I. Kenesei

REVISITING *WH* MOVEMENT

1. Trace theory has been one of the most interesting attempts at renewing generative grammar. It is, in part, an outgrowth of efforts to reduce the power of transformations by limiting their number and scope. Elsewhere I have already discussed one of the major transformations of the type 'move α': NP Movement (Kenesei 1982). In this paper I will examine Noam Chomsky's claims concerning a unified *Wh* Movement transformation as made in his paper "On Wh Movement" (1977) and in an article written jointly with Howard Lasnik on "Filters and Control" (1977).

NP Movement was one of the few major transformations that have swallowed up a number of old operations. Another is *wh* Movement, which was thought to be limited to cases where the *wh*-word (whether an interrogative or a relative pronoun) actually occurred in the structure or it was immediately recoverable. The second half of the disjunction refers to the optional deletion of relative pronouns, as shown in (1):

(1) a. the man (who) I saw
b. the book (which) I read about
c. the money (which) I believe he has stolen

The deletion of question-words is of course prohibited:

(2) a. *(Who) saw you?
b. *(Who) did you see?
c. *(Who) does John believe met you?
d. *(Who) does John believe you met?

However, there is another type of structure in which both embedded questions and relatives may appear: infinitival constructions:[1]

(3) He didn't know where to go
He wondered what to do
He told me who to see
He asked me how to do it

Of course question-words cannot be ommitted but relatives do require deletion of *wh*-phrases in some of the cases:

(4) a. This is the man to whom to give the presents.
b. This is the man (*who) to give the presents to.
c. *This is the man (*who) for us to give the presents to.
d. *This is the man (to whom) for us to give the presents.
e. This is the man (*who) to tell us the secret.
f. *This is the men for who(m) to tell us the secret.

Undeniably, *wh* Movement shows a number of common characteristics. One of them is the 'gap' it leaves in the sentence from which a *wh*-phrase is moved. Another is the possibility

of moving a *wh*-phrase more than one cyclic node 'up' in the tree if allowed by the intervening verb (a 'bridge'), as in (1c) and (2c-d). Thirdly, *wh* Movement from noun phrases is blocked as predicted by Bach and Horn's NP Constraint or some other device to the same effect.[2] Finally, no *wh*-phrase can be promoted from a construction introduced by another *wh*-phrase (*wh*-island constraint), as in:

(5) a. *What do you wonder who killed *t*?
b. *the man who you didn't know what happened to *t*

where *t*'s are the traces of *what* and *who*, respectively.

In order to generalize *wh* Movement Chomsky selects these four characteristics as a specific configuration and asks the following question:

> "Where we find this configuration in some system of data, can we explain it on the assumption that the configuration results from *wh*-movement?" (Chomsky 1977, 86).

To elaborate: *wh* Movement has five characteristic properties altogether; the fifth one being the occurrence or recoverability of a *wh*-phrase. What Chomsky asks is simply whether, if this fifth property is neglected and the other four are demonstrable, we can be justified in speaking of *wh* Movement.

2. The first domain of *wh* Movement to be examined is comparative constructions, a topic which has intrigued many

excellent minds. Even trace-theoretically oriented works on it alone fill hundreds of pages, mainly due to Joan Bresnan's interests and her dispute with Chomsky and others.

Comparative constructions used to be described as arising through an 'identity of sense' type deletion process more or less along the lines of (59a-b):[3]

(6) a. John is $[_{AP}$ strong $[$er $[_{S'}$ than $[_{S}$ Mary is strong$]]]]$

b. John is stronger than Mary (is).

Now on a structure similar to (6) Chomsky intends to demonstrate that the properties of *wh* Movement are observable:

(7) a. Mary isn't taller than [she was five years ago]

b. Mary isn't taller than [John believes [Bill claims [that she was five years ago]]]

c. *Mary isn't taller than [John believes [the claim [that she was five years ago]]]

d. *Mary isn't taller than [I wonder [whether she was five years ago]]

and draws an analogy between the sentences of (7) and those of (8):

(8) 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]]

Chomsky also asserts that certain dialects have the sentences of (7) with *what* in the clause under the same conditions of grammaticality.

As was suggested above, the ongoing discussion on comparative constructions indicates that their structure is far from clear at present. In fairness to Chomsky, however, it must be granted that there does occur a *wh*-word in at least some of the examples. Yet postulating that particular *wh*-word (*what*) is perhaps an unfortunate choice, since Chomsky's rule (47), cited here under (9), was originally meant to apply to ordinary relative constructions (with *who* or *which* in their COMP nodes), and not to 'headless' ones -- even though the relative pronouns may coincide in some of the cases (eg. *where*).

(9) *wh*-phrase becomes null

Rule (9) is part of a series whose task is, roughly speaking, to eliminate unnecessary constituents of the COMP node or even the whole node, and is therefore complemented by two more rules:

(10) a. *that* becomes null
for becomes null

of which one must apply since the COMP nodes of relative constructions are doubly filled as a result of *wh* Movement

into COMP, the only restriction being that complex *wh*-phrases (eg. those which have lexical content: prepositions, possessives) cannot delete.

But the extension of (9) with a stroke of the pen to strings introduced by *what* seems uncalled for or unjustified to say the least. For that amounts to claiming that *what* can be deleted anywhere, which surely is not the case. The only concession Chomsky makes in order to distinguish between *what*- and *which*-relatives concerns their sematic interpretation (Chomsky 1977, 92 and passim). Furthermore, in contradiction to his principle of *wh*-deletion, it seems that Chomsky does allow deletion of *wh*-phrases with lexical content as is clear from the following derivation. Sentences like (11) are supposed to derive from structures like (12), a procedure which evidently involves the deletion of lexically filled constituents:

(11) more students flunked than thought they would flunk

(12) more students flunked than [[*wh*-many (students)] [*t* thought [they would flunk]]]

Chomsky of course realizes the danger of assuming there to be *wh*-phrases where there is none to be found on the surface but shrugs off all possible objections by saying that "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 with the first over a subdomain of structures [...]" (89).

