Remarks and Replies

On Some Rules That Are Not Successive Cyclic

Paul M. Postal

A. Background

A grammar that contains transformational or transformation-like rules must be subject to metaconditions of applicability of various sorts[1,2] determining how various rules apply, especially in relation to each other. Among such conditions are those which would, for example, have rules linearly ordered, partially ordered, or unordered. However, overriding "traffic rules" for different classes of rules have also been proposed. These include:

(1) a. conditions specifying a class of precyclic rules
b. conditions specifying a class of postcyclic rules
c. conditions specifying a class of last cyclic rules
d. conditions specifying a class of cyclic rules

It has been and continues to be a subject of considerable disagreement whether rules meeting some or all of these sets of requirements exist in natural languages. Most everyone would agree today, I think, that no evidence for precyclic rules exists. Similarly, most would agree that there is some strong evidence that a class of cyclic rules exist. The existence of either postcyclic or last cyclic rules seems likely, if less clear.

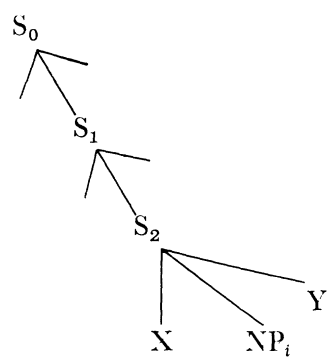
[1] I am grateful to G. Lakoff, W. Plath, J. Ross and several anonymous critics for comments on an earlier version of this paper. Needless to say, responsibility for the present contents is entirely the author's.

[2] The clumsy formulation "transformational or transformation-like" is necessary because there is now considerable evidence that many rules are not properly speaking transformations, in that they involve either or both global or transderivational environments. Compare Lakoff (1970; to appear), Postal (1971). More generally, almost all of the evidence taken to show that rules are transformational shows at best really only that they are at least transformational; that is, that the facts cannot be described with weaker devices. But evidence that the theory of transformational rules provides an upper bound on rule power is not only nonexistent but, as noted above, in reference to global and transderivational environments, evidently impossible.

As Emmon Bach has stressed in unpublished work, however, the normal account of cyclic rules covers up an important ambiguity at least with respect to one class of rules. If one considers those rules like English Wh Relative Movement (Wh Rel Movement), Wh Question Movement (Wh Q Movement), and Topicalization (henceforth, U-Rules), which have the property of reordering NP constituents over an unbounded section of tree structure, the notion of cyclic application can interact with particular formulations of the rules themselves to yield quite contrastive modes of application. One formulation consistent with cyclic application would be that on every cycle the NP is moved to the front of the clause defining that cycle, ending up finally in its maximally fronted position after movement on the last cycle. I shall refer to this mode of cyclic application as *successive cyclic*.[3] Under another conception, application would be cyclic, but the constituent would move only when some element higher up in the tree which "triggers"[4] the movement is reached in the cycle. Under the latter formulation, even though a very deeply embedded constituent were moved by a cyclic rule, it might not be reordered until a very late cycle. Let us call this mode of application *higher-trigger cyclic*.

The contrast between these two modes can be illustrated as follows:

(2)



Given an underlying structure like (2), which is supposed to be the basis for a surface structure in which the correspondent of NP_i ends up in S_0, as a function of the application of some U-Rule like Wh Q Movement, the two contrasting modes of application would claim the following. Under successive cyclic application, NP_i would be moved to the front of S_2 on the S_2 cycle, moved to the front of S_1 on the S_1 cycle, and finally moved to the front of S_0 on the S_0 cycle. Hence to move up into the highest clause,

[3] Although, as noted, several authors advocate a successive cyclic account of such rules, I am aware of no explicit statement of rules in this framework. Certain nontrivial problems immediately arise in such a formulation; in particular, how a statement is to be given formally indicating where the element which must move *n* times is to stop. The problem is considered further at the end of Section C.

