit{Bill}] \textit{by} [_{NP_k} \textit{e}]]]]$

The rule of NP postposing moves NP_i , replacing the terminal identity element e , in NP_k . It is natural to assume that the moved NP, *John*, retains its index, so that in place of NP_k , we have NP_i 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) $[_S [_{NP_i} \textit{e}] [_{VP} \textit{beten kill} [_{NP_j} \textit{Bill}] \textit{by} [_{NP_i} \textit{John}]]]]$

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

(29) $[_S [_{NP_j} \textit{Bill}] [_{VP} \textit{beten kill} [_{NP_j} \textit{e}] \textit{by} [_{NP_i} \textit{John}]]]]$

We may now define the substructure $[_{NP_i} \textit{e}]$ of (28) as the “trace” of NP_i (= $[_{NP_i} \textit{John}]$), and represent it by convention as $t(i)$ (read: “trace of NP_i ”). Similarly, the substructure $[_{NP_j} \textit{e}]$ of (29) is the trace of NP_j , represented as $t(j)$.¹⁴ 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_x , an NP without a fixed index. The index is then assigned by a rule of control. E.g., in (20), if $v = \textit{persuade}$ and *the man* is NP_i , then PRO will become NP_i ; and if $v = \textit{promise}$ and I is NP_j , then PRO will become NP_j . 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*¹⁵

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:¹⁶

(38) Given an \bar{S} of the form:

[_{COMP} --[*wh*- \bar{N}]-[+WH]] [_S ... t ...]
 where t is the trace of [*wh*- \bar{N}], rewrite it as:
 [_{COMP} for which x , x an \bar{N}], [_S ... [- x -] ...]

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,¹⁷ 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):¹⁸

- (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, *vbl*), 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.¹⁹

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 subadjacency, 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 subadjacency (implying (c) and (d)²⁰), 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)21). 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.²² 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) $[\bar{S} [\text{TOP } \textit{this book}] [\bar{S} [\text{COMP } \textit{what}] [\textit{I asked Bill to get his students to read t}]]]$
 (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} [$_{TOP}$ *this book*] [\bar{S} [$_{COMP}$ *what x*] [*I asked Bill to get his students to read x*]]]

Deletion of the *wh*-phrase leaves an open sentence,²⁵ 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} [$_{TOP}$ *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} [$_{TOP}$ *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).²⁶ 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.²⁷ 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*²⁸; *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*.²⁹

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*³⁰
 b. *(purely) out of spite, he assigned it that number*
 c. *only reluctantly he agreed to swim at all*³⁰

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.³¹ 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. **who₂ did you wonder [who₁ t₁ saw t₂]*

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) [[COMP+WH] [- to VP]]

In this context, the value of *x* of *t(x)* is determined by a rule of control or NP_x 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 order³² Mary [to promise [to visit]]]*
 b2. *I wonder [who to persuade Mary [that she should promise [to visit]]]³³*
 c. **I wonder [who to insist on [the principle [that Bill should visit]]]*
 d. **who₂ do you wonder [what₁ to give t₁ to t₂]; *what₂ do you wonder [[to whom]₁ to give t₂ t₁] (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):³⁴

- (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 (≠ 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* [_S *which for*] *him to read t*]]
 b. *we found* [_{NP} *books* [_S [*which for*] *each other to read t*]]
 c. **who*₂ *did he find* [_{NP} *a book* [_S [*which*₁ *for*] *t*₂ *to read t*₁]]

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*₂ in (c) should, on our assumptions, be accessible to interpretive rules, for which the subadjacency 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).³⁵ On the assumptions we are investigating, bound anaphora (rule (2b)) may also apply in the position of the anaphor (*him*, *each other*, *t*₂) in (103), giving, e.g., (104), which becomes (105) by EQUI (cf. note 4 and references cited there):

- (104) *John found* [_{NP} *a book* [_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 subadjacency. 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)).³⁶

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,³⁷ is less readily available in the case of infinitival relatives.³⁸ 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.³⁹ 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. **who₂ was the job good enough for us to offer t₁ to t₂ (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. *who₂ was the job good enough* [(*which*₁) *for us to offer t₁ to t₂*]
 c. [*to whom*₂] *was the job good enough* [(*which*₁) *for us to offer t₁ t₂*]

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. *what₂ job was John famous enough* [(*who*₁) *for us to offer t₁ to t₂*]

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.⁴⁰ 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):⁴¹

- (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):⁴²

- (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.⁴³ 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*

- c. *John is easy (for us) to convince Bill that he should meet t*
 d. *John is easy (for us) to convince Bill to tell Mary that Tom should meet t*
 e. (i) **John is easy (for us) to convince Bill of the need for him to meet*
 (ii) **John is easy (for us) to describe to Bill a plan to assassinate t*
 f. (i) **what₂ is John fun (for us) [(who₁) to give t₂ to t₁]* (from a source like: *John is fun (for us) to give presents to*
 (ii) **who₂ are the presents fun (for us) [(which₁) to give t₁ to t₂]* (compare: *the presents are fun (for us) to give to him*
 (iii) **[to whom]₂ are the presents fun (for us) [(which₁) to give t₁ t₂]* (compare: *the presents are fun (for us) to give to him*)

As in other cases discussed, cases (c) and (d) are marginal.

In short, the basic properties of *easy-to-please* constructions follow directly from the assumptions we have already made, assuming that here too *wh*-movement is crucially involved. The latter assumption is particularly natural in this case, since we have analogous forms in which the *wh*-phrase may directly appear. Thus following our analysis, (121) is analogous to (128), and in such cases, we may have the full *wh*-phrase, as in (129)⁴⁴:

- (128) *John is an easy person to please*
 (129) a. *this is an easy violin on which to play sonatas*
 b. *this is a pleasant room in which to work*

Whatever the correct analysis of these structures may be, it seems clear that they involve, at some level, a phrase such as (130), as an adjectival modifier:

- (130) a. *easy – on which to play sonatas (violin)*
 b. *pleasant – in which to work (room)*

Our analysis simply assumes that the same is true quite generally of *easy*-complements. In the case of (130), the structures are embedded (presumably, in some manner, as relatives) within an NP with a head; in the case of (121), there is no NP antecedent and the structure must be interpreted as an open sentence, as in topicalization and other examples discussed above. Thus *wh*-deletion is obligatory, as in the other cases discussed, and forms analogous to (130) do not appear in the *easy-to-please* structures, just as we do not have (112), etc.

It should follow that in general, *easy-to-please* constructions have the relevant properties of *wh*-movement. Parallels have been observed in the literature. E.g., Lasnik and Fiengo (1974) note such parallels as (131)⁴⁵

- (131) a. *what did you give to John*
 b. **who did you give a book*
 c. *who did you give a book to*
 d. *John is dumb enough to sell the Brooklyn Bridge to*
 e. **John is dumb enough to sell the Brooklyn Bridge*
 f. *John is easy (for us) to sell the Brooklyn Bridge to*
 g. **John is easy (for us) to sell the Brooklyn Bridge*

Notice that the cases (127f) are exactly analogous to other examples of *wh*-island constraints, on this analysis, e.g., as in (132) and many examples already cited:

- (132) a. **who₂ do you wonder [what₁ t₂ saw t₁]*
 b. **I wonder [who₂, this book, [(which₁) t₂ really likes t₁]]*
 c. **who₂ is John more friendly to Mary than [(what₁) he is t₁ to t₂]*

In all of these cases, the sentences are ruled out on the assumption that *wh*-movement has taken place, by the *wh*-island constraint, which, as noted, follows from the conditions postulated. In the form immediately underlying case (132a) there is a residual *wh*-phrase indicating that *wh*-movement has taken place; in the case of (127f), (132b,c), and many others discussed above, there is no such residual phrase, but the effects of *wh*-movement are still evident.

There is a well-known puzzle concerning application of *wh*-movement to the sentences (133):

- (133) a. *the sonata is easy to play on this violin*
 b. *the violin is easy to play sonatas on*

Consider first (b). The phrase *sonatas* appears to be in a position susceptible to *wh*-movement; compare (134):

- (134) a. *John was told to play sonatas on his violin*
 b. *what was John told to play on his violin*

But in (133b), *wh*-movement is impossible. We cannot have (135):

- (135) a. **what sonatas is this violin easy to play on*
 b. **the sonatas that this violin is easy to play on—are in your book*

We now have an explanation for this fact. In terms of our analysis, *sonatas* in (133b) is within a *wh*-island, just as *t₂* is within a *wh*-island in the topicalization and comparatives of (132). The structure to which *wh*-movement must apply to give (135) is (136):

- (136) *this violin is easy [_S(which) for PRO to play sonatas on t]*

But *sonatas* in (136) is not subject to *wh*-movement because of SSC, as in the cases discussed earlier. Consequently, the examples of (135) are ruled out. While superficially (133b) is analogous to (134a), in the mental computation underlying (133b) there is, we now assume, a *wh*-phrase blocking the application of *wh*-movement.

Consider now (133a). Suppose that we apply *wh*-movement to *this violin*. The result is (137):

- (137) a. *what violin is the sonata easy to play on*
 b. *the violins that the sonatas are easy to play on—are being repaired*

Many speakers find these acceptable, in contrast to (135), which are universally rejected. By our analysis, the underlying structure for (133a) is (138), which should be immune to *wh*-movement just as (136) is:

- (138) *the sonata is easy [_S(which) for PRO to play t on this violin]*

why, then, should the examples (137) have a different status, for some speakers, than those of (135)?

Notice that in other contexts, the embedded \bar{S} s of (136), (138) are, as expected, both immune to *wh*-movement. Consider (139a,b), with the same embedded sentences as (136), (138), respectively:

- (139) a. *you found a violin* [\bar{S} (*which*) *for PRO to play sonatas on t*]
 b. *you found a sonata* [\bar{S} (*which*) *for PRO to play t on this violin*]

Application of *wh*-movement gives (140), impossible in both cases:

- (140) a. **what sonata did you find a violin to play on*
 b. **what violin did you find a sonata to play on*

Example (140a) is analogous to (135); example (140b) is analogous to (137). Comparing these cases, we see that it is the acceptability of (137) (for some speakers) that is the exceptional case, somehow to be explained.

