Review of Paul Postal On Raising: One Rule of English Grammar and its Theoretical Implications

Emmon W Bach
REVIEW ARTICLE


Reviewed by EMMON BACH, University of Massachusetts

1. INTRODUCTION. In this review, §1 outlines the major issues raised in Postal's monograph. In §2, I evaluate a number of P's arguments and conclusions. In §3, questions are posed about several controversial assumptions made by P. In §4, I raise a question of explanatory adequacy, in connection with an alternative analysis available in the revised version of the 'extended standard theory'. In §5, I compare a number of alternative theories about complement structures of English. In §6, I return to the question of frameworks, with a diatribe against grammars that lack rules. Finally, §7 gives a somewhat pessimistic assessment of what we can conclude from the whole debate so far.¹

1.1. THE CENTRAL QUESTION to which P's monograph is dedicated is whether or not there is a rule of Raising from subject to object position in the derivation of sentences like

(1) We hold these truths to be self-evident.

Chomsky 1973, P's principal opponent, holds that there is not. Both scholars agree that the sequence these truths to be self-evident is a single constituent (a SENTENCE) in remote structure. P holds that it is also dominated by NP; Chomsky does not. Chomsky maintains that the sequence is still an intact sentence in surface structure. P's rule takes as input a structure in which a sequence like these truths be/are self-evident is the sentential object of a verb like hold, and gives as output a structure in which the former subject (these truths) is now the direct object of the verb and is followed by a separate constituent, the infinitival phrase with to. The bulk of his book is devoted to arguments for the existence of this rule, and relatively little space to the actual form of the rule. Indeed, the rule itself is never explicitly stated—a fact that I consider to be of some significance, and to which I will return.

1.2. AN ALTERNATIVE ANALYSIS was proposed by Rosenbaum 1967. That analysis agrees with P's in claiming that, in the surface structure of 1, these truths is an object of hold. It also agrees with P's in attempting to give a uniform treatment both to

¹ In the fall of 1973, I was given the task of outlining the major arguments from an earlier version of P's monograph in a seminar at the University of Massachusetts, Amherst. Almost nothing survived. On the other hand, almost none of the counter-arguments survived either, and it must be said that many of those counter-arguments have been dealt with in P's published version. I would like to thank the following participants in that seminar, since much of what is in this review had its origin there: Adrian Akmajian, Joan Bresnan, Ellen Broselow, Mary Clark, Robin Cooper, Elan Dresher, Frank Heny, George Horn, Jay Keyser, John Nerbonne, Barbara Hall Partee, Tom Roeper, Muffy Siegel, Susan Tavakolian, and Tom Wasow. I would also like to thank Barbara Hall Partee for valuable comments on my review; many of the ideas included here had their origin in discussions with her. Completion of this review was made possible by the Center for Advanced Study in the Behavioral Sciences and a grant from the National Endowment for the Humanities.
sentences with Subject-to-Object Raising (S-to-O) and to sentences like 2, with Subject-to-Subject (S-to-S) Raising:

(2) These truths seem to be self-evident.

It differs from P’s in the assumption that the rule of Extraposition is involved in the derivation of sentences like 1, 2, and 3:

(3) These truths are held to be self-evident.

In his sketch of the history of analyses for such sentences, P reviews the arguments given by G. Lakoff and by J. R. Ross against the Extraposition analysis. Four arguments against the Extraposition analysis are held by P to be no longer valid; two he takes to be still valid, and sufficient to refute Rosenbaum’s hypothesis. (I will take up these arguments in §5.1.)

It is difficult to write a single rule which will accomplish the mappings from underlying subject and object complement sentences to structures of the sort assumed by P and others for sentences like 1–2, given the usual assumptions about the underlying structure of English sentences. This fact and others led McCawley 1970 to postulate the VSO HYPOTHESIS about the underlying structure of English sentences.

1.3. AN EVEN STRONGER CLAIM than McCawley’s is made by P, who suggests that it is impossible ‘within the normal constraints on rule structures’ (23, fn. 23) to write a uniform rule for S-to-O and S-to-S Raising under the SVO assumption about English. (I will try to cast some doubt on this claim in §5.) Apparently, Chomsky and P agree on the logic of the following argument:

(4) a. It is desirable to provide a uniform rule for S-to-S and S-to-O Raising.

b. It is impossible (or difficult) to provide a uniform rule under the assumption that the underlying order of English sentences is SVO.