[4] This would assume a general theory in which all unbounded left movement rules are uniquely triggered by some element in the tree. I believe this is a plausible hypothesis but it is irrelevant to the arguments of this paper.

three applications of the rule are necessary. Further, NP_i does not remain in its original position at the end of any cycle on (2). Under the higher trigger cyclic mode, however, the triggering element would have to exist in S_0, and NP_i would move to the point of that trigger (to the left or right as the rule indicates) in a single movement on the S_0 cycle. Hence in this case NP_i would remain in its original position at the end of both the S_2 and S_1 cycles, unless other rules apply.

It is an important descriptive and theoretical problem to determine whether actual U-Rules like English Wh Rel Movement and so forth are cyclic or not, and if so, whether they are successive cyclic, higher trigger cyclic, or some other variety. In what follows I shall give grounds for concluding that none of the U-Rules can be successive cyclic.

B. The Preposition Dangle Argument

In the 1968 version of my *Cross-Over Phenomena*[5] I provided an argument that Wh Rel Movement and Wh Q Movement could not be successive cyclic, although I did not use this terminology. The argument was that prepositions optionally accompany NPs moved by this rule, and that successive cyclic application would predict that the preposition could be left behind at any of the intermediate sentence initial positions, when in fact the only possibilities are that the preposition either stays in its original position or moves all the way to the front along with the NP which is moved:

(3) a. I believe Mary thinks Joan talked to someone.
b. Max, who I believe Mary thinks Joan talked to . . .
c. Max, to whom I believe Mary thinks Joan talked . . .
d. *Max, who/whom I believe to Mary thinks Joan talked . . .
e. *Max, who/whom I believe Mary thinks to Joan talked . . .

(4) a. I believe Mary thinks Joan talked to someone.
b. Who do you believe Mary thinks Joan talked to?
c. To whom do you believe Mary thinks Joan talked?
d. *Who/Whom do you believe to Mary thinks Joan talked?
e. *Who/Whom do you believe Mary thinks to Joan talked?

I regarded this Preposition Dangle argument as decisive. But some linguists, including Chomsky and Jackendoff, who wish to maintain that these rules are successive cyclic for various reasons,[6] disagree. Thus Jackendoff (1969, 51), after giving various

[5] In *Specification and Utilization of a Transformational Grammar* (Scientific Report No. 3, IBM Research Center, Yorktown Heights, New York. Compare now the published version (Postal 1971)).

[6] One of them, in Jackendoff's case, seems to involve a wish to formulate a cyclic-interpretive theory of pronominalization. Chomsky's interest in a successive cyclic account of such rules is motivated in part by a desire to constrain extraction rules in certain ways, in particular, such that no element can be extracted from at least a certain class of clauses (finite clauses) unless it is in the initial "complementizer" position in such clauses. Cf. Chomsky (to appear c). This principle collapses if U-Rules like the *wh* movements are not successive cyclic.

examples which exemplify the Preposition Dangle argument, states the following:

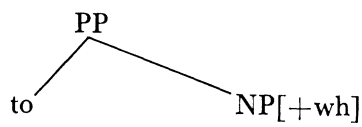
> Chomsky has suggested (class lectures, winter 1969) that optionality of preposition movement is due to the optional attachment of the feature *wh* to the PP instead of the NP. Then Wh-preposing applies to whatever node is marked *wh*, and there is no way for a step-by-step process of Wh-preposing to make a mistake like (52) or (53). Postal himself . . . suggests that such a feature . . . is necessary, so his earlier argument is nullified. Therefore it can be maintained that Wh-preposing is cyclic (*read* successive cyclic: PMP), and hence that pronominalization is also.

This argument is approvingly repeated in Bresnan (1971), although she herself prefers a conception in which the rule is not successive cyclic.