A possible explanation is that there is another structure underlying (133a), namely, (141), where the PP *on this violin* is associated with the VP rather than the adjective phrase:

- (141) *the sonata is* [$_{AP}$ *easy* [\bar{S} (*which*) *for PRO to play t*]] *on this violin*

If (141) is taken to underlie (133a), under one option, then (137) will be derivable by *wh*-movement. No such alternative analysis is possible in the case of (133b), (139). Therefore, no *wh*-movement is possible in these cases.

If this is the correct explanation, then we should find that in forms analogous to (133a) but where the PP is not separable from the embedded verb, forms analogous to (137) should be on a par with (135) rather than (137).⁴⁶ Compare (142), (143):

- (142) *the book is easy* [\bar{S} *for PRO to put t on the table*]

- (143) a. *what table is the book easy to put on*
 b. *the table that the book is easy to put on*

It seems to me that the prediction holds; that is, the examples (143) are excluded, in contrast to (137). The question deserves fuller investigation, but in a large class of rather puzzling cases it seems that we have an explanation for the facts in terms of a *wh*-movement analysis, given the framework of conditions and rules outlined earlier.

Consider next the examples (144):

- (144) a. *it is a waste of time for us* [*for them to teach us Latin*] (= (122a))
 b. **Latin is a waste of time for us* [*for them to teach us*]

In Chomsky (1973) the distinction was explained in terms of conditions on rule application, but that approach is ruled out in the present analysis. The correct explanation for the ungrammaticality of (144b), I think, lies in a base condition. In the underlying structure (145), the subject of the embedded infinitival must be PRO, as in the *persuade-promise* cases and others that we have discussed:

- (145) NP *is* Predicate (*for us*) [\bar{S} *for* – *to VP*]

The conditions on NP and Predicate in (145) must be specified in the base. Note that where NP is a dummy element *it* (however this is introduced), there is no constraint on –; cf. (122). The restriction to PRO applies only when the matrix NP subject is lexically specified. Furthermore, in such examples as (123a,b), where *wh*-movement is excluded in the embedded clause, there is also no constraint on – in (145). Thus we seem to have either the base condition (146a) or (146b):

- (146) a. in (145), – is PRO if \bar{S} is subject to *wh*-movement
 b. in (145), – is PRO if \bar{S} is obligatorily subject to *wh*-movement.

Of these two conditions, (146a) is preferable, if it is tenable; it is more general and can, I believe, be reformulated so as to fall together with other cases with obligatory PRO subject under generalizations relating *it* and choice of complementizer. It seems to cover all cases except for (123c).

The argument that the *for*-phrase in (123c) is within the complement offered by Bach and Horn (1976) does not seem to me entirely compelling. They note that the *for*-phrase is not preposable in (147), though it normally is when part of the matrix. Thus we have (148) but not (149):

- (147) *the house is ready for John to buy*
 (148) *for the rich, it is pleasant for the poor to do the hard work*
 (148) **for John, the house is ready to buy*

But this argument seems inconclusive, since even in the case of (149), where *for John* is surely a PP of the matrix, it is not preposable, for some reason:

- (149) *the house is ready for John*
 (150) **for John, the house is ready* (cf. *for John, the problem was easy*)

They give supplementary arguments in terms of right node raising and gapping, arguing that (151), (152) are acceptable:

- (151) *the moussaka is ready and Mike says that the egg-lemon soup is almost ready – for us to eat*
 (152) *the kidney pie is ready for us to put in the oven, and the salad – for you to put on the table*

Assuming that the *for*-phrases in *easy*-structures are in the matrix, a point that they do not contest, the strength of these arguments depends on the distinction between (151), (152) and (153), (154):

- (153) *young children are quite difficult, and Bill says that older children are still more difficult – for untrained teachers to control*
 (154) *the young children are difficult for Bill to control, and the older children – for Mary to teach*

I am not convinced that there is any relevant difference. Consequently, it is possible that the *for*-phrase associated with *ready* is also in the matrix sentence where the

complement is subject to *wh*-movement, contrary to (123c). If so, then (146a) may be the correct principle.

Whichever case of (146) holds, (144b) is ruled ungrammatical on the grounds that it requires a base form not generated by base rules (or a corresponding surface condition).

Notice that if the subject NP of the complement in (145) is PRO, then it can never be assigned *wh*- or moved by *wh*-movement. Thus it follows that the rule applying to *easy-to-please* structures is limited to an NP in the embedded predicate.

Consider again the form (126), repeated here as (155), underlying (121):

(155) *John is easy (for us) [\bar{s} [who for] PRO to please t]*

Suppose that *wh*-movement were to apply to (155), as in the COMP-COMP case of *wh*-movement, giving (156):

(156) *who is John easy (for us) to please*

Plainly (156) is ungrammatical. We might account for this fact by rule-ordering, i.e., requiring that the obligatory deletion of *who* precede *wh*-movement on the matrix cycle. But there is in fact a simpler approach that requires no such stipulation. Thus note that the resulting structure corresponding to (156) is (157), after interpretation of the *wh*-quantifier, in contrast to (158), underlying (121):

(157) *for which x, x a person, John is easy (for us) [for PRO to please x]*

(158) *John is easy (for us) [for PRO to please x]*

We have assumed that (158) is interpreted by the general rule of predication described for topicalization and other forms, with an open proposition taken to be satisfied by the referent of the focused NP, in this case, the matrix subject. But the rule of predication is inapplicable to (157), since there is no open proposition: the variable *x* is bound in (157) by the quantifier "for which *x*." Thus the sentence is uninterpretable, just as "John is easy to please Bill" is uninterpretable. This seems a natural way to account for the ungrammaticality of (156).

Some might object that (156) must be excluded as ungrammatical on syntactic grounds rather than on grounds of uninterpretability. I have argued elsewhere that, whereas speakers can make judgments of acceptability, they have no direct access to the grounds of these judgments. Thus I have no intuitive insight into the source of the unacceptability of (156). Only if these acceptability judgments come marked as "syntactic," "semantic," etc., can the objection be sustained. It seems to me that there is no merit to the contention.

Suppose that in fact convincing arguments can be given that in (123c) the *for*-phrase is embedded even where *wh*-movement takes place in the embedded clause, so that we have the underlying structure (159), where either subject or object of the embedded clause is accessible to *wh*-movement:

(159) *the house is ready [for NP to buy NP]*

Applying *wh*-movement to the object, we derive (160):

(160) *the house is ready [(which) for John to buy]*

Application of subsequent *wh*-movement to an NP in the position of *John* is impossible for familiar reasons.

Suppose that we apply *wh*-movement to the embedded subject of a structure like (159), obtaining (161):

(161) *the house is ready [[which for] t to fall down]*

With obligatory deletion of *which* followed by *for*-deletion before *to*, we derive (162):

(162) *the house is ready to fall down*

If, in contrast, applicability of *wh*-movement to the embedded clause is taken to correlate with PRO subject, as in (146a), then (162) would derive only from (163) by EQUI, just as (165) derives from (164):

(163) *the house is ready [for itself to fall down]*

(164) *John is eager [for himself to please]*

(165) *John is eager to please*

Assuming that we do derive (161), consider the effect of applying the COMP-COMP rule of *wh*-movement to give (166):

(166) **what is the house ready to fall down*

But this is ungrammatical on the same grounds that rule out (156). Thus nothing much seems to depend on where the *for*-phrase appears in (123c), apart from the generality of the base principle or corresponding surface filter.

Other structures similar to (121) are much more restricted in scope, e.g., (167):

(167) *Mary is pretty to look at*

In this case, we do not have the full range of properties (49). Thus there is no form (168), analogous to (127b):

(168) *Mary is pretty to tell Bill to look at*

Furthermore, in such structures as (167) there are very narrow restrictions on the choice of the matrix adjective and embedded verb. We may propose the same analysis as in the *easy-to-please* cases, but with idiom interpretation rules associated with the adjectives in question. Note that there are structures such as (169), but in this case the embedded complement is not associated with the adjective but with the adjective qualifier, *too*:

(169) *Mary is too pretty to expect anyone to look at (her)*

As has long been known, structures of the *easy-to-please* type do not appear as nominals, in contrast to the superficially similar forms with *eager*: compare (170):

- (170) a. *John's eagerness to please – surprised me*
 b. **John's difficulty to please – surprised me*

Various explanations have been proposed, relying on particular analyses of movement or deletion in the *easy* cases. Under the assumptions of EST, including the lexicalist hypothesis, the distinction between (170a) and (170b) must be formulable without reference to ordering of transformations and the like, on the assumption that *eagerness*, *difficulty*, etc., are drawn from the lexicon.⁴⁷ On our assumptions, the NPs of (170) have the underlying structures (171a,b), respectively:

- (171) a. [_{NP} *John's* [_N *eagerness* [_S *for himself to please*]]]
 b. [_{NP} *John's* [_N *difficulty* [_S *(who) for PRO to please t*]]]

The form (171a) is analogous in structure to nouns with sentential complements, as in (172):

- (172) a. *John's certainty that Bill will leave*
 b. *John's desire for Bill to leave*
 c. *the fact that Bill left*

In contrast, (171b) has the formal structure of a relative, as in (173):

- (173) a. *the certainty that you feel*
 b. *the desire (for Bill to leave) that you expressed*
 c. *the fact that Bill cited*

But the rule of interpretation for relatives plainly cannot apply in (171b), any more than it can in (174):

- (174) a. *the eagerness [(who) (for Bill) to visit t]*
 b. *the certainty [(who) that Bill will visit t]*
 c. *the desire [(who) for Bill to visit t]*
 d. *the fact [(who) that Bill visited t]*

More precisely, if the rule of relative interpretation were to apply in these cases, it would take the relative to hold of the head, as in (175):

- (175) a. *a book [(which) for you to read]*
 b. *the book [(which) that you read]*

Cf. the discussion of relativization above. This interpretation is senseless in (171b); furthermore, infinitival relatives (or relatives altogether) do not occur in general with such determiners. Thus expressions such as (171b) are ungrammatical. Perhaps this is the explanation for the absence of derived nominals corresponding to the forms of (170b). We might proceed further, in terms of the X-bar system, to assign sentential complements of nouns, which are immune to *wh*-movement (or, perhaps, to relative interpretation), a different position in the hierarchy than relatives. Cf. Jackendoff (forthcoming) for some suggestions.⁴⁸

To summarize, I have suggested that we can eliminate from the grammar rules of comparative deletion, topicalization, clefting, object-deletion and "tough movement," rules for adjective and adjective-qualifier complements, and others, in favor of the general rule of *wh*-movement that also yields direct and indirect questions (finite and infinitival) and finite and infinitival relative clauses, several rather general rules of interpretation, and some language-specific properties of base and surface

structures. If this analysis proves tenable, we can drastically reduce the grammatical apparatus for the description of English; but more important, we can drastically limit the class of possible rules. Some curious and otherwise unexplained phenomena fall into place quite naturally, under this simplification of grammatical theory and the description of English. The properties (49), which appear (with the provisos noted) in a wide range of cases, fall together naturally, as a consequence of independent and, I think, rather natural conditions on rules: the subjacency condition, which in effect limits the "memory" available to transformational rules; SSC, which selects a most prominent NP in an embedded cyclic category that is alone accessible to rules if it is present; and PIC, which immunizes a certain category of propositions from rule application, subject to the language-specific proviso that permits COMP-COMP movement over a "bridge." Each of these conditions may be thought of as a limitation on the scope of the processes of mental computation that ultimately determine phonetic and logical form.