What follows (weakly) from these premises is that the underlying order of English is something other than SVO. But in fact, Chomsky rejects 4a, claiming that there is no such thing as S-to-O Raising. Of course, as P notes, it does not follow that the underlying order of English is SVO. In 4b itself there are many hidden assumptions about the form of rules, the nature of underlying structures, and so on.

1.4. SHifty FRAMEWORKS make it difficult to reach firm conclusions. The debate about Raising has been carried on within no fixed framework. One can discern, even in P’s own argumentation, a number of different positions.

Most of the book is written from a standpoint of what we may call ‘generative semantics’, vintage 1970–72 (xiv–xv). Thus, (1) ‘underlying structures are subparts of the logical structures of sentences ...’; (2) ‘the only symbols usable ... are the set of symbols necessary for the formation of surface structures and the set of symbols required to describe logical structures’ (P rejects ‘all so-called syntactic features, doom markers, other abstract syntactic markers, coding devices, “empty nodes”, “doubly filled nodes”’—and presumably also trace elements of the sort hinted at in Chomsky 1973, and more fully developed in later work by Chomsky and that of some of his students, e.g. in Fiengo 1974.) In addition, P assumes ‘the
availability of "global" statements in grammars, rules that range over more than one (not necessarily contiguous) tree in derivations' and 'even more peripherally ... "transderivational" rules, principles defined over sets of derivations.'

Shimmering through in various footnotes and pages (especially in P's §8.4) is a later position which postdates most of the writing of the book, that of 'relational grammar' (P refers to several works to appear by himself and D. Perlmutter.) This framework is somewhat outside standard transformational theory. The basic component of a relational grammar is a set of specifications of sentences employing such notions as Subject, Object, and Indirect-Object as primitives. Relational rules are defined on such structures, which are assumed to have no left-to-right order. Within such a framework, rules like Equi, Raising etc. (the clearly cyclic rules of transformational grammar) are replaced by rules that change grammatical relations. Thus the verb-initial hypothesis becomes meaningless, along with an alternative verb-final thesis advanced by Ross 1973.

The third main framework discussed in the book is that of Postal's opponents, alluded to in the quotations above—the position variously called 'extended standard theory', 'interpretive semantics', and the like. In the final sections of this review, I shall return to the question of the debate about frameworks. Along the way, I will try to point out places where indeterminacy of frameworks makes it impossible or difficult to evaluate arguments.

1.5. Four basic classes of verbs in English are dealt with by P: 'B-verbs', the classical object-raising verbs like believe, hold, allege, and rumor (verbs of propositional attitude); 'W-verbs' like want, wish, and prefer, which differ from the first class in accepting complements with for; 'A-verbs' like seem and begin, with which the subject of the complement becomes the surface subject of the (usually) aspectual or modal verb (P includes the modal auxiliaries here); and 'N-verbs' like prevent, which have negative import and take a complement verbal with from ... -ing. There is also a short section on verbs, like regard, which take complements with as; P believes that these participate in Raising also. His analysis assumes a uniform Raising rule which entails the application of a number of other rules or principles: insertion of to and for (and from?); have-substitution for an underlying past tense; to-be deletion for verbs like consider construed with adjectives; and the operation of various global and surface constraints, such as the principles determining quantifier scope, different entailments or presuppositions for raised vs. non-raised structures, and a 'derived object constraint' to account for the fact that certain verbs lead to unacceptable collocations unless the derived object is removed by some rule such as Passive (rumor).

After his brief historical sketch, P outlines the main alternative he wishes to refute: Chomsky's proposal (1973) that no Raising has occurred at all in sentences like 1. Chomsky's explanation depends on the correctness of conditions like the 'tensed sentence' and 'specified subject' conditions; so P, throughout his book, tries to undermine alternative explanations for relevant facts that depend on those conditions. (See Bach & Horn 1976 for some remarks on Chomsky's conditions.)