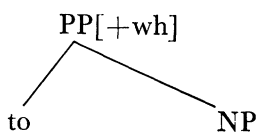
The proposal by Chomsky which underlies this attempt to maintain the successive cyclic nature of Wh Rel Movement and Wh Q Movement is essentially this. The underlying structure[7] of sentence pairs like (4b, c) would not be identical, as my formulation of the Preposition Dangle argument assumed, but would be differentiated by means of features with respect to the constituents which move. Roughly, the relevant NPs would differ as:

(5)

a.



b.



The movement rules would then be made sensitive to constituents marked [+ wh]. In this way, it can be guaranteed that no prepositions will ever be stranded incorrectly in intermediate positions.

An anonymous commentator has pointed out that one immediate objection involves a necessary effect the proposal has on Pied Piping structures in the sense of Ross (1967). In particular, in cases such as *reports the height of the lettering on the covers of which the government prescribes . . .*, since the sequence *the height of the lettering on the covers of which* is moved, it is this which Chomsky's proposal requires be marked as [+ wh]. However, the morphological consequences of this feature show up only on the micro NP *which*. Consequently, some extra rule is required to "project" the feature marking [+ wh] from the macro constituent, where it is used to trigger movement, to the included constituent whose morphological form is effected.

[7] It is irrelevant here whether these differentiating markings are taken to be present in the "base structures" or whether they are transformationally introduced.

Moreover, the whole feature-marking proposal has no independent justification. The point is not that descriptive adequacy is unachievable in this way, but rather that it is achievable under the assumption of successive cyclicity only at the cost of having available the overly powerful device of marking arbitrarily selected nodes with arbitrary[8] rule behavior coding features. It is strange that this powerful device should be appealed to by authors who are often at pains to stress the need for restricting the power of syntactic theory, and who have often objected to other approaches on just this ground.[9] Jackendoff's remark that I once proposed just such features is, of course, correct, but not relevant. My mistaken use and temporary advocacy of arbitrary features offers no support for this theoretical move. I stress now the claim that such arbitrary constituent differentiators should be eliminated from linguistic theory, providing an important and necessary restriction of the power of linguistic rules. It is, I think, one of Chomsky's more important contributions to have made us aware of the importance of such restrictions, that is, aware of the fact that a linguistic theory is incorrect if too powerful just as it is incorrect if too weak. The importance of this contribution is, of course, not at all diminished by its author's occasional deviation from its consequences any more than it is by the present writer's or anyone else's similar deviations.

I have just claimed that arbitrary node markers of the sort which are necessary even to preserve descriptive adequacy for an account in which Wh Rel Movement and Wh Q Movement are taken to be successive cyclic involve an illegitimate appeal to overly powerful devices. If this is correct, it should be possible to find empirical arguments of a different sort which show that the rules in question are not successive cyclic. That is, if the theoretical argument is correct, features like the [+ wh] required in (5)

[8] By "arbitrary" here I mean that the features in question reflect neither surface structure properties nor semantic properties. Thus they are equivalent to the use within phonology of features which are neither phonetic reflections nor reflections of independent syntactic properties (like surface constituent categories). On the other hand Chomsky has strongly stressed the need for restricting arbitrary features in phonology.

[9] Thus Chomsky (to appear a) says:

> The gravest defect of the theory of transformational grammar is its enormous latitude and descriptive power.
>
> Improvements from the worst possible case will come by placing more restrictive conditions on the choice of grammars, limiting the kinds of rules that can appear in them and the way in which these rules can operate.

Jackendoff states (1971, 28):

> One way out would be to posit a hypothetical indefinite mass pronoun, which would be required to undergo a deletion rule. Such lexical items have been proposed extensively in recent literature. However, the free use of this device constitutes a powerful extension of linguistic theory; one should try as hard as possible to eliminate the need for it in each individual case.