3. Even more interesting is the second alleged context of *wh* Movement, topic formation.[4] It comprises three types of structures: topicalization (13a), cleft-sentences (13b), and pseudo-clefts (13c):

(13) a. This book, I have read.

b. It is this book that I have read.

c. This book is what I have read.

The structures Chomsky assigns to (13a-b) are given under (14a-b), respectively:

(14) a. $[_{S''}\ [_{TOPIC}$ this book$]\ [_{S'}$ COMP [I have read what]]]

b. It is $[_{S''}\ [_{TOPIC}$ this book$]\ [_{S'}$ COMP $[_{S}$ I have read what]]]

where S" is a cyclic node as defined by the initial rules:

(15) R1: S" ⟶ TOPIC S'

(16) R2: S' ⟶ COMP $\{^{S''}_{S}\}$

In the (b) type sentences we can again discover a rather unusual rule of *wh*-deletion, but we are for the moment more interested in the (a) type sentences, since it is on the basis of these that the dubious rules of (15-16) are set up. Our first objection is similar to one made against Chomsky's derivation of comparatives; it differs only in that in the present case there is no material trace of the *wh*-phrase and the justification of the alleged *wh* Movement hinges entirely upon the presence or absence of the four characteristic properties of *wh* Movement. One might be inclined to think that

this analysis constitutes an exemplary case for Occam's razor. But then it may equally well be said this is a minor metatheoretical point with no empirical import.

Our next objection will then be more of an empirical nature. According to the rules of (15-16), S" is a cyclic node, a view supported in Chomsky's opinion by sentences such as (17):

(17) I believe that the books, John gave away to some friends.

However, Emonds (1970, 1976) expressly claims that Topicalization, which is the rule that produces sentences like (13a) and (17), is a root transformation and demonstrates it on the following sentences:[5]

(18) a. *I fear (that) [each part John examined carefully]

b. *We are going to the school play because [our daughters we are proud of]

c. *Are you aware (of the fact) that [poetry we try not to memorize]?

d. *Do you think [socialist theory many Czechs would deny]?

e. *That [this house he left to a friend] was generous of him.

The topicalizations in complement sentences are no better than those in relative clauses or questions, which are the only structures in which Chomsky disallows topicalization;

compare the following example from Emonds:

(19) *I have shown you the broom (that)
[these steps I used to sweep with]

Emonds' analysis refutes Chomsky's claim to the cyclicity of S", and consequently the generality of *wh* Movement in topicalizations and clefts, since only the latter can be truly cyclic.

Our last objection, which will be made more use of below, follows from a very simple fact: by no means can Topicalization involve the subject of a matrix sentence, although it is perfectly possible with the other two processes of topic formation, clefts and pseudo-clefts, cf.

(20) a. *this book, is interesting
b. It is this book that's interesting.
c. This book is what's interesting.

(21) a. *John, arrived yesterday
b. It is John that arrived yesterday.[6]

Now it is worth noticing that another type of topic formation, usually called Left Dislocation, does allow subjects as well as other complements to be topicalized:

(22) a. This book, it's interesting.
b. John, he arrived yesterday.

Observe also that (20a) will not improve if the alleged *wh*-phrase is retained:

(23) *the book, what is interesting.

For the sake of our final argument let us accept Chomsky's analysis, and suppose that in the following structure:

(24) $[_{S''}$ $[_{TOP}$ this book] $[_{S'}$ COMP $[_{S}$ I have written $[_{PP}$ about what]]]]

the whole PP node rather than just *what* moves into COMP, yielding (25):

(25) $[_{S''}$ $[_{TOP}$ this book] $[_{S'}$ $[_{COMP}$ about what] $[_{S}$ I have written *t*]]]

The question then is what can happen to the COMP node: it cannot be deleted, since it contains lexical material, but it must be -- for otherwise the result is ungrammatical. This dilemma cannot be quibbled by making use of an ad hoc filter: it could not choose between well-formed and ill--formed structures containing complex *wh*-phrases in COMP.

For the time being I have nothing to say about the next two topics in Chomsky's discussion of *wh*-phenomena, indirect questions and infinitival relatives, since no query can be raised about *wh* Movement within them, especially as far as questions are concerned.

4. Before turning to the last two contexts discussed in Chomsky (1977) a short digression on certain aspects of trace theory will be necessary. Elsewhere I reviewed in some detail the reasons for introducing, as well as the functions of, two new elements in the grammar:

empty nodes and traces (Kenesei 1979). Empty nodes are generated by the base rules and there is a rather simple mechanism to block surface structures which contain empty nodes. Traces in turn are the result of movement transformations, which mark them for identity, and are interpreted as anaphors of a special kind.

Now a third innovation is the node PRO represented by the terminal identity element *e* under NP. It is thus syntactically indistinguishable from empty nodes and traces and is marked for coreference by means of being coindexed with another constituent by rules of control in the semantic component. The notion of PRO is certainly akin to the deep pronoun hypothesis, with the important difference that PRO, unlike 'ordinary' pronouns such as *he*, *it*, etc., must undergo control or otherwise the structure will be semantically uninterpretable. PRO may be controlled by some other constituent or may be assigned arbitrary reference. Thus in the following three sentences, which are all structures of obligatory control:

(26) a. John asked Bill [what PRO to do]
b. John told Bill [what PRO to do]
c. It is unclear [what PRO to do]

PRO is coreferent with the matrix subject (26a), the matrix indirect object (26b), or is arbitrary in reference (26c).