This discussion provides evidence in support of a positive answer to the query (50), and specifically, to the thesis that the phenomena that fall under CNPC and the *wh*-island constraints are to be explained in terms of more general properties of rules. But the evidence does not (and could not) suffice to establish the thesis, even if everything suggested here proves to be correct. This is, it remains an open possibility that some of the phenomena that fall under these constraints must be explained in other terms. Before turning to this question, I want to consider the effects of some modifications of the conditions discussed at the outset.

In our formulation of the basic conditions on rules, the notion "cyclic node" plays a crucial role. The cyclic nodes were taken to be NP and S (and perhaps \bar{S}) in the foregoing discussion. Suppose that we were to add S to the category of cyclic nodes. A slight reformulation of PIC is then required, but it is otherwise unaffected. There are interesting consequences in the case of SSC and subjacency, however.

Consider the effect on SSC. Given a structure of the form (176), no rule can now involve *X* and *Y* if *S* contains a subject not containing *Y* and not controlled by *X*:

- (176) ... *X* ... [*S* ... *Y* ...] ... *X* ...

Suppose in particular that *Y* is NP. Then a rule such as *wh*-movement, extracting an NP to the COMP position *X* outside of *S*, can apply to *Y* only if *Y* is the subject of *S*. In general, only subjects are accessible to movement rules involving an element outside of *S*, on this interpretation of SSC. It is well known that in many languages only subjects are accessible to many rules. Cf. Ross (1972); Keenan and Comrie (1973). Perhaps this fact can be explained by a modification of SSC for such languages in the manner just suggested. Note that if such a language also has COMP-to-COMP movement, the effect will be that only the subject of a subject sentence will be accessible to rules. For an apparent example, see Bell (1976).

The effect of incorporating *S* among the cyclic nodes is more far-reaching in the case of subjacency. It now follows that in a structure of the form (177), *wh*-movement cannot extract *Y* to COMP:

(177) [\bar{S} COMP [$S \dots$ [$NP \dots Y \dots$] \dots]]

In particular, it follows that *wh*-movement cannot extract anything from the subject of a sentence. Since the earliest work on transformational grammar, it has been clear that *wh*-movement must somehow be restricted in this fashion. E.g., it is noted in Chomsky (1955) that the rule of *wh*-movement must be prevented from applying to (178), to give (179):

(178) [*your interest in him*] *seemed to me rather strange*

(179) **whom did* [*your interest in*] *seem to me rather strange*

In the earliest work, it was assumed that the structural description of the rule must be designed to exclude this possibility. Later, general conditions were proposed on the functioning of rules, e.g., the Subject Condition of Chomsky (1973).⁵⁰ The Subject Condition follows at once from subjacency, when *S* is taken to be a cyclic node.

Of course, it follows as well that *wh*-movement cannot extract a phrase from a nonsubject NP, as in (180):

(180) *who did you see* [*a picture of t*]

But the sentence (180) is grammatical. It was for this reason that subjacency was not extended to include *S* in Chomsky (1973). We return to this problem directly. Let us assume that it can be overcome and that subjacency is correctly formulated with *S* as one of the cyclic nodes.

As a consequence of this decision, we now have the general property (181):

(181) In the structure (177), *Y* cannot be extracted from *S*; in particular, *wh*-movement cannot move *Y* to COMP.

If the general approach sketched earlier proves tenable, then perhaps the special case of (180) is the only case.

Notice that nothing prevents extraction of *Y* outside of NP within *S*, in (177). Thus there is now no barrier against the rules indicated in (182):

(182) a. [\bar{S} COMP [S [NP *a review t*₁] *was published* [of $Bill's$ $book$]₁]]
 b. [\bar{S} COMP [S [of the $students$ in the $class$]₁] [NP *several t*₁] *failed the exam*]]

Whether (182b) is the correct surface structure may be questioned. Note that extraction of PP as in (182b) is incompatible with *wh*-movement:

(183) a. **of the students in the class, which exam did several fail*
 **which exam, of the students in the class, did several fail*
 **I don't know which exam, of the students in the class, several failed*⁵¹

The impossibility of (183a) follows directly from the suggested analysis, but not that of (183b,c). These examples suggest that the PP is extraposed to COMP, contradicting our assumption, or perhaps that the PP is adjoined to *S*, creating a new *S*-category in the usual way, so that subjacency blocks *wh*-movement. In

support of the latter alternative (or 182b)) are such structures as (184):

(184) *I told Mary that of the students in the class, several will fail*

Let us suppose tentatively that (182b) is correct in essence, assuming that the problem posed by (183b,c) can be overcome as suggested. Note that extraction of phrases from the subject, as in (182), contradicts the Subject Condition of Chomsky (1973), as noted by Postal (1974a). But it is compatible with the reformulation of this condition in terms of subjacency, which of course has the added advantage of eliminating a rather ad hoc condition. Let us tentatively assume, then, that the Subject Condition is dropped in favor of subjacency as just amended. Cf. note 50.

A different approach to these questions is suggested by Bach and Horn (forthcoming). They propose a general constraint that they formulate as follows:

(185) The NP Constraint. No constituent which is dominated by NP can be moved or deleted from that NP by a transformational rule [apart from free deletions, if such exist].

The NP Constraint differs in its empirical consequences from the modified subjacency condition in that it excludes all movement from NP, whereas the subjacency condition excludes only those movement rules that extract an element from *S* as well as NP; just *wh*-movement, if the foregoing analysis is correct.

The NP Constraint is immediately falsified by such examples as (182).⁵² In fact, if the foregoing analysis is correct, the apparent generality of (185) is illusory: the only rule subject to it is *wh*-movement, which is also the only rule extracting a constituent dominated by NP from *S* as well as NP. All other extraction rules, it seems without exception, apply freely to subparts of NPs, as do all interpretive rules (subject to SSC, of course, as in *"*we read* [*Bill's stories about each other*],"
 "*they read* [*Bill's stories about them*]" with coreference of *they*, *them*).⁵³ The unique status of *wh*-movement from NPs is exactly what is captured by the analysis in terms of subjacency, since only this rule extracts a phrase not only from NP but also from *S* (on the assumptions of the foregoing analysis).

Let us now turn to the remaining problem, namely, *wh*-movement from nonsubject NPs, as in (180). Bach and Horn argue, very plausibly I believe, that the interrogative (186) derives from (187), with the structure as indicated, rather than from (188) (see also Cattell, 1976):

(186) *who did John write a book about*

(187) *John wrote* [NP *a book*] [*about who*]

(188) *John wrote* [NP *a book about who*]

They argue further that (187) is base-generated alongside (188), as shown by the fact that we can have such sentences as (189) and by the unambiguous interpretation of (190a) as compared with the ambiguity of (190b,c):

(189) a. *John wrote it about Nixon*

b. *a book was written about Nixon by John*⁵⁴

- (190) a. *John destroyed [his first 5 books about Nixon], in 1965*
 b. *John wrote [his first 5 books about Nixon], in 1965*
 c. *John wrote [his first 5 books]about Nixon, in 1965*

Correspondingly, on their assumptions, we can question "a book" in (187), obtaining (191), but we cannot form (193) from (192):

- (191) *what did John write about Nixon*
 (192) *John destroyed a book about Nixon*
 (193) a. **who did John destroy a book about* (cf. (186))
 b. **a book was destroyed about Nixon by John* (cf. (189b))
 c. **what did John destroy about Nixon* (cf. (191))

Suppose that we follow Bach and Horn in assuming that when *wh*-movement has taken place in nonsubject position, it has not extracted from inside an NP but rather from a PP that is not dominated by NP, but directly by VP, as in (187). This eliminates the remaining problem in the formulation of subjacency suggested above.

It remains to determine how structures of the form (194), which are subject to *wh*-movement of each NP, are derived:

- (194) COMP NP [_{VP} V NP [P NP]]

Bach and Horn assume that all of these are base-generated. The contention is plausible in the special case of (187), where we have the corresponding pronominal form (189a), but not, I believe, in many other cases, e.g., (180) or many such cases as (195):

- (195) a. *who did he find a picture of t*
 b. *what books did he write reviews of t*

In these cases we cannot have forms corresponding to (189). Thus:

- (196) a. **he saw it of John*
 b. **he found it of John*
 c. **he wrote them of three novels*

But in these cases we can question the NP in the PP. Thus the properties that Bach and Horn consider do not correlate, contrary to what they assume. Base-generability seems to me plausible only in such cases as (187), where "write a book" is treated virtually as a verb, and in fact possessive determiners are impossible; see below; also (vi), (vii) of note 10.

Departing now from Bach and Horn's analysis, suppose that we postulate a rule of extraposition from NP to give (198) from (197), perhaps related to the familiar rule (cf. (21)), though more likely, a kind of "readjustment rule."

- (197) *he saw [_{NP} a picture [_{PP} of John]]*
 (198) *he saw [_{NP} a picture t] [_{PP} of John]*

The conditions on the choice of the matrix verb are obscure; thus the rule can apply to *see*, *find*, but not *destroy*; There appears to be some vacillation and

disagreement in informant judgment on this matter, as one might expect in the case of a marginal rule such as this.