But with regard to the problem at hand the situation is this. A theory which bans arbitrary syntactic features is stronger than one which allows them, hence to be preferred in the absence of concrete evidence showing the need for weakening the theory, following the principles which Chomsky has long stressed. But, at least as far as U-Rules are concerned, no such evidence exists since, for example, the preposition-stranding facts are easily handled without such by specifying that the rules operate once "in a single bound" rather than successive cyclically. Thus the use of such features is an unnecessary weakening of linguistic theory. Compare below footnote 11 for a further indication of the extent of this weakening.

do not exist. Hence the rules cannot be successive cyclic because of the Preposition Dangle argument. But, if the rules are not successive cyclic, this should reveal itself in other properties of the set of well-formed sentences involving such rules. First note that Topicalization is also subject to the Preposition Dangle argument:

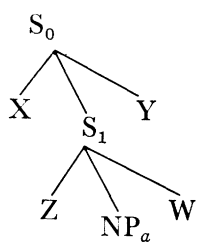
(6) a. I believe Mary thinks Joan gave a book to Melvin.
b. Melvin, I believe Mary thinks Joan gave a book to.
c. To Melvin, I believe Mary thinks Joan gave a book.
d. *Melvin, I believe to Mary thinks Joan gave a book.
e. *Melvin, I believe Mary thinks to Joan gave a book.

Hence the [+ wh] feature device does not suffice, and a new feature would have to be invented for Topicalization, as well as a new rule to assign it. Or is it to be claimed that, in contrast to the *wh* rules, the "no stranding" property of Topicalization derivations is to be attributed to Topicalization's not being successive cyclic, so that the similarity between this rule and the other two is accidental?

C. A Further Argument That U-Rules Are Not Successive Cyclic

If U-Rules are successive cyclic and obligatory, an NP which is moved by them is transported from its original position to the front of the subclause containing it as soon as the cycle defined by that clause is reached. Consider then a tree like:

(7)



A consequence of the successive cyclic hypothesis is that if a rule applies to NP_a, it does so first in structures like (7) on the S_1 cycle. Now observe that there exist cyclic rules which raise NPs out of subordinate clauses into main clauses. In cases like (7) such rules would move an NP out of S_1 into S_0. But by their very nature, these rules are not applicable in a configuration like (7) until the second (or S_0) cycle, since their environment cannot be met until that point. English has at least two such rules: Raising, which lifts complement subjects NPs, and Tough Movement, which lifts complement nonsubjects. Thus it is generally recognized that sentences like (8) are the result of the application of these two rules.

(8) a. Melvin seems to love Claudia.
b. Melvin is tough (for Betty) to please.

Thus, in a configuration like (7), the successive cyclic hypothesis about U-Rules

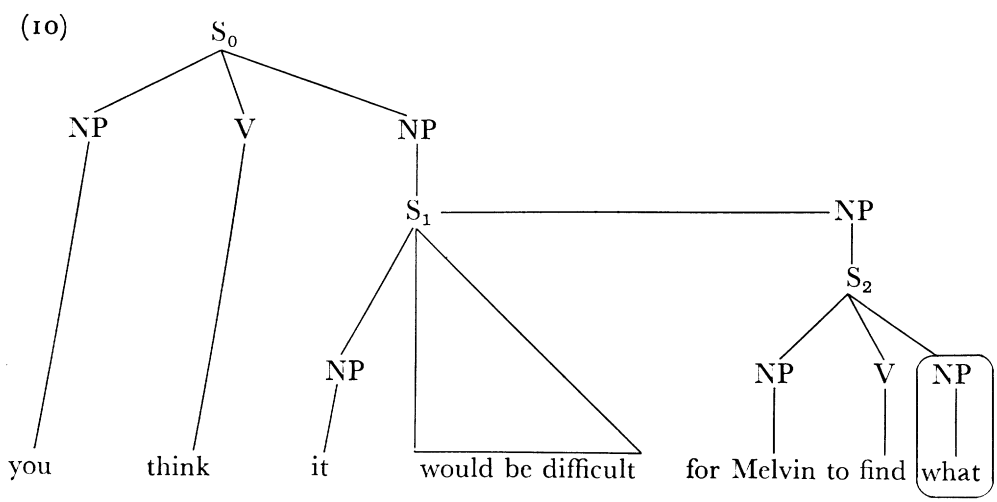
means that these will apply on the S_1 cycle before any rule like Raising or Tough Movement can be applied. Therefore, the successive cyclic hypothesis must claim either (i) that application of U-Rules to an NP in S_1 in a structure like (7) never destroys the environment for subsequent application of a rule like Tough Movement or (ii) that such application does destroy the environment for Tough Movement (that is, application of the U-Rule on S_1 to an NP blocks application of Tough Movement on the relevant S_0 cycles). But both of these claims are false. If obligatory U-Rules apply successive cyclically, they will necessarily apply to S_1 type structures in (7) in such a way as to block application of Tough Movement. But, in fact, U-Rule application in English cannot be allowed to block Tough Movement application,[10] since there exist sentences which are formed from the interaction of these rules on a single NP.