The nodes $[_{NP}$ e] exhibit the common property of having no phonetic outcome. But while traces are on a par with the

lexical NPs they are coindexed with as regards filters (ie. none of the structures in (26) are grammatical if a lexical NP or its trace replaces PRO), PROs are invisible for filters: they simply do not count as NPs as far as filters are concerned. It follows, Chomsky and Lasnik argue, "that PRO and lexical NPs (including trace) are in complementary distribution in surface structure. Where we find PRO, we can find neither a lexical NP nor a trace, and conversely." (Chomsky and Lasnik 1977: 441)

5. The next context on which the effects of *wh* Movement are claimed to be observable is *too/enough* plus infinitival complements of adjective phrases as in

(27) John is tall enough [for us to see]

The sentence (27) is synonymous with (28) on one of its readings, that is, when *him* is coreferent with *John*:

(28) John is tall enough [for us to see him]

Structures of the type of (27), but not those of (28), exhibit the characteristics of *wh* Movement, most importantly the property of blocking extraction from complex NPs and questions, according to Chomsky, cf.

(29) a. *The job is important enough for us to insist on the principle that the committee should advertise.

b. *Who was the job good enough for us to offer to?

Consequently, the structure underlying (27) is, continues Chomsky, (30b), which is ultimately derived from (30a), whereas (28) essentially has the structure as shown by the brackets.

(30) a. John is tall [enough [$_S$, [for us to see who]]
b. John is tall [enough [$_S$, who for us to see *t*]

The structure (30b) will then undergo obligatory *wh*-deletion to get (27). But even Chomsky admits that multiple embedding of such structures under 'bridge' verbs results in sentences which are "dubious or starred", while their alternative sentences containing pronouns in place of the trace are "highly preferred," compare:

(31) a. ?the job is important enough for us to order them to insist that the committee should offer *t* to John
b. the job is important enough for us to order them to insist that the committee should offer it to John

But even if we ignore this and other minor difficulties, such as the unexplained occurrence of *which* instead of earlier *what*, there still remains ample room for criticism.

The complement subject, it is claimed, must always be PRO, thereby distinguishing between (32a) and (32b-c), which underlie ungrammatical sentences:

(32) a. John was clever [enough [for PRO to run away]]
(⟶ John was clever enough to run away)

b. *John was clever [enough [for him to run away]]

c. *John was clever [enough [for who to run away]]

Incidentally, unless the above restriction about the embedded subject were in force, (32c) would underlie the grammatical sentence derived from (32a) after *wh* and *for* deletions both have applied. In other words there is no independent motivation for such a restriction except that only in this case would the insertion of a *wh*-phrase and that of a 'real' pronoun result in surface sentences of distinct grammaticality. Besides, the requirement that the complement subject always be PRO (cf. Chomsky 1977: 103) is blatantly false in view of sentences like *John is too short for Peter to see*, so it must be rephrased as a prohibition on the occurrence of *wh*-phrases in subject position. That alone would suffice to justify the revision of the initial assumption, but there are other reasons too.

If, in turn, there is a *wh*-phrase in the complement sentence alongside with a subject PRO, this PRO will now have to be blocked from being interpreted as coreferent with the matrix subject, cf. (33a-c):

(33) a. John is clever [enough [for PRO to understand who]]

b. John is clever [enough [who for PRO to understand *t*]]

c. John is clever [enough [to understand *t*]]

And while that is not impossible, relying on compatible data like (34):

(34) the man [who he thought [*t* would win]]

where the *wh*-phrase, but not *he*, can be associated with the head NP *the man*, it is at variance with Chomsky's treatment of infinitival relatives, in which embedded subjects are represented by *wh*-phrases -- all to be deleted subsequently; cf. the derivation (35a-c):

(35) a. [the man [for who to fix the sink]]
b. [the man [who for *t* to fix the sink]]
c. [the man [to fix the sink]]

In short, if we want to avoid inconsistency, either all slots (whether subject or complement) are filled in by *wh*-phrases in *too/enough* plus complement constructions or none of them are. We have shown that the former option entails an erratic relationship between 'real' pronouns and *wh*-phrases in these structures, whereas the latter would amount to giving up the idea of a general *wh* Movement transformation.

Another difficulty arises in connection with *wh*-extraction from the embedded sentence. The sentence (39b) is certainly ungrammatical, but questions formed of the analogous construction which contains a pronominal anaphor are not a trifle better, although Chomsky's analysis predicts that they should be grammatical; compare (36) below with

(29b):

(36) *Who was the job good enough for us to offer it to?

Since there is no *wh*-phrase in the embedded sentence, there can be nothing to prevent the movement of *who* into the matrix COMP node.

It is perhaps becoming increasingly clear that the configuration meant to characterize *wh* Movement (cf. 1.) simply does not fit the data (with the possible exception of comparatives). But, in accordance with what may be called Chomsky's principle of grammatical argument,[7] we still have the problems of (32-35) as well as the ungrammaticality of (36) to account for. As far as the structures (32-35) are concerned, a possible solution could be conceived by invoking Equi-deletion or analyses of similar effect. Suppose, for example, that underlying (37a-b) are (38a-b), respectively:

(37) a. John is tall enough to climb the wall.
b. John is tall enough for him to climb the wall.