The bulk of the book is taken up with three kinds of arguments for the existence of a rule of Raising: (a) 'previously discussed' arguments, depending on facts of the Passive, reflexives, reciprocals, and the inclusion constraint, which is supposed
to 'explain' the oddity of sentences like We like me; (b) new arguments which P considers relatively firm, in the sense that enough is known or seems plausible about the structures, rules, or principles involved to justify the conclusion that the facts speak for rather than against a Raising analysis; and (c) potential arguments, which are weakened by the absence of any very firm ideas about the premises. Other chapters deal with arguments against Raising, which P attempts to refute; remarks on why, if Raising exists, it works the way it does; Raising in nominalizations; meaning and Raising; plus various odds, ends, and afterthoughts ('a late argument', 'the next-to-last argument', 'the last argument'). Alas, if we could only count arguments to reach conclusions, P would surely have won this monumental battle in which almost 450 pages in favor of Raising are pitted against some scattered counter-arguments, taking up perhaps a page in all.

One finishes the book with a real appreciation of the complexities of language and the difficulties of reaching a firm conclusion about anything. Given that the book is such a rich mine of facts, alleged facts, semi-ideas, ideas, and refutations, it is probably good that P seems to have left everything in that he, or almost anyone else, ever thought about certain kinds of complementation. One section ends: 'These facts have no bearing on the issues then' (184).

2. SOME ARGUMENTS. Without writing a book at least as long as P's, I cannot deal with every bit of argumentation. I shall pick out a few of what seem to me the more convincing arguments for each of several claims made in the book.

2.1. THE NP-TO-PRED IS NOT A CONSTITUENT. The rule of Right Node Raising applies in clear cases to single constituents including complement sentences:

(5) I know but no one else knows—that the world is flat.
(6) I don't know WHEN but I do know THAT—we will go to Boston.

But it cannot apply to the NP-to-Pred part of a sentence like 1:

(7) *We hold but no other nation holds—these truths to be self-evident.

This follows from Postal's analysis, but not from Chomsky's.

2.2. THE NP IS NOT A SUBJECT. Complex (heavy) NP shift does not apply to subjects of complement sentences (in clear cases). It does apply to the NP of sentences like 1:

(8) We hold to be self-evident those truths for which our forefathers fought.

2.3. THE NP IS AN OBJECT. Adverbs cannot come between verbs and object NP's, but can come between verbs and complement sentences:

(9) I believe with all my heart that you are innocent of any evil intent.
(10) *I love deeply women.

In this respect, Raising sentences behave like 10, not 9:

(11) *I believe with all my heart you to be innocent. 3

2.4. THE SURFACE STRUCTURE of 1 is then like that of sentences with persuade-type verbs, as the above and many other arguments show. But could not this structure be directly generated in the base by an expansion of VP, as follows?

(12) VP —> V NP to Pred

3 Bresnan 1976b has attempted to undermine most of P's main arguments. In a number of cases, she has shown that the Raising hypothesis is not necessary to account for the facts; but her counter-arguments do not show that there is no such rule, and do not affect the evidential nature of P's arguments insofar as they are taken to be statements about what follows from the hypothesis and not from the alternatives he discusses.
P gives a number of arguments purporting to show that the NP cannot be independently generated (tabs, there, 'weather' it etc.) But given the looseness of the various frameworks available, it is not at all clear that these arguments are decisive. Suppose that one pointed to the presence of a passive 'downstairs', as in 13, to show that there must be a sentence in deep structure:

(13) I believe my cat to have been swallowed by a guppy.

For Chomsky, whose transformations are blind to grammatical relations, and are simply structure-dependent in a relatively 'stupid' way, Passive could very well apply within the VP to get 13 from a structure like Figure 1. (The structural condition is modified slightly—by the inclusion of AUX—from the one given by Chomsky 1973:236.)

![Figure 1.](image)

Similarly, if lexical insertion has no constraints in terms of sentence boundaries etc., one could perfectly well get the pieces of idioms, weather it etc. into such structures. As far as I can see, there simply is no argument against this analysis within the revised standard theory. Note that what might be an argument within some other theory, namely the semantic difference between sentence pairs like those in 14–15, is inapplicable under Chomsky's recent assumptions that the semantic interpretation is read directly from a 'surface structure' (now redefined to have an encoding by means of various trace elements of parts of the derivational history of a sentence):

(14) a. I persuaded Mary to groom the kinkajou.
   b. I persuaded the kinkajou to be groomed by Mary.
(15) a. I expected my daughter to mow the lawn.
   b. I expected the lawn to be mowed by my daughter.