The crucial sentences are those like:

(9) What do you think would be difficult for Melvin to find?

Ignoring McCawley's (1970) demonstration that English has underlying VSO order, a relevant remote form of this would approximate:

(10)



[10] The only alternative to Tough Movement which seems possible is one in which a rule of Tough Deletion is posited such that examples like (8b) would be derived from schematic remote structures like:

(i) Melvin_x is tough (for Betty) to please x

But this hypothesis is disconfirmable by several arguments. One such involving the general complement deletion rule Equi is given in Postal and Ross (1971). Another can be constructed on the basis of the special kind of reflexivization that occurs after "picture nouns" like *picture*, *story*, or *joke*. In general, this reflexivization is subject to the constraint that within the cycle the reflexive pronoun be commanded by its antecedent at some point:

(ii) a. *Mary_i is beautiful but Bob has no pictures of herself_i.
b. *People who bug Bob_i are never given pictures of himself_i.
c. *The pilot who helped Jack_i called the girl who painted a picture of himself_i.
d. *The fact that Bob_i loved Martha made Joan draw a picture of himself_i.

Within the complement of *think* in (10) there is a model of the structure in (7). Successive cyclic application of obligatory Wh Q Movement means that the circled NP in (10) must move on the S_2 cycle. But Tough Movement cannot even be tested for application until the S_1 cycle, since it is a rule which will move a nonsubject NP in S_2 into the position of *it* in S_1. Consequently, when the grammar tests for application of Tough Movement, it will find in the S_1 position a structure with no postverbal NP. This will necessarily block Tough Movement application.[11] Hence the derivation can

---

(iii) a. Bob_i said pictures of $himself_i$ were not worth having.
b. That picture of $himself_i$ made Bob_i angry.
c. That picture of $himself_i$ proved Bob_i to be a werewolf.
d. *That picture of $himself_i$ proved that Bob_i was a werewolf.

This is, of course, only a necessary and by no means a sufficient condition:

(iv) *Bob_i said I should buy pictures of $himself_i$.

The key fact, however, is that Tough Movement constructions permit such reflexivization:

(v) a. That picture of $himself_i$ was tough for Bob_i to paint.
b. That picture of $himself_i$ was tough for me to imagine Bob_i being willing to sell.
c. That picture of $himself_i$ was tough for me to get Bob_i to agree to sell.

Only the Tough Movement analysis and not the Tough Deletion analysis provides these sentences with cycle internal structures which meet the command condition, namely:

(vi) a. It was tough for Bob_i to paint that picture of $himself_i$.
b. It was tough for me to imagine Bob_i being willing to sell that picture of $himself_i$.
c. It was tough for me to get Bob_i to agree to sell that picture of $himself_i$.