(38) a. John is tall [enough [for PRO to climb the wall]]
b. John is tall [enough [for him to climb the wall]]

where in (38) PRO undergoes the usual assignment of control and is in effect deleted together with *for*, while in (38) *him* will be marked as non-coreferent with the matrix subject *John* by the rule of disjoint reference (Chomsky 1976). This is a distinction at work in essentially similar structures, cf.:

(39) a. John was eager [for PRO to win]

b. John was eager [for him to win]

That is, strictly speaking, a non-deletion approach. One could imagine replacing PRO with PRO-*self* and rely on Equi--deletion, but such alternatives do not affect the point made here. The only consequence one must take into account is whether the NP Constraint is construed as one which ranges over transformations only or one which may involve semantic interpretation as well. If the latter option is accepted, PRO embedded in a complex NP (as in (40)) will be assigned no control by a constituent outside that NP.

(40) *The job is important enough [for us to insist on [$_{NP}$ the principle [$_{S}$, that they should offer PRO to John]]]

Besides, the Specified Subject Condition (or its equivalent) can also do the job.[8]

Such a broad interpretation of a constraint on transformations is not at all alien to trace theory. Indeed the very first conditions on transformations were formulated more or less in this fashion (Chomsky 1973), and some of them have survived since. Assuming PRO to replace *wh*-phrases will have other additional advantages. For example, Chomsky explains in another connection that the paradigm in (41) demonstrates the workings of *wh* Movement in that whenever a *wh*-phrase has been moved from an indirect object into COMP, it must leave the preposition *to* stranded in the original position, that is, no Dative Movement is possible prior to *wh* Movement (Chomsky

1977:104):

(41) 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.

(Note that the last two sentences are meant to be synonymous.)

That is all right for American English. However, in British English (41b) and related constructions pass the grammaticality test with no difficulty, yet sentences like (41e) are considered unacceptable. In other words, a general constraint on *wh*-phrases in indirect object position which prohibits their participation in *wh* Movement will not do here. In British English the constraint must involve PRO only and we have no reason to suppose that the situation is markedly different in American English, where, we may conjecture, there can be an additional constraint on *wh*-phrases.

We have seen that none of the *too/enough* plus complement constructions allow question words to be extracted, whether or not they contain anaphoric pronouns, cf. (29b) and (36). But this is not a property peculiar to the structures under discussion. Consider the following sentences:

(42) a1. John was so happy that he sang madrigals.

a2. *What was John so happy that he sang?

b1. John polished the floor so hard that you could see your face in it.

b2. *What did John polish the floor so hard that you could see in it?

c1. Mary left early to catch the train.

c2. *What did Nary leave early to catch?

No matter what their exact constituent structure may be, the sentences in (42) all contain some kind of adverbial complements in the form of infinitival or tensed clauses -- and all of them prohibit the extraction of *wh*-phrases. That is to say, we may as well rely on *wh* Movement in order to explain the ungrammaticality of (29b) and (36), but we still have to face up to the ungrammaticality resulting from the extraction of *wh*-phrases from adverbial clauses. Thus we can rather contemplate another general restriction, this time a prohibition on the movement of any material out of sentences embedded in adverbials.[9]

6. Due to lack of space I will now give a rather cursory review of the last alleged domain of *wh* Movement. The phenomena referred to as *tough* Movements have traditionally puzzled a number of outstanding linguists. In the previous discussions of sentences like (43a-b) I followed their customary analysis, according to which (43b) is derived from the structure immediately underlying (43a)

through the movement of the embedded object NP into the matrix subject position (cf. *enough* etc. and complements above):

(43) a. It is easy (for us) to please John.
b. John is easy (for us) to please.

We can now proceed, without further ado, to examine the structures Chomsky claims to underlie the sentence in (43b). For any arguments supposedly in favour of the analysis the reader is referred to Chomsky (1977:103ff). Thus in (44)

(44) a. John is easy (for us) [$_S$, for PRO to please who]
b. John is easy (for us) [who for PRO to please *t*]

(44b) is derived from (44a) through *wh* Movement into COMP followed by *wh* and *for* deletion, ultimately yielding (43b).

No argument will be advanced here about the restriction of the occurrence of *wh*-phrases in the embedded subject position -- although threefold requirements of the kind invoked in these constructions (demanding the insertion of a *wh*-phrase in complement position, PRO in subject position in addition to a *for* complementizer) are certainly rare. Nor will I discuss the legitimacy of stretching the deep subject -- predicate relationship so far as to claiming that *easiness* is predicated of *John* or any other subject in the same way as *happiness* or *readiness* is in the sentences: *John is happy (for us) to leave., The soup is ready (for us) to eat.*

But how can the problem sentences of the following sort raise resolved?

(45) [$_S$, that the theorem is true] is difficult to prove

Recall that Emonds (1970, 1976) made out a good case for a root transformation of Intraposition to handle structures of this kind, a view fully accepted by Chomsky (for details see Chomsky and Lasnik 1977). If in (45) the bracketed S' was originally not adjacent to *prove*, it follows that it must have been intraposed from somewhere outside the sentential complement of *difficult*, as in the derivation (46a-b):

(46) a. NP is [$_{AP}$ difficult [$_S$, for PRO to prove which] [$_S$, that the theorem is true]]

b. [$_S$, that the theorem is true] is [difficult [$_S$, for PRO to prove which]]

However, for that or any other similar analysis to be correct there ought to be an analogous construction containing *it* as matrix subject, cf. the following examples:

(47) a. NP is illegal for John to leave

b. It is illegal for John to leave.

c. NP was proved that the theorem was true

d. It was proved that the theorem was true.

But no alternative exists for (46a), since the parallel sentence containing the 'dummy' *it* must be derived from something like (48), according to Chomsky:

(48) It is [difficult [for PRO to prove [that the theorem is true]]]

Such constructions are claimed to be syntactically unrelated to their 'topicalized' counterparts, ie. (45) in this case.