The extraposition rule forming (198) produces a structure just like the base-generated structures, apart from the trace *t* in (198). We can at one explain the impossibility of pronouns in the NP position of (198), (196); these are not base-generated structures. For the same reason, we cannot have (199):

- (199) *what did he see of John*

Application of *wh*-movement and passive to (198) gives the forms (200):

- (200) a. *who did he see a picture of t*
 b. [*what picture t*] *did he see [of John]*
 c. [*a picture t*] *was seen [of John]*

The status of (b) and (c) is obscure; cf. note 54. Pending further investigation, I will put them aside.

We now have the following three cases, with the deep structures indicated:

- (201) *he took [_{NP} a picture] [_{PP} of John]*
 (202) *he destroyed [_{NP} a picture of John]*
 (203) *He saw [_{NP} a picture of John]*

The lexically governed extraposition rule gives (204) from (203), but does not apply to (202):

- (204) *he saw [_{NP} a picture t] [_{PP} of John]*

The cases are differentiated in the following way:

- (205) a. Possibility of pronoun in place of *a picture*: (201) but not (202), (203-4)
 b. Applicability of *wh*-movement to *John*: (201), (204) but not (202)
 c. Possibility of a possessive NP in place of *a*: (202), (203-4) but not (201)

Notice that we cannot have (206):

- (206) **who did he see Bill's picture of t*

The reason is that extraposition from NP is impossible in (207) because of SSC (cf. (21)).

- (207) *he saw [Bill's picture of John]*

Since extraposition from NP is impossible in this case, subjacency (and also SSC) will prevent *wh*-movement; hence (206).

Since possessives are in any event impossible in the quasi-idiomatic case (201), we do not have (208):

- (208) a. **who did he take Bill's picture of*
 b. **who did he write Bill's book about*

Bach and Horn argue that the forms underlying (208) are blocked by base rules. But their analysis does not extend to case (203) (*see*, *find*, etc.), where *wh*-movement is possible from the PP, but we do have possessive forms, as in (207). They note the problem for their analysis in the special case of (190b,c), leaving it unsolved,

but in fact the problem is considerably more general, as we have seen. The problems all seem to be overcome in a natural way along the lines just sketched, with essential reliance on SSC and the modified version of subjacency.

There seems to be some reason, then, to take \bar{S} to be a cyclic node for the definition of subjacency (and for some languages, perhaps, SSC as well). The basic insight of Bach and Horn makes it possible to overcome what seemed to be a fundamental objection to this approach, and when incorporated within the framework outlined earlier, provides a natural explanation for an interesting class of phenomena.

There are further consequences that should be investigated. Thus, it is no longer clear that \bar{S} must be taken as a cyclic node for subjacency. The question has consequences with regard to preposing rules and other matters. Furthermore, the standard argument for the relative rather than absolute interpretation of the A/A principle—namely, that NP can be extracted from NP by *wh*-movement—disappears, leaving open the possibility that this principle can be interpreted differently. Cf. Kayne (1975) for some ramifications. Cf. also note 54. I will have to leave these interesting questions open.

I will conclude this discussion with some remarks about the adequacy of the general thesis (50): specifically, can we appeal to *wh*-movement and the conditions assumed for a general explanation of CNPC and *wh*-island constraints? Do these and similar phenomena appear outside of the domain of rules of construal in the sense suggested (including movement rules, under the trace theory)? I cannot hope to review the substantial literature on this question here, but will consider a few cases.

Some examples in the literature allegedly illustrating conditions on rules may have been wrongly analyzed. For example, I have just been arguing that the analysis of (209) in Chomsky (1973) was incorrect:

(209) **who did you see John's picture of*

To take another example, it is argued in Chomsky (1973) that (210b) is blocked by SSC applying to the rule associating *not*, *many*, giving essentially the meaning "few":

- (210) a. *we didn't see pictures of many of the children*
 b. **we didn't see John's pictures of many of the children* (* on the relevant interpretation)

But consider (211):

(211) *we didn't believe that Bill had seen pictures of many of the children*

It seems that in this case *not* can be associated with *many*, violating SSC and PIC if the rule is a rule of construal. In our present framework, there is no reason to suppose that it is. Thus we are left without an explanation for (210).

Perhaps what blocks (210) is not SSC but rather a prohibition against associating *not* and *many* when the latter is within a "specific" NP, whether definite or specific indefinite. Consider (212):

- (212) a. *we didn't see the pictures of many of the children*
 b. *we didn't see certain pictures of many of the children*

In all such cases, association of *not*, *many* seems difficult or impossible. Perhaps, then, the problem with (210b) is simply that the possessive NP *John's* is definite.⁵⁵ Thus what appeared to be a case of SSC fails under a different principle.

Analogous questions arise in the case of the quantifier *any*, often held to be subject to island conditions on scope determination. Fauconnier (1975) argues that (213b) is prevented by CNPC from having scope outside of NP, as compared with (213a):⁵⁶

- (213) a. *I didn't see anyone's husband at the meeting*
 b. **I didn't see the man anyone is married to at the meeting*

However, a further look suggests that specificity of the NP, not CNPC, may be what is involved. Consider (214):⁵⁷

- (214) a. *we can't find books that have any missing pages*
 b. **we can't find the books that have any missing pages*
 c. **we can't find certain books that have any missing pages*

The cases differ just as (210)–(212) do. One may interpret (214a) with wide scope for *any*, as for example, in despair after a search for certain missing pages has failed, even though *any* is within a complex NP.

Some discussions purporting to show that island constraints hold without movement seem to me to be based on rather questionable data. For example, Bresnan (1975) argues that CNPC applies in nonmovement rules on the grounds of such examples as (215):

- (215) a. *who was planning to buy what*
 b. *who was arguing about a plan to buy what*

As she notes, we must exclude the interpretation as echo questions. We can do this, for example, by embedding (215), as in (216):

- (216) *I wonder (don't remember)* a. *who was planning to buy what*
 b. *who was arguing about a plan to buy what*

To demonstrate that CNPC holds in these cases, where there is plainly no movement rule applying, we must argue that (216b) is starred but not (216a).

I do not myself perceive any significant difference in acceptability between (216a) and (216b). But even if there is such a difference, it does not suffice to show that CNPC holds in this case. To establish that CNPC holds it is necessary to show that structures of equivalent complexity with a cyclic node \bar{S} in place of NP are acceptable, while the structures with NP are not. Thus to establish that CNPC holds of direct questions does not suffice to compare (217) with (218), where brackets bound cyclic categories:

- (217) *who do you believe [that John saw]*
 (218) **who do you believe [the claim [that John saw]]*

These examples suffice only to establish the weaker "complex phrase condition."

To show that the relevant condition is, rather, CNPC, it is necessary to contrast (218) with (219):

(219) *who do you believe [that Bill claimed [that John saw]]*

Noting that (219) is grammatical while (218) is not, we establish that the “complex phrase condition” does not suffice and that in fact CNPC is operative. This is the course we have followed in the foregoing discussion.

Returning now to (216), to establish that CNPC holds we must consider such cases as (220):

- (220) *I wonder (don't remember)*
- (a) *who was arguing [that Bill planned [to buy what]]*
 - (b) *who was arguing about [a plan [to buy what]]*

Only if (a) and (b) differ crucially in grammaticality is there an argument for CNPC from these cases. But I see no difference, certainly nothing comparable to the distinction between (219), (218), which is the relevant analogue. It seems to me that double-*wh* structures are fairly free, in violation even of such constraints as the coordinate structure condition (cf. (221), subject to some qualifications about increasing complexity and its effect on naturalness, which may very well hold quite generally (e.g., in such cases as (219)), and therefore belong to an independent component of the full system of language and language use.

(221) *I wonder (don't remember) who went to the store to buy wine and what*

(222) *I don't remember who wondered how to do what to whom*

Similarly, Bresnan argues on the basis of (223) that the rule in question observes CNPC, but a satisfactory argument would require a basic difference between (223b) and (224):

- (223) a. *who saw pictures of whom*
 b. *who heard claims about pictures of whom*

(224) *who heard that Bill saw pictures of whom*

I am not at all convinced that there is a relevant difference. My judgments on these examples are not at all firm, but I would tend to take them as evidence that non-movement rules do not observe the constraints in question.

Bresnan's most interesting and important argument, however, deals with another matter, namely the rule she calls “comparative subdeletion” (C-Sub), which yields such sentences as (225), from Bresnan (1975):

- (225) a. *they have many more enemies than we have friends*
 b. *she seems as happy now as she seemed sad before*
 c. *my sister drives as carelessly as I drive carefully*

Elimination of the boldfaced word in (225) gives the corresponding comparatives, which Bresnan takes to be derived by a deletion rule falling together under a single generalization with C-Sub.

Bresnan argues further that C-Sub observes CNPC, as illustrated in (226)–(228):

- (226) a. *this policy has been as harmful to our interests as people believed it would be beneficial*
 b. **this policy has been as harmful to our interests as people believed the claim that it would be beneficial*
- (227) a. *I'll have to give as many Fs as you've proposed to give As*
 b. **I'll have to give as many Fs as you've discussed a proposal to give As*
- (228) a. *it has done no less harm than you say it has done good*
 b. **it has done no less harm than you have the opinion that it has done good*

To show that CNPC is involved, rather than just a “complex phrase constraint,” we must compare not the (a) and (b) cases of (226)–(228), but rather the (b) cases and such examples as (229)–(231):

- (229) *this policy has been as harmful to our interests as people believed that Tom claimed that it would be beneficial*
- (230) *I'll have to give as many Fs as you've mentioned that Bill proposed to give As*
- (231) *it has done no less harm than you informed me that it has done good*

Certainly (229)–(231) are much less acceptable than the corresponding comparatives, with the boldface phrases removed. This is characteristic. To take another, simpler case, consider (232):

- (232) a. *the desk is as high as it is wide*
 b. **the desk is as high as they believe the claim that it is wide*
 c. *?the desk is as high as they believe that Bill claims that it is wide*
 d. *the desk is as high as they believe that Bill claims that it is*

In (232), the basic judgments seem to me to be that (a) and (d) are fully acceptable, whereas (b) and (c) are not. Case (232a) is C-Sub; case (232d) is comparative formation, which I have argued is *wh*-movement. If there is no further difference between (232b) and (232c), then we may simply say that a “complex phrase constraint” applies to C-Sub. If we take the difference to be significant, with (232c) considerably more acceptable than (232b) (and comparably in the examples (226)–(231), then we might decide to accept (232c) as grammatical, explaining its relative unacceptability in some other terms, say, in terms of some performance factor—though why such a factor should apply in C-Sub but not in comparatives is unclear.