[11] Actually, it is not completely true that successive cyclic application of U-Rules will block application of Tough Movement. This conclusion can be avoided by making use once more of arbitrary syntactic features like the [+ Wh] advocated by Chomsky for making Wh Rel Movement successive cyclic in the face of the no-preposition-stranding facts. In this case, what would be needed would be some special feature, call it [+ Tough], plus a cyclic rule, either unordered with or ordered before successive cyclic rules, to assign this to NPs. This new rule would mark NPs that are in the appropriate positions for undergoing Tough Movement with the appropriate marking. Then it would be claimed that Tough Movement applied not to postverbal NPs but to NPs marked [+ Tough]. The marking rule would also have to be ordered with respect to other rules, for instance it would have to follow Passive and other rules which create objects since Tough Movement operates on derived rather than underlying objects:

(i) Max is tough for me to imagine Sally being kissed by.

Hence the total analysis would involve all of the following excrescent apparatus:

(ii) a. an arbitrary feature [+ Tough]
b. a rule to assign [+ Tough]
c. ordering statements restricting the rule in (b) with respect to Passive and other rules that create derived nonsubjects.

One needs only to look at lists like (ii) to see the incorrectness of such an arbitrary feature "solution". But in fact (ii) is not the worst problem with the proposal.

Observe that a feature like [+ Tough] simply serves to code part of the structure of early trees into late trees. I have discussed such features elsewhere (Postal, 1972), arguing that they build into linguistic theory in a surreptitious way the power of global rules in the sense of Lakoff (1970). It is odd that Chomsky (to appear a) over and over stresses his opposition to allowing a weakening of linguistic theory to include global rules. Yet simultaneously he permits the use of arbitrary features like [+ Wh], which can easily be used to build in a wide range of global power. In short, the use of arbitrary syntactic features is not a minor wrinkle in grammatical description but is equivalent to adding in an obscure way a largely unconstrained apparatus of global rules. The extent of obscurity here is revealed, I think, by the fact that one finds the same group of grammarians opposing global rules on the basis of their supposedly excessive power while simultaneously advocating and using arbitrary syntactic features which permit much of the power of global rules as well as considerable additional power of other (undesirable) sorts.

only continue with further successive cyclic applications of Wh Q Movement, yielding ultimately the sentence:

(11) What do you think it would be difficult for Melvin to find?

This is also a perfectly good sentence. However, while (11) can be derived, (9) cannot be, since (9) can be generated only by application of Wh Q Movement to the output of Tough Movement. But successive cyclic application of Wh Q Movement guarantees that Tough Movement and Wh Q Movement can never apply to the same NP, contrary to fact. This argument can, of course, be duplicated for both Wh Rel Movement and Topicalization:

(12) a. The thing that you think would be difficult for Melvin to find . . .
b. Gold, I think would be difficult for Melvin to find.

As to the latter, a most obvious excess of arbitrary features is that they permit any arbitrary collection of elements whatever to be turned into a *natural* class, that is, a class all of the members of which can be designated more economically than any single member. In this respect then, such features both justify their name and reduce the notion of natural class to vacuity. To illustrate this power, consider the class of English morphemes consisting of the preposition *of*, the noun *grizzly* the verb *dazzle*, the determiner *the*, and the conjunction *or*, hardly a natural grouping. Nonetheless, these can be made a natural class by recognizing that they all share the feature [+ Bananas], which is not assigned to any other elements. In this way, one could elegantly state a rule referring to just these five forms.

It might be objected that even allowing arbitrary syntactic features like [+Wh] or [+ Bananas] does not open the way to absurd analyses like that in the previous paragraph since the cost of assigning the features to the five items and no others would outweigh any savings in the rules referring to such collections. This is not correct, however. In fact, there need be *no cost whatever* in the relevant assignment. This can be eliminated by noting that the assignments can be removed from English grammar and made part of universal grammar. It can be done by strengthening universal grammar to include such principles as the following:

(iii) a. all prepositions of the form *of* are assigned [+ Bananas]
b. all nouns designating grizzlies are assigned [+ Bananas]
c. all verbs ending in *-zl* meaning "dazzle" are assigned [+ Bananas]
d. all definite determiners beginning with nonstrident dental fricatives and not ending with consonants are assigned [+ Bananas]
e. all disjunctive conjunctions ending with a liquid are assigned [+ Bananas]
f. no other elements are assigned [+ Bananas]

In this way, obviously any arbitrary collection can be made a natural class without cost. The point is, of course, that the excessive power of arbitrary features is in no way lessened—and in fact is rather increased—by claiming that the features are required to be universal, or even by claiming that the universal account gives necessary and/or sufficient conditions for assignment of the features. Evidently, what is wrong with arbitrary features has nothing to do with whether they are regarded as universal or not, since their arbitrariness renders this vacuous.