That is rather a paradoxical situation for any proponent of general *wh* Movement who is opposed to postulating sentences in subject position. The only way out seems to be to disregard Emonds' findings and allow S' nodes to occur in subject position.[10] This possibility is argued forcibly by Halitsky (1975), who maintains that there is no symmetry in the X-Bar Theory as developed by Chomsky (1970) and demonstrates that rules of the form:

(49) [Spec, V'] ---→ V''

must also be possible if there is already a rule [Spec, N']---→ N''. The rule (49) would be capable of generating tensed and infinitival sentential subjects on the one hand, and, on the other, these sentential subjects would not be dominated by the category NP, thereby reconciling Rosenbaum's (1967) and Emonds' (1970, 1976) positions.

Although Halitsky's idea is absolutely feasible, it will contribute nothing whatever to solving the dilemma of (46). It is an immediate consequence of Halitsky's amendment (49) that the rule of Extraposition is resurrected as a structure preserving (cyclic) transformation, so the derivational relationship between the relevant structures *NP--Pred--S*

and *S--Pred* is, as it were, reversed with the latter now underlying the former. If then we were to believe that (46b) ultimately underlies (45), we would still have to account for the impossibility of Extraposition in (46b), since after Extraposition in (46b) the result is something like (46a), in which the *wh*-phrase *which* precedes its antecedent. But the sentence which approximately has the structure of [48] must somehow be obtained. Thus the paradox prevails and cannot be resolved in the framework of general *wh* Movement.[11]

We will now continue discussing *easy* plus complement constructions by citing an example from Chomsky (1973). One of his counterarguments against Raising was based on the asymmetry between (50a) and(50b). At that time these two sentences served to show that the surface subject *Smith* in (50b) cannot have been the object of *expect* at any point of the derivation (including the one prior to its movement into matrix subject position), since it does not behave as a proven object of an embedded verb, eg. *Smith* in (50a). But having changed his view of the underlying structure of these constructions, Chomsky must now find a different reason why (50b) is not grammatical. According to him, (50a) is derived from (50c), so the structure underlying (50b) must be (50d):

(50) a. Smith was easy for Jones to force to recover.

b. *Smith was easy for Jones to expect to recover.

c. Smith was easy for Jones [for PRO to force who [PRO to recover]]

d. Smith was easy for Jones [for PRO to expect

[who to recover]]

In order to block (50b), *wh* Movement in the embedded sentence must be prohibited. However, the structures in (51) show that *wh* Movement is perfectly possible in the relevant contexts:

(51) a. Who do you expect to recover?
b. the man (who) I expect to recover

The only solution Chomsky's framework allows would follow from the requirement that "the underlying structure must contain an embedded $\bar{S}$ as complement to *easy*, with an obligatory PRO subject, as in the case of infinitival complements [of *too* and *enough*] already mentioned." (Chomsky 1977, 103) That restriction is necessary to block the occurrence of *wh*-phrases in embedded subject position, thus preventing ungrammatical sentences such as (52b) from being generated:

(52) a. John is easy [for who to talk]
b. *John is easy to talk

This is simply a new way of stating the familiar restriction on the transformation *tough* Movement, *viz.* that it is not permitted to operate on embedded subjects. Suppose that we extend this restriction to subjects in 'lower' sentences saying that however deeply embedded the subject is, it cannot be moved if it is a *wh*-phrase in the complement to *tough*-predicates.

Putting aside all possible reservations which may arise from formulating the constraint and conceding that it would handle one type of irregularity, let us now see another that no constraint mentioning subjects could deal with. Neither of the pair of sentences (53a-b) seems to go through, although both of (53c-d) are grammatical:

(53) a. *The police are easy to arrest us.
b. *The police are easy for us to be arrested by.
c. It is easy for the police to arrest us.
d. It is easy for us to be arrested by the police.

Even if all proposed analyses of the constructions in question can block (53a), none of them (including Chomsky's) can prevent (53b).

One possible explanation we can envisage would rely on referring to the information 'deep subject', since in the cases of (103b),(106b), and(107b) it is apperently the deep subject of some one of the embedded sentences in the complement to *easy* that cannot be associated with the surface subject of *easy*. Owing to its neutrality with respect to the existence of Raising, this solution will, incidentally, deprive (50a-b) of their value as a counter-example to Raising. Now whether *easy* plus complement constructions involve a transformational or an interpretive rule is, I believe, a question that remains to be decided.[12]

To close this section I will put forward a perhaps complicated though persuasive argument based on problems

of interpretation as related to syntax. Hankamer and Sag (1976) demonstrate among others that null anaphora can arise in either of two ways: in the case of 'deep anaphora' the anaphor is not derived transformationally, but is present in underlying representations and is interpreted in deep structure; in the case of 'surface anaphora' the anaphor is derived transformationally by deletion and is interpreted in surface structure. Deep anaphora must represent semantic rather than syntactic units and, especially "in the case of sentential deep anaphora [...,] any attempt to assign an interpretation to a sentential deep anaphor at a later than precyclic stage runs afoul of the possible disintegration of its antecedent." (423)

It is exactly sentential deep anaphora that we have to do with in the case of (54a) as against (54b), in which the anaphor arises through syntactic deletion (for discussion see Hankamer and Sag 1976:414).

(54) a. John thought I must have passed the test but I didn't even try.

b. John thought I must have passed the test but I didn't even try to.