So one flaw in P's reasoning is apparent. His arguments show that Chomsky's surface structure is wrong; but they do not show that this must result from the operation of a rule, rather than being directly generated in the base. However, I do not think it a virtue of the revised standard theory that it has no argument against the no-rule analysis; more on that later.

2.5. Passive facts turn out to be inconclusive. Both Chomsky and P assume that 16 is the source of 17, via Passive:

(16) Everyone holds these truths to be self-evident.
(17) These truths are held by everyone to be self-evident.

The difference is, of course, that P assumes Raising in 16. To account for the failure of Passive to apply in sentences like 18, Chomsky invokes a constraint that says you can't do things to tensed sentences:

(18) Everyone thinks (that) these truths are self-evident.

But since S-to-S Raising also exists (for both P and Chomsky), there is another possible derivation for 17, by applying Passive first and then Raising, i.e. through 19:

(19) These truths to be self-evident is held by everyone.
Of course 19 is ungrammatical, but never mind. On P's account, the sentence would surface as 20 if Raising didn't apply (for Chomsky, presumably S-to-S Raising could obligatorily apply, to yield 17):

(20) It is held by everyone that these truths are self-evident.

Curiously, neither Chomsky nor P discusses such possibilities. Moreover, P must allow this derivation—since he holds that there is no extrinsic ordering—or else he must specifically prohibit Raising to apply from passivized subject position (more on ordering below).

Now here is a curious situation. It's not even clear that there is a question, but quite strong theoretical claims are made about the answer, as follows:

(21) a. Chomsky: The theory of transformations would be weakened if we were allowed to prevent a rule like Passive from dipping into a clause.

b. Postal: The theory of transformations would be weakened if rules could be allowed sometimes to operate on clause mates, sometimes into clauses (aside from unbounded movement rules).

But it appears that this very interesting theoretical question can arise only if there is a basic agreement about the structures involved in these sentences. P assumes (I gather, from some of his diagrams) that the relevant structure of 16 is roughly like Figure 2.

Now, given this structure, Chomsky must allow that Raising exists, or else his own A-over-A principle would disallow the application of Passive to a structure like Fig. 2, quite apart from any questions about tensed or tenseless sentences. On the other hand, if Chomsky follows Bresnan (see §5.2) in assuming a structure like Figure 3 to underlie 16, and a structure like Figure 4 to underlie 18, there is again nothing to explain by any special conditions, since A-over-A would again block application of Passive to structures like Figure 4.
But there is a problem here for the proponents of a theory like Chomsky's, who adhere to the hypotheses expressed in Figures 3-4. Bresnan 1976a has shown that the A-over-A principle applies not to isolated categories like NP, VP, S, but rather to categories defined as clusters of features. She also advances and supports the hypothesis that transformations can be constrained so as to contain only conjunctions of structural predicates, with no disjunctions (i.e. curly brackets) or negations. These hypotheses, taken together, make very strong claims about possible structural conditions for transformations. As far as I can tell, no combination of feature specifications will single out just S and NP as a natural class. Hence, if Bresnan is right, no single Passive transformation can account for sentences like 17 (or ordinary passives like 22) and also apply to S to produce perfectly good passives like 23:

(22) Harry was hit by a flying saucer.

(23) That these truths are self-evident is believed by everyone.

This problem is solved if S in such structures is dominated by NP. But if it is, then (as we have seen) there is an independent principle that keeps Passive from applying to the subject of the complement sentence.

I noted above that it's difficult to draw conclusions about analyses within varying frameworks. The considerations just given show that you can't really raise questions about frameworks without structural analyses. (I'll come back to these alternative structures in §5.)

So far as I am aware, the only argument against the derivation Passive-Raising is one given a long time ago (in La Jolla) by Ross and Lakoff, based on P's Crossover Principle. (The argument was actually used against an epicyclic theory of rule application, but it works here too.) Examples 24 and 25 seem to be equi-bad:

(24) ?Jane was kissed by herself. (= Jane kissed herself.)

(25) ?Jane was believed by herself to have kissed her neighbor.