There thus seems no alternative to the view, long accepted in phonology, that a linguistic theory which is to provide the basis for explaining the child's ability to choose a grammar on the basis of limited and degenerate evidence cannot make available an unconstrained set of arbitrary symbols for constructing structural descriptions of sentences. In phonology, the obvious basis for constraining the class of available symbols lies chiefly in phonetics; in grammar it lies no less obviously in semantics. In both cases, however, the determination is not complete and a residue of categorizations exist which are not so accountable. It is these which are evidently understood the least. That is, how should one characterize such semantically unmotivated (and linguistically variable) categories as auxiliaries, adjectives, prepositions, conjugation classes, nonsemantic gender classes, and so forth? What has been stressed here is that whatever the solution it must be radically distinct from one that allows appeal to arbitrary syntactic features.

It is clear, then, that whereas explicitly global rules simply permit rules to refer to nonadjacent tree structures in derivations, arbitrary features permit this and in addition an unconstrained extension of the class of possible tree structures themselves.

It is obvious then that the operation of Tough Movement shows that NPs moved by U-Rules must, in cases like (9) and (12), remain in their original positions until the cycle on which Tough Movement applies. But this is just what they cannot do if obligatory U-Rules are successive cyclic. Hence, there is a clear argument independent of the Preposition Dangle argument which indicates that U-Rules are not successive cyclic.

George Lakoff has pointed out to me that the preceding argument depends heavily on the assumption that U-Rules are obligatory. He notes that if U-Rules are optional, then sentences like (9) could be derived even with successive cyclic application on the branch of derivation in which optional Wh Q Movement is not applied on the first appropriate cycle. An advocate of successive cyclic application might, Lakoff then observes, simply claim that U-Rules are optional.

However, U-Rules cannot in general be regarded as optional. Thus observe that with a variety of verbs which trigger Wh Q Movement (for example, *wonder*), ill-formed sentences result if Wh Q Movement fails to completely front a *wh*-marked NP:

(13) a. I wonder who Bob thinks Joan married.
b. *I wonder Bob thinks who Joan married.
c. *I wonder Bob thinks Joan married who.

The facts in (13) follow automatically from the formulation of Wh Q Movement as an obligatory rule. But under an optional formulation, some ad hoc apparatus must be invented to draw the distinctions in paradigms like (13). The conclusion follows that at least in such contexts Wh Q Movement must be obligatory. Therefore, the existence of sentences formed by embedding analogues to (9) under *wonder* shows that the Tough Movement argument against a successive cyclic formulation of Wh Q Movement obtains:

(14) a. I wonder what Harry thought would be tough for Melvin to find.
b. *I wonder Harry thought what would be tough for Melvin to find.
c. *I wonder Harry thought it would be tough for Melvin to find what.

An analogous argument holds for Topicalization. In this case, the point is that Topicalization moves an NP up a tree only when triggered by a certain kind of higher verb, probably one involving verbal expression. Thus many speakers, including the present writer, find both (15a) and (15b) acceptable:

(15) a. Max, Harry said Joan would never be willing to marry.
b. Harry said that Max, Joan would never be willing to marry.

(15a) would be accounted for by claiming that the deleted performative verb falls

with *say* into the category of verbs which can trigger Topicalization. Consider now facts first brought to my attention by John Ross:

(16) a. Max, Harry said Joan realized I hated.
b. Harry said that Max, Joan realized I hated.
c. *Joan realized that Max I hated.