Bresnan notes that “acceptability of sentences involving Subdeletion seems to decay more rapidly as length and complexity increase than with [Comparative Deletion].” She also notes that “natural contrasts” or “foci” are required in C-Sub to a much greater extent than in comparatives, citing Akmajian. But these considerations raise some doubt as to whether in fact C-Sub observes CNPC and whether it can be coalesced with comparative deletion or regarded as a rule operating over a

variable. The "variable" in question must be subject to some condition indicating that it is not too complex, in some sense, and that the appropriate parallelism holds. This notion of "complexity" is qualitatively different, it seems, than the performance factors that may apply in cases of *wh*-movement (comparatives in particular). It may well be that the conditions of complexity and parallelism, when properly formulated, will simply rule out such cases as the (b) examples of (226)–(228) and (232) as being particularly bad. If so, we do not have a case of CNPC, just as (210) does not illustrate SSC, though superficially it appears to do so.

It seems, in fact, that very slight modifications suffice to cause decay of acceptability of C-Sub. Consider such cases as the following:

- (233) a. *the desk is wider than it is high*
 b. *the desk is wider than it used to be high*
 c. *the desk was wider than it seems to me to be high now*
- (234) a. *she seems as happy now as she seemed sad before*
 (= (225b); Bresnan's (83))
 b. *she seems as happy now as she* $\left\{ \begin{array}{l} \text{was sad before} \\ \text{has ever been sad} \\ \text{will ever be sad} \end{array} \right\}$
- (235) a. *John is happier today than he usually is sad*
 b. *John is happier than he looks healthy*
 c. *John looks more satisfied than he is happy*
 d. *John is more healthy now than he has been happy for many years*

It seems to me that (233b,c), (234b), and (235) are very low in acceptability, hardly better than (232b) (if at all), although the comparatives formed by removing the italicized word in these examples are perfectly acceptable and the modification that gives the unacceptable C-Sub forms is rather slight. Thus it seems to me difficult to establish that C-Sub meets CNPC, that it falls under the same generalization as comparative formation, or that it is a rule operating over a variable.

To summarize so far, I have argued that comparatives are formed by *wh*-movement, and that there seems no reason to postulate a second rule of comparative deletion that is extensionally identical (as a function) to *wh*-movement over a sub-domain of the latter. I see no reason to believe that C-Sub constructions challenge that conclusion. However, it remains to determine how C-Sub relates to the general thesis (50). Specifically, is C-Sub a rule of deletion over a variable meeting the conditions (49)? If the answer is positive, we must permit a new category of rules, deletion over a variable, thus expanding the class of permitted grammars. Furthermore, we must abandon the thesis (50) and with it the explanation for CNPC, *wh*-island constraints, and cross-over.⁵⁸ But the crucial data seem to me relatively unconvincing. Until some formulation of the relevant notion of "complexity" or "parallelism" is advanced, we really have no way of knowing whether the restrictions on C-Sub bear on the thesis (50) at all.

But in fact further analysis shows, I think, that little hinges on the question of whether C-Sub is taken to observe the conditions (49). We can see why by considering more carefully the applicability of Bresnan's relativized A-over-A condition (RAOAC) to the case of C-Sub.

Recall that RAOAC guarantees that application of *wh*-movement to (236) will give (237), not (238):

- (236) a. *John read* [[*how many*] *books*]
 b. *John is* [[*how (much)*] *tall*]
- (237) a. *how many books did John read*
 b. *how tall is John*
- (238) a. **how many did John read books*
 b. **how (much) is John tall*

The condition guarantees that the larger bracketed phrase of (236) is extracted, in these cases. Bresnan argues that the same condition is applicable in the case of C-Sub. Given (239) we form (240) by C-Sub, deleting *X*:

- (239) *the desk is as high as it is* [[*X*] *wide*]
 (240) *the desk is as high as it is wide*

Bresnan takes *X* to be a QP, say, *that much*. Why doesn't RAOAC apply, deleting (241a), as it moves (241b) in (236b) or (according to Bresnan's analysis) as it deletes (241a) in (242):

- (241) a. [[*X*] *wide*]
 b. [[*how (much)*] *tall*]
- (242) a. *this desk is as wide as that one is* [[*X*] *wide*]
 b. *this desk as as wide as that one is*

The reason, Bresnan argues, lies in the principle of recoverability of deletion. Thus RAOAC requires that we apply the rule to the maximal appropriate phrase that is not distinct from its antecedent; (241a) is distinct from its antecedent in (239) but not (242a); therefore only *X* is deleted in (239). It is this assumption that permits Bresnan to take comparative deletion and C-Sub to be the same rule.

But the assumption seems to me questionable. Notice in the first place that on this approach, we must take (243) to be analogous to (244) rather than (245).⁵⁹

- (243) a. *John is taller than Bill is tall*
 b. *John is taller than he is tall* (take *he* to refer to John)
- (244) *how is John tall* (= (238b))
- (245) a. *John's height exceeds Bill's height*
 b. *John's height exceeds his height* (take *he* to refer to John)

The reason is that under this analysis, (243a,b) derive by the same violation of RAOAC that gives (244). But this conclusion seems to me highly counterintuitive. Rather, it seems to me that (243a,b) are quite analogous to (245a,b) and very

different from (244); specifically, (243b) seems to be simply a logical contradiction, like (245b).

But in fact there is additional evidence that Bresnan's analysis of C-Sub is defective. A crucial requirement of this analysis is that (243) must be marked ungrammatical, as a violation of RAOAC. But in fact, neither (243a) or (245a) (nor, for that matter, (243b) and (245b), which I take to be just contradictory) is ungrammatical, as we can see readily by constructing an appropriate context. If this is correct, then we can understand why (243) seem analogous to (245) rather than (244), which really is ungrammatical. As relevant contexts, consider the following:⁶⁰

- (246) Speaker A: *John is more courageous than Bill is intelligent*
 Speaker B: *No, you've got it all wrong; John is more courageous than Bill is courageous*
- (247) Speaker A: *this desk is higher than that one is wide*
 Speaker B: *What is more, this desk is higher than that one is high*
- (248) Speaker A: *this desk's height exceeds that desk's width*
 Speaker B: *In fact, this desk's height exceeds that desk's height, too*

Similar examples can be constructed for (243), (245), apart from the difficulty of finding a natural contrast to "tall":

- (249) Speaker A: *John is taller than Bill is heavy*
 Speaker B: *What is more, (243a)*
- (250) Speaker A: *John's height exceeds Bill's weight*
 Speaker B: *Furthermore, (245a)*

In short, when context supplies an adequate reason for placement of the required stress on the compared form in C-Sub constructions, examples such as (243), (245) (but never (244)) are quite all right. The simplest explanation for this fact, avoiding any elaborate complication of rules to distinguish somehow between cases of phonetically identical stress, is that C-Sub simply removes *X* in (239), etc., and is not subject to RAOAC. The remaining element is stressed, but for reasons having nothing to do with C-Sub; cf. (245). If the remaining element happens to be identical with the paired phrase that is its "antecedent," then the sentence is either contradictory (as in (243b) and (245b)), or else must be understood as in the discourses cited. All of this will form part of the rules of interpretation for foci in C-Sub constructions.

Notice now that there is no basis at all for generalizing C-Sub and comparatives, which is just as well in the present framework, for reasons already discussed. But we can go further. Consider the choice of the element *X* eliminated in C-Sub, as in (239) or more generally, (251):

- (251) . . . *than (as) NP is* [[_{QP} *X*] . . .]

There are several kinds of familiar deletion phenomena. Typical examples are VP-deletion, as in (252); *wh*-deletion, as in (253); and *for*-deletion, as in (254):

- (252) *John left early but Bill didn't (leave early)*
- (253) a. *the man (who) you met left early*
 b. *John is taller than (what) Bill is*
- (254) *John wants (for) to leave*

In such examples as (252), there is typically a variant with the deleted phrase unstressed. It may well be that this is the only kind of deletion that involves lexical items; namely, deletion "under identity" (cf. p. 81, above) of a phrase that can appear unstressed.⁶¹ Examples (253)–(254) illustrate another major class of deletions—perhaps the only other case—namely, deletion of designated terminals, sharply restricted, and often with optional or dialectal variants without deletion.

Let us now return to C-Sub and ask where deletion of *X* in (251) fits into this pattern. Plainly, it is not a case like (252); there is no variant with an unstressed expression. Nor are there optional or dialectal variants. The deleted element *X* must simply be absent; period. The rule of C-Sub, as we have seen must refer specifically to *X*; it does not fall under RAOAC, as in (237). Assuming Bresnan's analysis, *X* is simply some representative of QP that is obligatorily eliminated.

We do have an element that is obligatorily deleted under some conditions, namely, *wh*-. Suppose, then, that we were to take *X* = *wh*- or to take *wh*- to be a feature of *X*. This choice allows us to express the relation between comparatives and C-Sub constructions in terms of presence of *wh*-. Furthermore, the obligatory deletion might fall under a broader generalization or might require no rule at all, given that *wh*- in isolation has no phonetic content. And we can easily formulate RAOAC so that it does not apply to "bare" *wh*- but only to phrases *wh*-*Y* (*Y* some terminal string) of the form *X*-bar (with the right number of bars; three in Bresnan's theory). We might, for example, limit RAOAC to cases where *wh*- is a specifier, in the sense of *X*-bar theory, of some lexical category, as it is in all the cases where RAOAC applies but not in (251), where it does not.⁶²

Pursuing this approach, we will have *wh*-movement followed by the familiar *wh*-deletion in C-Sub constructions.⁶³ It will follow, then, that C-Sub has the properties (49).