This is accounted for by something like P's Crossover Principle AND a derivation from structures like 16 to ones like 17, with or without Raising. But if Passive applies before Raising, there is no way that Crossover can block 25. Now it's true (as P says in another connection) that, 'given the possibility of global rules', there are a number of ways to state a restriction that would give the results without Crossover and the derivation Raising-Passive; but as someone said in another connection, 'With global rules you can build the Panama Canal.' Until we get a better, more general account of Crossover facts, this argument stands. Note that it depends on extrinsic ordering. Many linguists, of course, would take this as sufficient reason to reject it (see below, §3.3.)

3. SOME CONTROVERSIAL ASSUMPTIONS, made by P, are not really supported. I will take up a number of them here.
3.1. Complementizers were introduced transformationally by Rosenbaum 1967, but Rosenbaum himself points out that this was a largely arbitrary choice (p. 25). It is not entirely clear how P intends to treat all the complementizers of English, but he is explicit (p. 6, text and fn. 8) in claiming that that-clauses and infinitivals have the same underlying form, and the difference between them is 'induced during derivations, rather than being a function of distinctions drawn in the maximally underlying structures'. It is curious that P does not deem it necessary to defend this assumption, since he himself cites Bresnan 1970, where the view that complementizers are present in underlying structures is explicitly argued for.

Apparently, P believes that he must make this choice because of his theory: 'under reasonable assumptions, in particular under the assumption that sentences with the same meaning have the same underlying structures, the latter conclusion [i.e. that that-clauses and sentences with Raising have the same source] is supported by the lack of semantic contrast between such pairs' (6-7). Bresnan devotes an entire chapter of her dissertation to a discussion of the semantic contrasts between these types, and P himself devotes Chapter 11 to a discussion of differences in meaning between raised and non-raised sentence pairs such as these:

(26) I found Julius Caesar to be boring.
(27) I found that Julius Caesar was boring.

As far as I can tell, none of P's arguments requires such sentences to have the same remote structure, and none of Chomsky's arguments against Raising requires them to have different underlying structures. (I'll return to the question of complementizers in §5.2.)

3.2. 'Logical structure' is invoked at various points as if it were given in advance (cf. the quotation given above in §1.4). For example, in discussing an alternative analysis of seem-sentences, under which they would be derived from the same type of structures as sentences with eager etc., P writes (34): 'It does not seem that (7) [an Equi-type structure] gives a semantically relevant analysis, since there is no reason to assume that the logical structure of (6)a [28 below] involves two occurrences of elements designating an individual named Melvin':

(28) Melvin seems to speak fluent Gwambamambese.

Earlier, he has said: 'thus the logical structure of (6)a must reflect the fact that it consists of a part [SEEEM] and another part [MELVIN SPEAK FLUENT GWAMBAMAMBSE], which are related in a certain way' (33).

It is hard to know what to do with statements like this. There is no hint of what logical structure is supposed to be. (Surely it must be more than writing capital letters in square brackets.) But even if we knew what it was, is there any a-priori reason to believe that the underlying structures of the syntax have any particular relation to it—rather than to, say, phonological structure?

It would seem to be elementary that we can't make any good arguments about the relationship between syntactic and logical form without providing an explicit theory for both. But it isn't even clear that a representation of logical form is necessary (Cooper 1975 has shown how to get directly from syntax to semantic interpretation without the intermediate step of a disambiguated language.)

3.3. Extrinsic ordering of transformational rules is rejected by P, as noted above. It has become quite popular to take this position (cf. Koutsoudas, Ringen, T. Lehmann, Sanders, Perlmutter, G. Lakoff etc.) Actually, it is somewhat misleading to discuss this issue in terms of ordering versus non-ordering. The proponents of grammars without ordering are not so much rejecting extrinsic ordering as such, but rather attempting to find language-independent principles from which particular orderings would follow—and no one can quarrel with that aim. This is not the place to enter into an extended discussion of these issues, but I will mention one puzzle that arises from the no-ordering hypothesis and P's analysis. Consider predicates which can occur with either extraposed or raised structures:

(29) It seems that the earth is round.
(30) The earth seems to be round.
(31) *That the earth is round seems.

As these examples show, some verbs require either Raising or Extrapolposition. Recall that, for P, 29–30 have the same remote structure. In a grammar with extrinsic ordering, we can order
optional Extraposition before Raising, and mark *seem* as requiring Raising (or do it the other way around). In a grammar without ordering, we must state a global condition. It is interesting that this problem does not arise in Rosenbaum's analysis (where something like 29 is the immediate source of 30, i.e. the rules are intrinsically ordered). We can simply say that Extraposition is obligatory for *seem*. (In a grammar where 29–30 have different sources, but both have sentential subjects, both rules are obligatory for *seem*.)