In the following sentence only deep anaphora is possible:

(55) John is too difficult to please for us to try ($\#_{to}$)

According to Chomsky's treatment of these constructions

utilizing Hankamer and Sag's observations, the deep structure of(54) is (56a), from which the surface structure (11b) is derived through *wh* Movement:

(56) a. $[_{S_0}$ John is [too [difficult $[_{S_1}$ for PRO to please who]]
$[_{S_2}$ for us to try $[_S$ PRO]]]]
b. $[_{S_0}$ John is [too [difficult $[_{S_1}$ who for PRO to please *t*]] $[_{S_2}$ for us to try $[_S$ PRO]]]]

Hankamer and Sag's theory of anaphora can be reconciled with more recent developments of trace theory, since the antecedents of deep anaphors no longer can be said to disintegrate due to the effect of traces and the fact that all deletions are postponed until after surface structures receive semantic interpretation. But even so, *wh* Movement has a weak case here. The interpretation of the surface structure (56b) can choose between the following two options. It can first assign control to the (free) sentential anaphor in S_2 by, as it were, substituting S_1 for $[_S$ PRO]. In this case the *wh*-phrase must remain uninterpreted in the sentential anaphor. Or it may first assign control to both the *wh*-phrase and PRO in S_1 and only then 'substitute' S_1 for $[_S$ PRO]. In this case the PRO subject in S_1 has already been assigned arbitrary reference, so the subject PRO in the sentential anaphor cannot now be controlled by *us*, the matrix subject in S_2, which again renders (56) ungrammatical.

The alternative solutions will face none of these problems. If we consider (54) as the result of the trans-

formation *tough* Movement from the structure (57)

(57) NP is [too [difficult [$_{S_1}$ for PRO to please John]] [$_{S_2}$ for us to try [$_S$ PRO]]]

through a transformation which places *John* in the matrix subject and leaves a trace in its original position, we can see that there is an easy way to associate [$_S$ PRO] with S_1, since the object of *please* is the trace of *John*, a bound anaphor, while the subject of S_1 is the unbounded anaphor PRO, which can be controlled by the S_2 subject *us* in [$_S$ PRO], and assigned arbitrary control in S_1.

Another possible solution postulates a lexically filled matrix subject and a PRO in place of *John* in S_1. The assignment of control is relatively straightforward, but its description is too long to be included here.[13]

We have seen that neither Chomsky's four criteria nor the data he relies on are sufficiently clear to prove that the transformation *wh* Movement can indeed account for the phenomena presented. Furthermore we have come across a number of misstatements or perhaps even confusions, such as the cyclicity of Topicalization or the insertion of PRO in subject but *wh*-phrases in object positions when the resulting sequences are indistinguishable. A number of counterarguments were given, which in the important cases lean towards a reformulation of Chomsky's analyses in terms of the referentially empty constituent PRO, especially as regards the last two sets of data (5. and this section). However the notion of PRO

to be adopted here is somewhat different from what seems to be generally accepted.

7. PRO was characterized in 4. as represented by the terminal identity element *e* and is therefore not distinct syntactically from empty nodes and trace. But this last statement is a little too crude. A moment's reflection will show that empty nodes and trace cannot be moved but PRO can. For if empty nodes were moved the operation would be pointless, since they could only be transposed to other empty nodes according to the requirements of the structure preserving constraint, and such an operation would produce a tree identical with the input. If a trace were moved trace theory itself would become useless, since the essential information traces convey would be lost. But PRO does move around; no matter whichever analysis we subscribe to in the cases discussed, the derivation of (58a-d) must involve the movement of PRO.

(58) a. John persuaded Bill [the doctor to be examined PRO by [$_{NP}$ e]]

b. John persuaded Bill [t_1 to be examined PRO by [$_{NP_1}$ the doctor]]

c. John persuaded Bill [PRO to be examined by [the doctor]]

d. John persuaded Bill to be examined by the doctor.

If there were another [$_{NP}$ e] in (58a), then of course the

movement of the NP containing PRO into the embedded subject would be skipped over, that is, (58c) would be omitted, and instead we would get something like (58b') :

(58) b'. John persuaded Bill [t_1 to be examined [$_{NP}$ e] by [$_{NP_1}$ the doctor]]

which underlies no grammatical sentence for at least three reasons. First of all, the NP coindexed with its trace is preceded by it, then there is an uncontrollable empty node in the structure, and finally the rule of control finds no candidate for PRO in the embedded subject position as is required by verbs of the *persuade* type. Thus even if the constituent [$_{NP}$ e] is regarded as PRO, the structure it is in will have to be starred. Before we analyse the consequences of the movement of PRO we will discuss the criterion of the insertion of PRO. Recall that Chomsky and Lasnik (1977) rely on the thesis of complementary distrubution of PRO and lexical NPs or trace to determine the positions PRO can occupy (cf. 4.). But the thesis has certain drawbacks.

First, it does not hold for the constructions treated in 5. as is clear from Chomsky's own analysis of the data (1977:100ff) and in the face of evidence such as (59a-b):

(59) a. John is too short [for PRO to see us]
b. John is too short for us to see him]

Then the thesis would require us to postulate PRO in some

of the constructions which, according to Chomsky, contain *wh*-phrases:

(60) a. *The chicken is ready for us to eat {the fish / it}

b. *John is easy (for us) to please {Bill / him}

Clearly, no lexical NP can occur in the curly brackets.

The trouble with the criterion is that Chomsky and Lasnik try to overgeneralize their findings in constructions like *persuade NP [PRO to VP]* in order to cover the domain of the earlier rule of Equi-NP Deletion. But some rule of Equi is still retained and operates on PRO-*self* NPs. These reflexives can occur wherever full lexical NPs do. Their deletion by Equi, however, leads to quite an unorthodox concept of deletion in cases like the following:

(61) It is illegal for PRO-self to leave

where PRO-*self* is necessary because lexical NPs can also occupy that position and after "the rule of control [...] assigns arbitrary reference to PRO [...] PRO-*self* deletes by Equi" (Chomsky and Lasnik 1977:472). Obviously, if there is any deletion here it is not by Equi, there being no other identical lexical node, but it is a free deletion, which must be rare or perhaps illegitimate in trace theory on account of the principle of recoverability of deletion.[14]

Now if PRO can be moved about, it is neither syntactically nor semantically on a par with the nodes represented by $[_{NP}\ e\ ]$. If then the thesis of complementary

distribution is refuted, PRO must be capable of entering any syntactically appropriate position and the grammaticality of the structures formed in this way will depend on the control requirements of individual verbs and adjectives and other relevant constituents in the matrix sentence which are in construction with the embedded sentence containing the PRO. It follows then that (115) can contain PRO instead of PRO-*self*, and conceivably (114a-b) also have PRO in place of the curly brackets.