Bresnan gives a number of arguments against the assumption that a movement rule such as *wh*-movement applies in C-Sub. There are two basic points. The first is that there are no dialectal variants with *wh*-words in the case of C-Sub; i.e., no examples such as (255) analogous to (256):

- (255) *John is more courageous than what Bill is intelligent*
 (256) *John is more courageous than what Bill is*

But this argument does not apply to the analysis just suggested. Under this analysis, there is no form such as (255) for the reason that no *wh*-word was moved, but only *wh*-, which cannot be phonetically realized. The second argument is that where there is a lexical string in place of *X* in (251), extraction of QP is impossible, as illustrated by (238). That is, "certain kinds of left-branch modifiers cannot be moved away from the constituents they modify." Bresnan explains this fact in

terms of RAOAC, and we have been relying on her explanation in the case of questions and relatives. But we have already seen that RAOAC does not account for C-Sub; rather, we must reformulate either RAOAC or C-Sub, perhaps along the lines just sketched, so that C-Sub does not fall under RAOAC. Therefore, this class of arguments against a movement rule no longer applies. Whether we have deletion or movement, the left-branch modifier involved in C-Sub is not subject to the general left-branch condition, which Bresnan convincingly explains in terms of RAOAC. In short, it does not matter whether we assume that the designated element *X* of (251) is deleted in place, or is moved by *wh*-movement and then deleted by an obligatory rule; in either case, either because it has no phonetic content in principle (and therefore, strictly speaking is not deleted) or as a subcase of the familiar rule illustrated in (253).

We can now see that C-Sub, though an extremely interesting phenomenon, does not seem to be relevant to our current discussion or to the thesis (50). If we decide to rule such "complex" examples as (232c) ungrammatical, then it follows that CNPC, etc., simply do not apply to C-Sub. In accordance with this decision as to the facts, we will formulate C-Sub as a rule deleting *X* of (251) in place; whatever conditions are established regarding complexity and parallelism will form part of the associated rule of interpretation. The rule is no longer "deletion over a variable"; we therefore do not have to admit this new category of rules into the grammar, and nothing follows concerning the general thesis (50). Or, if we decide, with Bresnan, that (232c), etc., are grammatical, then we will conclude that C-Sub does observe our general conditions subject to some extragrammatical factors that account for the rapid decline in acceptability with complexity and for the focus and parallelism requirements. In accordance with this interpretation of such constructions as (232c), we will stipulate that *X* of (251) is (or has the feature) *wh*-; We now have just another bit of evidence corroborating the general thesis (50), though very weak evidence because of the ambiguous status of (232c), etc.

The choice between these two alternatives will have to await a better understanding of the conditions on complexity and parallelism involved in C-Sub constructions. As far as the general thesis (50) is concerned, nothing seems to follow, either way.

If this line of argument is correct, we have then a very welcome outcome. Namely, there seem to be no clear counterexamples to the general thesis (50). The consequences have already been noted several times. We have an explanation for a variety of otherwise unexplained constraints in terms of rather simple conditions on rules, conditions that seem entirely natural as limitations on procedures of mental computation. Furthermore, we can reduce drastically the set of available rules. There will be no asymmetry between rightward- and leftward-movement rules; all are upward-bounded, in Ross's sense. There is no distinction between bounded and unbounded rules. All movement rules are simply subject to subjacency, if they are part of the cycle. There is no clause-mate constraint applicable to certain rules but not others. The only deletion rules are those of the type (252)–(254), and of these, only (252) are non-trivial. Rules of construal and no others are subject to the basic conditions

(4), (5); we thus have a rather natural formulation of an autonomy thesis for formal grammar, as noted earlier. More important still, we have some reason to believe that for the core grammar at least, the expressive power of transformational rules can be vastly reduced so that very few possibilities are available at all. Thus the class of possible grammars is significantly reduced and we have a natural and rather far-reaching explanation for phenomena of the sort under discussion here. Of course, these conclusions will only hold if the problems noted along the way and many others, no doubt, can be overcome.

Reduction of the class of available grammars is the major goal of linguistic theory. To account for the fact that language is acquired as it is, we must find ways to restrict the "space" of potential grammars to be searched by the language learner. Note that reduction of the class of grammars is not in itself an essential goal, nor is restriction of the class of generable languages; it is the class of "available" grammars that is important. We might in principle achieve a very high degree of explanatory adequacy and a far-reaching psychological theory of language growth even with a theory that permitted a grammar for every recursively enumerable language. The reasons are those outlined in Chomsky (1965), chapter 1, section 9. What is important is the cardinality of the class of grammars that are compatible with reasonably limited data and that are sufficiently highly valued. We achieve explanatory adequacy and approach a successful "learning theory" for language to the extent that this class is small, irrespective of the generative capacity of the class of potential grammars. We can try to keep this class "small" by restrictive conditions on the various components of the grammar (e.g., the X-bar theory for the categorial component of the base). The preceding discussion suggests other ways in which the variety of highly valued grammars can be reduced—quite significantly, if the suggestions developed here prove tenable.

Acknowledgments

I am grateful to Leland George, Norbert Hornstein, and Howard Lasnik for helpful comments on an earlier draft of this paper, and to many students whose suggestions and criticisms in the course of lectures and class discussion have been worked into the text.

Notes

1. As noted in Chomsky (1973), the principle of strict cyclicity as there formulated implies that *wh*-movement is cyclic. Bach and Horn (1976) state that they do not see why this principle implies that *wh*-movement is successive cyclic. The problem they perceive arises from their conclusion that when I wrote that the principle implies "cyclicity," I really meant "successive cyclicity"; cf. their note 23. But I did mean "cyclicity," and the problem they see does not arise.
2. Kayne suggests a possible deep structure for this case, but it seems rather artificial.

3. Note that there is no way to explain these facts in terms of a "clause-mate" constraint and a rule of raising to object. In the first place, reciprocal interpretation is not subject to a "clause-mate" constraint. cf. (7), (8); in fact, I think there is no credible evidence that any transformational rule or rule of construal is subject to such a constraint, i.e., that there is any reason to permit this option within linguistic theory. Furthermore, in many dialects we have such sentences as "they want very much for each other (themselves) to win," completely ruling out any such analysis. In general, even if there is a rule of raising to object position, which I doubt, it will not apply to *want*-type verbs, for reasons discussed in Bresnan (1970, 1972, 1976c); Lightfoot (1976a). See the latter for a general review of the matter.
4. Cf. note 3. Note that in all dialects, "*they* want very much for *them* to win" requires disjoint reference between the italicized positions. I will assume here that EQUI is correctly analyzed as deletion of "*X*'s self" (*X* a pronoun) in the context *for* - VP, optionally in some dialects, obligatorily in others, yielding the dialectal "they want for to do it" and the standard "they want to do it" with *for*-deletion before *-to*, under conditions that are moderately complex and somewhat variable across dialects. For discussion, cf. Chomsky (1975c), and for an independent argument, cf. Fodor (1975), pp. 141ff.
5. Bach and Horn (1976), in a criticism of Chomsky (1973), argue that "the total effect of the Specified Subject Condition . . . (etc.) . . . is to block extraction from" noun phrases. This is a rather selective reading. Examples of reciprocal interpretation and disjoint reference, not to speak of many others discussed in Chomsky (1973), have nothing to do with extraction from noun phrases. Thus even if they were correct in their proposals concerning noun phrases, to which I return below, the consequences for the analysis presented in Chomsky (1973) would be slight, it seems to me.
- presented in Chomsky (1973) would be slight, it seems to me.
6. The point of these examples is that by reliance on PIC and SSC, which are independently motivated for interpretive rules, we can significantly reduce the expressive power of the theory of transformations, perhaps even to such a level that basic rules can be formulated as in (1). For discussion, cf. Chomsky (1975b,c). Even if this reduction is unattainable, the effect of the conditions discussed is considerable. This is important, since naturally we are concerned to reduce the class of grammars potentially available.
7. These examples are from Kayne (1975), as reanalyzed by Quicoli (forthcoming a,b,c).
8. Such examples as (i) have been suggested as counterexamples:
- (i) *los hombres parece [que t estan cansados]*
 But as Quicoli observes, this appears to be a case of topicalization with subsequent deletion of the subject pronoun rather than a case of raising. Under the analysis of topicalization presented below, PIC is irrelevant here. This is another example of the irrelevance of unanalyzed examples to confirmation of conditions on rules.
9. Cf. Kayne (1975), Quicoli (forthcoming b), Pollock (1976). The English examples are cited in Chomsky (1973) from Fauconnier (1971), who gives the French equivalents.
10. One crucial assumption in this analysis is that in English COMP cannot be doubly filled. It follows that no more than one element can be extracted to the matrix sentence from a complement clause. Postal (1976a) argues that this assumption is incorrect, as shown by (i) and (ii):
- (i) *under those conditions, what do you think I should do*
 (ii) *if he comes, what do you think I ought to do*
- He argues that "both of the italicized phrases have been extracted from the complement of *think*." Postal does not formulate the rules that he believes to be operative here, but presumably he is assuming that certain phrases *X* are preposable to sentence initial position in the context (iii):
- (iii) *what you* VP [_S NP V . . . *X*]

Assuming that some such rule is what he has in mind, we see at once that it is incorrect. Cf. (iv), (v):

- (iv) **under those conditions, what did you tell Mary [that I should do t]*
 (v) **if he comes, what did you tell Mary [that I should do t]*

To be precise, (iv) and (v) are not starred, but rather cannot be interpreted as extraction from the position marked with *t*, but only as preposing from the matrix clause. Thus the rule that seems to be presupposed by Postal's discussion is wrong. This leaves us with the problem of explaining (i), (ii). Whatever the explanation may be, notice that the phenomena cited have no direct bearing on the conditions on rules that Postal is discussing, for reasons already elaborated several times. Postal's discussion of alleged counterexamples to SSC is a good example of the fallacy that I have noted several times: phenomena do not bear directly on conditions on rules; only rules do. In no case does he propose a rule that violates (or confirms) these or any other conditions. Similar criticisms with regard to Postal (1974a) appear in Lightfoot (1976a); Bresnan (1976c).

As far as (i) and (ii) are concerned, perhaps the explanation is that such phrases as "you think" are subject to a reanalysis as adverbials, so that none of the relevant conditions apply, just as we have "violations" of CNPC under the reanalysis indicated in (vi) and (vii):

- (vi) *what did he [make a claim] that John saw* (acceptable, for many speakers)
 (vii) *what did he [have an opportunity] to do*

Such reanalysis is motivated by the familiar analysis of tags; e.g., (viii), (ix):

- (viii) *I think that John will come, won't he*
 (ix) **I told Mary that John will come, won't he*

Whether or not this suggestion is correct, I stress again that such examples as (i) and (ii) have no relevance to the question of the adequacy of conditions on rules *in themselves*, but only indirectly, insofar as they indicate what the rules of grammar might be.