4. **SO WHAT IS THE RIGHT STRUCTURE** for sentences like *We hold these truths to be self-evident*? As I indicated in §2.4, on balance it seems to me that P has shown Chomsky's surface structure to be incorrect. There are arguments that show that NP-TO-PRED does not act like a constituent under various transformations, nor like a sentence under various tests (pronominalization); and intonational facts bear this out. As also indicated in §2.4, I don't think P has shown that there has to be a rule of Raising to object position (at least for Chomsky), since the appropriate structures could be directly generated in the base. That is, the underlying AND the surface structure could be roughly as in Figure 5.

![Figure 5](image)

Under this hypothesis, there might be no difference in syntactic behavior between *persuade*-type and *believe*-type verbs. Let’s call this the No-Rule Hypothesis.

There may be psycholinguistic evidence against the hypothesis that *believe* and *persuade* structures are the same. Bever, Lackner & Kirk 1969 have shown that click-locations are judged differently for the two types (cf. Fodor, Bever & Garrett 1974:339 ff.) The displacement of clicks to a position just past the main verb is greater in sentences like 32 than in those like 33:

(32) John expected Bill to leave.
(33) John persuaded Bill to leave.

Bever et al. assume that 32–33 have essentially the same surface structure, and interpret their results as an indication that deep-structure clausal units play a role in sentence perception. If P is correct about the surface structure of such sentences, then this experiment shows that there must be some difference between 32 and 33.
One such difference might be a difference in underlying structure and the operation of a Raising rule. (For this to be a very firm result, we would need to set up and test some hypotheses about the differences between purely semantic contrasts and syntactic ones, since the former might conceivably be the determinants of the differences in perception.)

The assumption that persuade and believe differ in underlying privileges of occurrence leads to many correct predictions, many of them incorporated into P's book. For example, it predicts that NP's that are limited in distribution will occur in the object slot for the Raising structure, but not the persuade structure (there, tabs, 'weather' it). The point here is not that one can't account for these facts under Chomsky's analysis, or under the alternative No-Rule Analysis mentioned above. It is rather that there is no prediction at all about this mass of facts.

Suppose it has been established that there is a rule of Raising which takes underlying sentential complements and promotes their subjects to object position in the matrix sentence. Is there any general hypothesis about linguistic form or meaning which would REQUIRE US to adopt this analysis?

Chomsky 1973 himself claims or suggests that there is a general principle which would PREVENT US from adopting this analysis, namely a requirement that no transformations be formulable which change the structural configurations of a sentence without affecting the linear order of the constituents. If P is right, this principle is false.

There are several things to be said about Chomsky's principle. First, there are fairly well-motivated rules which violate this stricture, notably the one proposed by Emonds 1970, and further motivated by Akmajian & Wasow 1975, which detaches be or have from the VP constituent and puts it into the AUX node. Second, given the possibility of using empty nodes, it is possible to escape the effect of this principle entirely. Thus, for believe-sentences, one can set up underlying structures like Figure 6 (indeed, Bowers ms has done so).

Postal’s Raising can now be effected by substituting the lower NP for Δ in the matrix.

I argued above that the No-Rule Analysis, which I am claiming to be wrong, is consistent with the revised extended standard theory. The only problem with the analysis that I can see is in accounting for the difference between sentences like these:
I believed there to have been a riot in the kitchen.

*I persuaded there to be a riot in the kitchen.

Since in this theory semantic interpretation is a function of (enriched) ‘surface structure’ (Chomsky 1975, Ch. 3), one can simply say that 35 is well-formed syntactically, but the semantics of persuade and believe account for the difference. (I’ll say more about this in §5.8.)

What would be needed to rule out this wrong analysis? In Chomsky’s language, how can we reach the level of explanatory adequacy here? I believe it is in the spirit of P’s general approach to expect an answer to this question (he asks similar questions about various properties of Raising in §8.4.) I also believe it is in P’s spirit to expect the answer to lie in a theory of logical form and its relationship to syntactic form. As I indicated above, I don’t think there is anything in P’s book that will help us here. The closest is the requirement ‘that underlying structures be subparts of the logical structures of sentences’, but this is much too vague.