PRO then will be distinct from empty nodes and traces in that it is subject to movement rules and requires control. In other words, it is an unbound anaphor with null phonetic outcome, and as regards movement rules (but not filters) on a par with lexical NPs. It is the control requirement that blocks both surface structures (52a-b), since there is no possible controller in them:

(52) a. PRO is hit John by NP
b. John is hit *t* by PRO

It follows that there may also be structures which contain two PROs in one clause, like (53), which underlies the sentence (53b):[15]

(53) a. The chicken is too hot [for PRO to eat PRO]
b. The chicken is too hot to eat.

For the rules of control two options are open: they may associate the matrix subject with the subject PRO or the ob-

ject (or in general the complement) PRO in the embedded sentence. If they take the first choice, the object PRO will be left uncontrolled. If they choose the second, the object PRO will be controlled by *the chicken* and the way is open for the embedded subject PRO to be interpreted as having arbitrary reference. Note that the assignment of arbitrary reference must observe (at least) two prerequisites: one is that the PRO in question be the subject of an infinitival construction, the other that the constituent commanding the infinitival construction be specified for allowing such assignment; eg. *unclear, too, enough, illegal,* as against *know + wh, force, happy*.

N O T E S

Why is the only question-word that may be used in matrix questions with non-finite (infinitival) predicates, eg.,

(i) Why bother about John?

On the other hand, it is the only *wh*-word which cannot be used in embedded infinitives, eg.,

(ii) *He doesn't know why to leave.

2 In his original formulation, Chomsky (1977.86) quotes the Complex NP Constraint (cf. Ross 1967). For us, any distinction is indifferent for the time being insofar as the constraint mentions deletion in addition to extraction.
Bach and Horn's NP Constraint is, incidentally, as follows:

> *"No constituent that is dominated by an NP can be moved or deleted from that NP by a transformational rule."* (Bach and Horn 1976:280)

3 According to one typology of deletion phenomena (Grinder 1976), Identity-of-Sense Deletion comprises several subcases besides Comparative Deletion, such as Gapping (i), Sluicing (ii):

(i) John killed the man and Mary --- the woman.
(ii) I know someone here plays tennis but I don't know who --- .

Other types of deletion are based on coreferential identity (eg., cases of Equi), or are free deletions.

4 'Topic-formation' is meant here as a term general enough to cover related phenomena; although there is no such label in Chomsky (1977), the three processes

taken here to belong under Topic formation are treated more or less in conjunction by him.

5 Grammaticality judgments are from Emonds (1970, 24), the square brackets have been added to show the domain of topicalization.

6 The pseudo-cleft variety is not given in (21.) since it is ungrammatical for a different reason, namely, the restriction on the distribution of *who*.

7 "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 grammar. *The confirmation or refutation will be as convincing as the fragment of rules presented.* [...] 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'." (Chomsky 1977, 74; emphasis added)

8 The original formulation of the Specified Subject Condition is this:
"No rule can involve *X*, *Y* in the structure
... *X* ... [$_{\alpha}$ ... *Z* ... -- *WYV* ...] ...
where Z is the specified subject of *WYV* in α." (Chomsky 1973, 239)

9 In addition, there are other structures of adverbial type which corroborate our analysis of *too/enough* plus complement constructions, since they also allow deletion of certain complements in the embedded sentence, eg.,

(i) The two chairs were brought in for us to sit on (them).

Neither does this construction allow question-word movement out of the embedded sentence, yet it would be futile to suppose that there is some kind of latent *wh* Movement underlying the derivation. Postulating PRO would surely be much simpler and closer to actual facts of grammar.

For the record it should be noted that there can be an independent reason for preventing sentences like (41a2-b2-c2) from arising. Ray Cattell (1976) proposes two constraints which will in effect limit the number of possible arguments, including *wh*-phrases, within the domain of the predicate. The constraint of immediate relevance here, ie. the Overcrowding Principle, runs as follows: "The number of NPs within a syntactic configuration must not exceed the number of argument places available for the verbs within it."

10 However, Chomsky and Lasnik (1977), who discuss the relevant structures at some length, take no cognizance of this option.

11 Even though Halitsky (1975) does not allow of any ambiguous interpretation, it may be asked in defence of Chomsky whether the assumption that there are both sentential and dummy NP subjects while S's are generated in a predicate-final position would help solving the problem. I think it would not -- at least not without a major modification of trace theory. For one thing, passivization should be completely altered. The now canonical treatment of passives takes (i) as the underlying form of (ii):

(i) NP $[_{VP}$ was $[_{AP}$ en $[_{VP}$ $[_{V}$ hit $[_{NP}$ Bill]]]]]
(ii) Bill was hit.