11. Compare **"what did they hear some funny stories about pictures of,"* blocked because of subadjacency. Cf. Chomsky (1973) for discussion of some complications.
12. A more familiar assumption is that relativization in these languages leaves a copy. I am assuming that pronouns are base-generated, and that the power of transformations is so restricted that pronouns (or, for that matter, lexical items in general) cannot be introduced by transformation. For discussion, see Wasow (1972), Lasnik (forthcoming). Perhaps the most convincing argument against a pronominalization transformation, in my opinion, is the one given by Dougherty (1969). He points out that *in positions where nouns and non-anaphoric pronouns can freely occur*, pronouns that can be understood anaphorically can also be understood nonanaphorically, a fact unexplained under a transformational analysis but immediately explicable on the assumption that pronouns are base-generated (his "anaporn relation"). Thus a transformational analysis is missing an important and obviously nonaccidental generalization. Postal (1972) argues that Dougherty's observation is false, on the basis of examples in which pronouns occur in positions where nouns and nonanaphoric pronouns do not freely occur. Since the proviso italicized above is perfectly explicit in Dougherty's discussion (cf. his note 13) Postal's rejoinder is completely beside the point.
13. It has been noted that English speakers sometimes use a construction with a pronoun where an island constraint would block relativization, as in (23ii); cf. Andrews (1975a) for some discussion. I suppose that this is an ancillary process, not to be incorporated, strictly speaking, within the grammar.
14. On the significance of erasure of trace, as in NP-preposing in (29), cf. Fiengo (1974), Chomsky (1974, 1975b). Note that trace theory introduces no "globality" in any unwanted sense, contrary to what is sometimes assumed. Cf. Chomsky (1975b) for discussion.
15. We may take *who*, *what* to be, in effect, *wh-person*, *wh-thing*, respectively. Thus *who* is analogous to "what student." Relations between interrogatives and indefinite pronouns, discussed in Chomsky (1964), Postal (1965), will be expressed, within this framework, as conditions on variables in LF, along the lines of Chomsky (1975c).

- 105
16. In Chomsky (1973), section 17, it is suggested that the rule of *wh*-movement might be replaced by an interpretive rule for *wh*-phrases generated in COMP position. The rule of interpretation would then be something like (38). I think that this is entirely possible, but I am not convinced that it is a meaningful alternative to the transformational analysis as a movement rule, for reasons discussed in Chomsky (1975c). The same may be said about the proposal to replace NP-movement rules by interpretive rules. It seems to me that we have three types of rules, each with their separate properties: NP-movement, *wh*-movement, rules of construal (and, of course, others: e.g., extraposition, quantifier-movement or interpretation, FOCUS, predication, etc.). If all are regarded as interpretive rules, we still have the same three collections of properties, which can, in fact, be explained (rather than stipulated) if we take the NP-movement and *wh*-movement rules to be movement transformations meeting the conditions discussed here.
 17. Direct application of *wh*-movement to give (41) is blocked by PIC and SSC. Application of *wh*-movement in the embedded clause will give "you told Mary [who she should meet]." Extraction of *who* on the next cycle is blocked by PIC, since the bracketed phrase is a tensed clause, and by SSC, under the present formulation (but not that of Chomsky, 1973), since it contains a specified subject. Cf. notes 37, 38. Bach and Horn (1976) state that in Chomsky (1973) a "special clause" is required "allowing extraction over a specified subject by movement into a COMP node." That is incorrect. Movement to the COMP node in the first cycle does not fall under (11), as explained in Chomsky (1973), because there is no internal cyclic node.
 18. Examples (a) and (b) are from Dean (1967). Example (c) is from Erteschik (1973), who gives a detailed discussion of the topic. The oddity of *wh*-movement from certain factives is noted in Kiparsky and Kiparsky (1970).
 19. In Chomsky (1973), note 22, I remarked that none of the arguments in the literature appear to apply to the formulation of cyclicity of *wh*-movement proposed. Bach and Horn (1976) dispute this observation, claiming that these arguments do apply. They present one example, which, as they note, is based crucially on the assumption that *wh*-movement is obligatory in the embedded clause. They fail to note, however, that I explicitly assumed the rule to be optional; cf. section 13. Thus the statement to which they object seems to me accurate. They assert that optionality of *wh*-movement undermines arguments for successive-cyclicity, but they present no grounds for this conclusion (they do offer arguments purporting to show that strict cyclicity is untenable, but whatever the merits of these may be, the question is a quite different one). In fact, optionality of *wh*-movement is irrelevant to the arguments for successive cyclicity. My assumption is that all rules of the "core grammar" (excluding what Bach, 1965, calls "housekeeping rules") are optional, the apparent obligatoriness deriving from filters and principles of interpretation, along the lines discussed in Chomsky (1973). Cf. Lasnik and Kupin (forthcoming), Chomsky and Lasnik (forthcoming).
 20. Note again that we also must presuppose the "superiority condition" of Chomsky (1973), independently motivated by "I don't remember who saw what," "I don't remember what who saw."
 21. Note that if a language has no transformations, this "escape hatch" is unavailable in principle. Therefore it follows that for such languages, extraction of *wh*-phrases from embedded clauses is impossible. Cf. Hale (1976) for an argument along these lines for Walbiri. Note that the sufficient condition cannot be strengthened to necessary and sufficient.
 22. Bach and Horn (1976) suggest a different explanation for CNPC, to which I will return.
 23. Bresnan notes that there is no convincing evidence for cross-over in the case of comparative subdeletion, and that some of the cross-over phenomena do not seem to hold for subdeletion. Reanalysis of subdeletion along the lines discussed below has as a consequence that these structures should differ from comparatives in this respect.
 24. The analysis that follows is essentially that of Chomsky (1974). It was suggested by material presented in Emonds (1976) based on ideas of Ann Banfield. I am indebted to Ivan Sag for emendations, as noted below.

25. How seriously one should take this last remark I am not sure. There is no particular reason to take the *wh*-phrase of the COMP in relatives to be a quantifier binding the variable, and it may be that a natural semantic interpretation of relatives, along the lines previously discussed, will regard the variable introduced as free whether or not the *wh*-phrase in the COMP is deleted. Cf. discussion following (24) and Vergnaud (1974).
26. E.g., subjects, extraposed clauses, pseudoclefts. Also such structures as "I'm sad he left" (cf. "I'm glad he left"), "I muttered he'd better leave," etc.
27. After writing this sentence, I noticed that it illustrates the property of left-dislocation discussed earlier without clear examples, namely, that the proposition need not be "open" but can be about the focused element of the TOPIC in some more vague way.
28. This is ungrammatical, but for independent reasons; namely, free relatives with *who* are excluded in general by a special restriction. Thus "this book is what I want Bill to tell Mary to read" is fine.
29. The last example is from Pinkham and Hankamer (1975). We should have also "pea green is what Tom ordered us to paint our boat"; "pea green, Tom ordered us to paint our boat"; "it was pea green that Tom ordered us to paint our boat." But the last two of these seem to me very questionable. The (d) cases also seem marginal, suggesting that PP might be excluded from subject or TOPIC. I am assuming here that the TOPIC is construed with the embedded clause.
30. With obligatory auxiliary inversion.
31. Pinkham and Hankamer, in their very interesting study of clefts, state that their analysis, which postulates two independent rules that form clefts (one of them structure-building), is intended "as a challenge to any analysis" that is not structure-building. But I think that the data they cite, where judgments are clear, is just as well handled by postulating the underlying structure (85) and no rule of cleft-formation at all, apart from the interpretive rules. Note that this analysis covers two cases, but these do not coincide exactly with their two cases. I am not convinced by some of the crucial data that they offer to demonstrate that the examples divide as they propose, but will not pursue the matter here. They note a parallel between PP-preposing and clefting (p. 438), but it is not exactly the parallel noted here. I am suggesting, in effect, that the parallel is far broader and not limited to extraction from the scope of a negative as they propose.
32. Note that we cannot have *believe* in this position as in (94b). The reason has to do with general properties of infinitives. Nonagentive constructions would be equally odd in "I ordered Bill to believe that Mary left," "John is hard to believe to have left," etc.
33. On the status of embedded finite clauses in these constructions, see below.
34. For discussion of infinitival relatives from a somewhat different point of view, see Emonds (1976).
35. Judgments vary, as is generally the case when *each other* is in subject position: cf. "they wanted each other to win," "they prefer for each other to win," "they would hate it for each other to win," "they would hate it for each other's pictures to be on sale," etc.: see Chomsky (1973) for some discussion. However, there seems no question that (103b) is incomparably more acceptable than (103c). Note that *books* must be plural in (103b); as is generally the case in reciprocal constructions for reasons that remain obscure: cf. "we saw pictures (*a picture) of each other," "we turned the arguments (*the argument) against each other," etc. Cf. Chomsky (1973) for discussion.
36. Similarly, "who₂ did you find a book (which₁) for us to give t₁ to t₂," "[to whom]₂ did you find a book (which₁) for us to give t₁t₂," "[what book]₂ did you find a person (who₁ for) to give t₂ to t₁," "[what book]₂ did you find a person [to whom]₁ (for) to give t₂t₁."
37. The proviso (46) permits escape from COMP in a tensed sentence. If the notion "subject" is so defined that the subject of S is also the subject of \bar{S} , then (46) is required to permit escape from COMP in infinitives as well. If the subject of S is not the subject of \bar{S} , then nothing will prevent movement from COMP in infinitives. In Chomsky (1973) I made the

latter assumption, in the foregoing discussion, the former, largely for expository reasons. If bridge conditions are limited to finite clauses, as appears to be the case, then "subject of" should be defined only for S, not \bar{S} , so that SSC is inapplicable to COMP-COMP movement.