Let me state the question in a slightly different way. What principle(s) might lead us to expect that the most descriptively adequate grammar for English is one in which persuade is associated in underlying structures with an NP + PRED (or NP + S) complement, but believe with just an S complement (ignoring the question of whether there is an NP above the S), rather than the opposite? What predictions can be made about the underlying structures associated with similar verbs in arbitrary languages? What does ‘similar verbs’ mean? Why aren’t there grammars in which persuade and believe have underlying structures like their surface structures, and surface structures like their underlying structures? Surely the answers to such questions must be based on the meaning of the verbs—or, as P terms it, their logical structure.

A hint at how to answer these questions is given by P (Ch. 7, fn. 11) when he argues against a derivation rather like the No-Rule Analysis, namely one in which structures are derived (via Equi) from underlying structures of the form NP V NP S: ‘The structure [... VERB + NP + S] for sentences like (i) [i.e., Bob showed Mary to be incompetent] is LOGICALLY inexplicable, since it seems that show (and other B-verbs) represent transitive predicates that relate an individual and a proposition in both sentences like (i) and (ii) [i.e., Bob showed that ...].’ Presumably the logical structure of persuade-type verbs would include some representation of the fact that persuade is a three-place relation holding between two entities and a proposition (or predicate in the VP analysis). But what theory of logical structure and its relation to syntactic structure is required if the syntactic structure of the two types of sentences is necessarily as P wishes? And what evidence is there that this desired theory is correct? We can’t simply assume that such a theory would be correct, as P appears to do.

It might be thought that a theory which incorporates a principle of Fregean compositionality would lead to a correct choice in this matter, and one such theory is that of Montague grammar and its derivatives. But even there we find opinions divided. Partee 1976 has provided a means for incorporating a rule of Raising into a Montague framework; but Thomason 1976 has worked out a Montague syntax in which persuade and believe are given exactly the same syntactic analysis. The semantic differences usually used to motivate a different syntax are to be accounted for by
meaning postulates (cf. below §5.8). Thus it would seem that, if a theory of syntax-semantics is to have the sort of predictive power that we want, some constraints must be placed on possible meaning postulates.

5. ANALYSES OF COMPLEMENT STRUCTURES. Here I will review a number of alternative analyses of sentences involving various classes of verbs discussed by P, including some not taken up in his book.

5.1. ROSENBAUM'S ANALYSIS, as indicated above, assumes (1967:7) that Extraposition applies in the derivation of both S-to-S and S-to-O sentences:

(36) IT for these truths to be self-evident seems
   IT seems for these truths to be self-evident
   these truths seem for to be self-evident
   these truths seem to be self-evident

(37) we hold IT for these truths to be self-evident
   we hold these truths for to be self-evident
   we hold these truths to be self-evident

(38) everyone hold IT for these truths to be self-evident
   IT for these truths to be self-evident be held by everyone
   IT be held by everyone for these truths to be self-evident
   these truths be held by everyone for to be self-evident
   these truths are held by everyone to be self-evident

Boiled down to essentials, Rosenbaum's rule looks like this:

(39) X, IT, Y [FOR, NP, to Z], W
     1 2 3 4 5 6 7 =>
     1 5 3 4 0 6 7

Note that this rule meets Chomsky's objection to vacuous movement (since the burden of vacuous movement is now thrown onto Extraposition). It also accomplishes S-to-S and S-to-O in a single rule.

In the opening sections of P's book, as mentioned above, he reviews a number of arguments of Lakoff and Ross against Rosenbaum's analysis (10-22). Four he rejects as invalid or inconclusive.

The first argument is that Extraposition is a last-cyclic rule, and hence Raising can not apply to its output—since, if any rule is cyclic, Raising must be. P notes that a number of linguists have shown the first premise to be invalid (cf., e.g., Baltin 1975).

4 Thomason 1976 gives only a syntax with hints at the semantics of the fragment. Thomason Msa provides a semantics, but ends by noting that his fragment does not give equivalent readings for the following:

(a) A unicorn is believed to be expected to go.
(b) It is believed to be expected that a unicorn will go.