Observe that the object NP of *hit* is in the same 'lower' VP. Consequently, the deep structure for (iii) can be

nothing but (iv):

(iii) That the theory was true was disputed.
(iv) NP $[_{VP}$ was $[_{AP}$ en $[_{VP}$ $[_{V}$ dispute] $[_{S}$, that the theory was true]]]]

where S' has the same complement status with respect to *dispute* as NP with respect to *hit*. If we opt for having S' as subject in (iv), which should be possible in this version of Halitsky's amendment, we must also change (i) by making *Bill* the deep subject. But then nothing will be left of the idea that traces reflect 'thematic' relations like 'agent', 'goal', etc. By letting the grammar decide which way to go about it, both (i) and (v) will underlie (ii):

(v) $[_{NP}$ Bill] [was [en $[_{VP}$ hit $[_{NP}$ e]]]]

just as (iii) can then be derived from both (iv) and (vi):

(vi) $[_{S}$, that the theory was true] [was [en [dispute $[_{S}$, e]]]]

The futility of such a position will even be clearer if we consider the multiple results of movement transformations on these double deep structures.

12 Since the preceding counterargument overlaps discussions of Raising, I have also made use of it in Kenesei (1982).

Another point worth noting is the ungrammaticality of the following sentence (from Chomsky 1973, 263):

(i) *John is fun to see pictures of.

although *wh* Movement on *picture*-nominals is generally possible, cf.:

(ii) Who did you see pictures of?
(iii) the man I saw pictures of

In view of Chomsky's analysis of constructions involving *picture*-nominals in Chomsky (1973, 114ff), where, as a consequence of earlier analyses, he makes the implicit claim that (i) must be grammatical, this very example represents another instance of internal contradiction.

Again, Chomsky (1973, 265) confronts (iv) and (v):

(iv) Books are difficult to believe that Tom reads.
(v) This is the book that it is difficult to believe that Tom read.

and adds that (iv) is blocked, implying that it is ungrammatical, while (v) goes through owing to multiple *wh* Movement. Now recall that according to Chomsky (1977) the sentence (iv) is grammatical, though 'marginal'.

13 These are the outlines of the solution. The surface structure of (54) is given under (i):

(i) $[_{S_0}$ John is [too [difficult $[_{S_1}$ for PRO to please PRO]]
$[_{S_2}$ for us to try $[_{S}$ PRO]]]]

The S_1 and the S_2 cycles will be passed. On the S_0 cycle there are two ways to assign control to $[_S$ PRO]:

(a) by first taking it to be identical with S_1, cf. (ii):
(ii) [John is [too [difficult $[_{S_1}$ for PRO to please PRO]]
$[_{S_2}$ for us to try $[_{S_3}$ for PRO to please PRO]]]]

Then the subject PRO in S_3 is controlled by *us* in S_2, and the object PRO is assigned control independently of S_1, according to the requirements of *too* plus com-

plement constructions, ie. it will be controlled by *John*. The subject PRO in S_1 is of course assigned arbitrary reference, while the object PRO, according to the rules applying to *difficult* plus complement constructions, will be taken to corefer with *John*; or

(b) it may happen that the two PROs in S_1 are first assigned control, and $[_S$ PRO] undergoes the assignment of control only afterwards. In this case the structure will be uninterpretable since the subject PRO in S_1 is assigned arbitrary reference, which, if taken over to the subject of S_3, will result in an ungrammatical sentence.

14 That is, free deletion of lexical constituents. Chomsky and Lasnik (1977) allow the deletion of arbitrary structures in arbitrary categories only if they are in COMP, which makes it possible for them to delete *wh*-phrases in COMP.

Thinking over the role of PRO-*self*, it might even turn out that since in effect PRO and PRO-*self* are also in complementary distribution, they are the representations of one and the same element. On the other hand, the free insertion, free deletion, and free interpretation (due to the assignment of arbitrary reference) of PRO-*self* makes it a uniquely powerful entity in the grammar. In other words, with PRO-*self* anything goes.

15 The other source for the ambiguous (53b) contains no object in the embedded sentence: *The chicken is too hot* [*for PRO to eat*].

R E F E R E N C E S

Bach, Emmon, and George G. Horn. 1976. "Remarks on 'Conditions on transformations'". *Linguistic Inquiry 7*: 265-299.

Cattell, Ray. 1976. "Constraints on movement rules". *Language 52*: 18-52.

Chomsky, Noam. 1970. "Remarks on nominalizations." In: Roderick A. Jacobs and P.S. Rosenbaum (eds.) *Readings in English transformational grammar*. Waltham: Ginn. 1970. 184-221.

Chomsky, Noam. 1973. "Conditions on transformations". In: Stephen Anderson and P. Kiparsky (eds.), *Festschrift for Morris Halle*. New York: Holt. 1973. 232-286.

Chomsky, Noam. 1977. "On *wh*-movement". In: Peter W. Culicover et al. (eds.), *Formal syntax*. New York: Academic Press. 1977. 71-132.

Chomsky, Noam, and Howard Lasnik. 1977. "Filters and control". *Linguistic Inquiry 8*: 425-504.

Emonds, Joseph E. 1970. *Root and structure preserving transformations*. Unpublished doctoral dissertation. Cambridge: M.I.T.

Emonds, Joseph E. 1976. *A transformational approach to English syntax*. New York: Academic Press.

Grinder, John T. 1976. *Ondeletion phenomena in English*. The Hague: Mouton.

Halitsky, David. 1975. "Left branch S's and NP's in English". *Linguistic Analysis 1*: 279-296.

Hankamer, Jorge, and Ivan Sag. 1976. "Deep and surface anaphora". *Linguistic Inquiry 7*: 391-428.

Kenesei, Istvan. 1979. "Transformations of generative grammar". *Computational Linguistics and Computer Languages 13*: 57-87.

Kenesei, Istvan. 1982. "The reincarnation of Raising". In: F. Kiefer (ed.), *Hungarian general linguistics*. Amsterdam: J. Benjamins (to appear).

Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. Cambridge: The M.I.T. Press.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Unpublished doctoral dissertation. Cambridge: M.I.T.