(16c) indicates that *realize* is not one of the verbs which properly trigger Topicalization. But the existence of sentences like (16a, b) shows that an NP can be topicalized out of the complement of *realize* when it is moving still higher to some proper verbal trigger. Therefore, if Topicalization were optional, the grammar would necessarily generate examples like (16c) and (17):

(17) *Max said that Mary realized that Helen you loved.

since these would result from applying Topicalization successive cyclically on the first cycle, but not taking the option on later cycles. Hence, Topicalization cannot be an optional successive cyclic rule.

In the case of Wh Rel Movement, with respect to either restrictive or appositive clauses, the fact that application is obligatory is obvious enough to require no stressing:

(18) a. *The wombat$_i$ Melvin thinks you like which$_i$ . . .
b. *The wombat$_i$ Melvin thinks which$_i$ you like . . .
c. The wombat$_i$ which$_i$ Melvin thinks you like . . .

(19) a. *Joan$_i$, Melvin thinks you like whom$_i$ . . .
b. *Joan$_i$, Melvin thinks whom$_i$ you like . . .
c. Joan$_i$, whom$_i$ Melvin thinks you like . . .

I conclude that the claim that U-Rules are successive cyclic cannot be rescued from the argument based on Tough Movement by claiming these rules are optional, since the evidence is clear that they are not optional.

Incidentally, facts like (13) above provide an obvious further difficulty for successive cyclical conceptions of U-Rules. Such examples show that Wh Q Movement must be formulated in such a way that it does not fail to prepose a *wh*-form with a verb like *wonder*. This could be an automatic consequence of a higher trigger cyclic formulation according to which the form would move once to the right of the element that triggered it. But with successive cyclical application, even of an obligatory sort, the question arises as to how the trigger is to be made to mark the fact that the *wh*-form must move (in *n* jumps) just that far and no further. (This is the problem referred to in footnote 3.) As far as I can see, the problem is not trivial and probably should count as the basis for an independent further argument against successive cyclic application of U-Rules.

D. Conclusion

It follows from the above discussion that U-Rules in general cannot be successive cyclic. Of course this negative conclusion leaves open the question of just exactly how these rules are applied, and is in no way incompatible with the claim that this mode is not the same for all three.

References

Bresnan, J. (1971) "On Sentence Stress and Syntactic Transformations," *Language* 47.
Chomsky, N. (to appear a) "Some Empirical Issues in the Theory of Transformational Grammar," in P. S. Peters, ed., *Goals of Linguistic Theory*, Prentice-Hall, Englewood Cliffs, New Jersey.
Chomsky, N. (to appear b) *Problems of Freedom and Knowledge*, Pantheon Books, New York.
Chomsky, N. (to appear c) "Conditions on Transformations," (to appear in *Studies Presented to Morris Halle*, Holt, Rinehart, and Winston).
Jackendoff, R. (1969) *Some Rules of Semantic Interpretation for English*, unpublished Doctoral dissertation, MIT (to be published by the MIT Press, Cambridge, Mass.).
Jackendoff, R. (1971) "Gapping and Related Rules," *Linguistic Inquiry* 2.
Lakoff, G. (1970) "Global Rules," *Language* 46.
Lakoff, G. (to appear) "The Global Nature of the Nuclear Stress Rule," *Language* 47.
McCawley, J. D. (1970) "English as a VSO Language," *Language* 46.
Postal, P. M. (1971) *Cross-over Phenomena*, Holt, Rinehart, and Winston, New York.
Postal, P. M. (1972) "A Global Constraint on Pronominalization," *Linguistic Inquiry* 3.
Postal, P. M. and J. R. Ross (1971) "Tough Movement Si, Tough Deletion No!" *Linguistic Inquiry* 2.
Ross, J. R. (1967) *Constraints on Variables in Syntax*, unpublished Doctoral dissertation, MIT.

IBM
Thomas J. Watson Research Center
Box 218
Yorktown Heights, New York 10598