38. The basic observation is due to Ross (1967). He notes that some restriction must be formulated to rule out "this rock is too heavy for us (to try) to claim that we picked up." Cf. also Ross (1973): "Gravel pizza is tough for me to prove that she thought of." Ross stars these examples. Lasnik and Fiengo (1974) note that the restrictions follow from PIC, according to their analysis. In our terms, they follow by withdrawing (46) in these cases. Consider the corresponding infinitivals: "this rock is too heavy for us (to try) to order her to pick up," "Gravel pizza is tough for me to prove her to have thought of." Lasnik and Fiengo block these by SSC. If, in fact, the finite and infinitival embedded clauses give approximately the same degree of "strangeness," then in the present framework we must accept the formulation of "subject of" assumed here rather than in Chomsky (1973) (cf. note 37), and relax the language-specific proviso (46) for all these cases. If the tensed S's are indeed of a different category, then the formulation of Chomsky (1973) must be accepted, and the language-specific proviso (46) relaxed. If there is some independent reason for the "strangeness" in all of these cases, then nothing follows with regard to subject of \bar{S} and nothing need be said about (46). Judgments are sufficiently obscure, to me at least, so that I hesitate to make a definite proposal. Note that all that seems to be involved is a language-specific proviso and the precise formulation of a general principle for a domain of facts that are rather marginal.
39. The following discussion draws heavily on Lasnik and Fiengo (1974), though a somewhat different analysis is proposed.
40. I have been assuming throughout that VP is introduced only under S; thus, that infinitival subjectless complements of *promise-persuade*, etc., are \bar{S} , with NP = PRO. Deletion of *for X-self*, as in EQUI, will leave VP under S (cf. note 4). Arguments in support of distinguishing VP infinitival complements in this way appear in Quicoli (forthcoming a,b,c); cf. Chomsky (1975c). There is a similar distinction in Kayne (1975).
41. For discussion of various adjectival constructions, cf. Lees (1960b) and much later work. That the *for*-phrase is within the embedded sentence in both cases of (123c) is argued by Bach and Horn (1976). We will return to this question.
42. These examples suggest that the phenomena that Bresnan discusses under the rubric of the "fixed subject constraint" might preferably be handled by a surface filter (a suggestion that she rejects, but on grounds that can be overcome in terms of trace theory, it seems) rather than as a condition on rule application, since in these cases *wh*-movement takes place after *for*, which then deletes before *to* in standard dialects.
43. A rather similar analysis, but without *wh*-movement, is suggested in Ross (1967), but he later rejected it on grounds that were later shown to be inadequate by Akmajian. Cf. Lasnik and Fiengo (1974) for a review.
44. Cited in Grimshaw (1975) from Berman (1974). Grimshaw attributes the original observation to Huddleston (1971).
45. They do not cite (f), (g), but these are implicit in their analysis. Similarities between *easy-to-please* constructions and others that we have discussed here are noted in the literature. Cf., e.g. Evers (1975).
46. I am indebted to Alan Prince for pointing out this consequence.
47. Postal (1974) argues to the contrary on the basis of such examples as "John's tendency to talk too much," which he takes to be derived by raising to subject followed by nominalization. But he overlooks the fact that the noun *tendency* must have a different source, as in "John's tendency towards violence," where there can be no raising. In fact, it seems that there is an NP of the form "NP's tendency . . ." wherever there is a structure "NP has a tendency . . .," suggesting either a transformational analysis or a redundancy rule, in

either case, relying on base-generated *tendency*, as implied by the lexicalist hypothesis. For discussion, see Chomsky (1974), Lightfoot (1976a).

48. Perhaps the latter approach will provide a principled explanation for the other major property of *easy*—as distinct from *eager*—constructions, namely, the fact that we have "an easy man to please" but not "an eager man to please." Again, various explanations have been proposed since the basic properties of these constructions were noted (cf. Chomsky, 1962), and the investigation has clearly been a very fruitful one in terms of insights attained along the way, though the original problem remains a challenging one.
49. But see notes 37, 38.
50. Note that the Sentential Subject Constraint of Ross (1967) is a consequence of subadjacency whether or not S is cyclic, but the Subject Condition is not.
51. Of course we have "of the students in the class, how many got As" by PP-extraction after *wh*-movement.
52. In their concluding remarks, Bach and Horn note that there are many rules that extract phrases from NP, violating the NP Constraint as they formulated it. They do not consider this a problem for their analysis, apparently because NP is explicitly mentioned in the structural description of these rules. I do not fully understand their point, however, and may have misstated it.
53. Compare "we read stories about each other," "they read stories about them" (cf. Chomsky (1973) for some discussion of the latter as compared with "they read stories about themselves"). Note that SSC is required for NPs quite apart from the reanalysis that Bach and Horn suggest. Cf. note 5.
54. Examples of this sort are difficult to evaluate, since they might arise from passivization of "John — wrote — a book about Nixon" followed by extraposition from subject NP and (perhaps) interchange of PPs. The same is true of *wh*-question; see below. It has sometimes been suggested that (i) is not so deviant as (ii):

- (i) *of whom was [a picture t] standing on the mantelpiece*
 (ii) *who was [a picture of t] standing on the mantelpiece*

One might attribute this difference, if it is systematic, to extraposition from NP yielding (iii) and then (i):

- (iii) *[a picture t] was standing on the mantelpiece of NP*

To me, (iii) seems to have approximately the status of (i). Alternatively, one might argue that (i) derives by *wh*-movement directly while (ii) is blocked, appealing to the absolute interpretation of the A-over-A condition to make the distinction. If so, then S need not be taken as a cyclic node for subadjacency, and (181) will be modified accordingly. Unfortunately, the examples that seem crucial to selecting among these alternatives seem rather marginal.

55. Similar question arise in connection with (209), as noted in Chomsky (1973). See Oehrle (1974) for some relevant discussion.
56. Fauconnier refers to Postal (1974b) for a possible explanation of why quantifier scope should be constrained by islands, but the basic data that Postal assumed seem to be incorrect. Cf. Jackendoff (1975a). That is, the cases he considered seem not to be governed by such conditions as CNPC (as would be expected in the present framework).
57. Cf. Hornstein (1975), note 33, citing observations by G. Horn.
58. Recall that cross-over conditions are in part inapplicable, in part violated by C-Sub, as Bresnan notes. See note 23.
59. Note that the italicized phrases in (243)–(245) must be stressed. However, this is no special property of C-Sub, as we can see from (245).
60. I owe this point to Ivan Sag, who cites the following sentence suggested by Larry Horn:
- (i) *John drinks more Scotch than Bill does Scotch*

As Horn observes, (i) is quite appropriate in the following discourse:

Speaker A: *John drinks more Scotch than Bill does Bourbon*

Speaker B: *No, you've got it all wrong, (i)*

Boldface type indicates stress throughout these examples.

61. One might consider the possibility that such rules as VP deletion do not belong to sentence grammar at all, but rather to a theory of discourse. Cf. Sag and Hankamer (1976), Sag (forthcoming, 1976), who do not draw this conclusion but provide arguments on which it might be based. If so, then deletions can be narrowly restricted in sentence grammar, perhaps just to deletion of certain grammatical formatives and pronouns. Other deletions, where a variant appears with the deleted string unstressed and the deletion is conditional on discourse factors (hence, in special cases, sentence-internal discourse factors), would then be regarded as on a par with the rules that generate bare NPs, say, as answers to questions. If this proves to be a reasonable course, there will be certain consequences with regard to the effect of deletion rules on generative capacity. Grammars must allow some deletion of designated elements; at least this is true of any grammar that derives "the man I saw" from "the man who I saw," etc. If no constraints are placed on such deletion, then for most classes of grammars it will follow that all recursively enumerable sets can be generated, not a particularly important fact, for reasons discussed in Chomsky (1965) and below. E.g., phrase structure grammars have the weak generative capacity of unrestricted rewriting systems (arbitrary Turing Machines) if one terminal symbol is taken to be "blank." Peters and Ritchie (1973) observe that the same is true of transformational grammars, and state some general properties of grammars with cyclic rules that would suffice to reduce weak generative capacity to recursive sets. Peters (1973) suggests a rather plausible general property of transformational grammars that would suffice for this purpose, namely, his "survivor property." A number of people have observed that there is no algorithm for determining whether an arbitrary transformational grammar has this property, again, neither a surprising nor particularly interesting fact; there is also no algorithm for determining whether an arbitrary rewriting system generates finite sets, but that would not lead us to conclude that a class of "grammars" generating only finite sets cannot be specified. Peters's suggestion poses the problem of finding decidable conditions for grammars that guarantee that the survivor property (or some other sufficient property) is met, if indeed it is true that natural languages are recursive, which is by no means obvious (or, again, particularly important, in itself). Perhaps an approach to deletions of the sort just briefly discussed might provide an answer to this problem, if worked out in detail.
62. As noted by Woisetschläger (1976), Bresnan's RAOAC might be modified so that it applies to all and only "mixed terms" with a designated specifier. Then her analysis would apply at once to such examples as "so tall a man, I have never before seen," *"so tall, I have never before seen a man," where the specifier is *so* rather than *wh*-; etc.
63. Note that *wh*-deletion is vacuous in this case, perhaps, since there may be no terminal string in the first place.

COMMENTS ON THE PAPER BY CHOMSKY

Emmon Bach

*Department of Linguistics
University of Massachusetts
Amherst, Massachusetts*

Introduction

In a paper published in 1972, Chomsky characterized the state of our field as follows:

There is an appearance of considerable diversity of points of view—and to some extent, the appearance is correct. However, I think that the dust is beginning to settle, and that it is now possible to identify a number of real, empirically significant theoretical questions that have been raised, if not settled, in this work. I also think much of the apparent controversy is notational and terminological—including many issues that appear to be fundamental and have been the subject of heated, even acrimonious dispute. This is unfortunate, because it sidetracks serious work, and because occasionally certain questions of some interest are embedded, sometimes buried in these terminological debates. (Chomsky, 1972b, 63 ff.)

I think this characterization holds today as well, and I would like to make my remarks in the same spirit. I will first list what I think are uncontroversial general areas of agreement. Then I will look at some specific points of agreement across various current (and not so current) frameworks that emerge from Chomsky's latest paper(s), trying to sort out what are differences of substance and what merely terminological. I have a number of questions about the interpretation of various aspects of Chomsky's current position and some criticisms. Finally, I will try to identify a very broad issue that needs to be worked on in the future.¹

I am not going to talk much about the parts of Chomsky's paper addressed to the question of comparative deletion since I think they have been adequately treated in the discussion of Bresnan's paper.