This defect is remedied in Thomason Msb for verbs like seem, and Thomason also attempts there to treat there, the stumbling block for most theories which give persuade and believe the same syntax.

5 Chomsky's principled exclusion of such rules, or the operation of general rules in a 'vacuous' fashion, is completely undercut by his most recent work, where 'readjustment rules' of exactly this sort are admitted. See the discussion of sentences like Who did you see a picture of? in Chomsky ms, where such a rule is assumed to break up the NP a picture of Wh.
The second is the argument against vacuous movement rules. The weakness of this argument has been mentioned above (§4 and fn. 5).

The third argument requires the assumption that a sentence like 40 has a source like 41—for P (and for me) a questionable assumption:

(40) It suggests itself to me that Harry is a liar.
(41) That Harry is a liar suggests that Harry is a liar to me.

The fourth argument is based on crossover phenomena of the sort discussed above (§2.5) in connection with the interaction of Passive and Raising. Lakoff and Ross argued, according to P, that the Rosenbaum analysis made it impossible to use the Crossover Principle to account for the unacceptability of sentences like 25 (*Jane was believed by herself to have kissed her neighbor), since Passive does not move an NP over a coreferential NP. Rosenbaum's derivation went like this:

(42) a. BASIS: Jane believe IT FOR Jane to have kissed her neighbor
    b. PASSIVE: IT FOR Jane to have kissed her neighbor be believed by Jane
    c. EXTRAPOSITION: IT be believed by Jane FOR Jane to have kissed her neighbor
    d. PRO-REPLACEMENT: Jane be believed by Jane FOR to have kissed her neighbor

For-Deletion and Reflexive then suffice to derive 25.

But this argument is invalid. It is perfectly true that Passive cannot be blocked by Crossover, and should not, since for P the direct reflex of step (c) is this impeccable sentence:

(43) It was believed by Jane that she had kissed her neighbor.

But Crossover would block Pro-replacement itself in derivation 42. Hence these facts are consistent with Rosenbaum's analysis and something like a Crossover Principle. Since they are also consistent with P's analysis, as pointed out in §2.5, they do not distinguish between the two possibilities.6

Two arguments against Rosenbaum are considered valid by P, and one of these is based on the assignment of derived constituent structures—though such arguments are notoriously weak. P assumes without discussion that a particular derived structure is correct for certain sentences, that there is no non-ad-hoc way to get this presumed structure from Rosenbaum's analysis. For this argument to work, we need to know the exact statement of the relevant rules, the exact principles for derived constituent structure, and the independent justification for the supposed readjustment rules. Further, for this argument to have any bearing on the choice between Rosenbaum's analysis and P's alternative, we also need all this information for the alternative analysis. None of these prerequisites is met.7

P's remaining argument against the Rosenbaum analysis turns on the grammar of examples like the following:

(44) There kept on being riots in the Sudan.
(45) *It kept on there being riots in the Sudan.
(46) Arthur began running.
(47) *It began Arthur's running.
(48) Arthur began to run.
(49) *It began for Arthur to run.

In general, gerundive (or ACC-ing) structures do not extrapose. (P shows that superficially similar structures like It's fun swimming have quite different properties.) The argument rests on the assumption that examples like 44 and 46 should be derived in the same way as 48, or all the previous raised structures we have been considering. But 45 and 47 show that such derivations would lead to unacceptable results unless Pro-replacement were made obligatory for such verbs. But whether or not there is an intermediate step of Extraposition, such verbs as begin must be marked as requiring Raising, so I can't see any basis for choice here.

---

6 P apparently accepts the logic of the purported Lakoff-Ross argument, but concludes that a Crossover explanation can no longer be maintained (cf. §2.5 above).
7 This is one instance of the consequences of discussing grammars without rules and without theories (see §§6.2, 6.3 below).
The logic of the argument is, then, something like this:

(50) a. Sentences with begin, keep on, believe, seem should all be analysed as undergoing the same transformations.
   b. Rosenbaum’s analysis of the sentences in question involves Extraposition and Pro-replacement.
   c. The results of applying just Extraposition to sentences with begin, keep on etc. are unacceptable.
   d. Ergo, Rosenbaum’s analysis is wrong.

Something is wrong with this reasoning. By the same paradigm, we could argue against P’s own analysis of B-verbs like believe. There are verbs which lead to unacceptable results if they undergo just Raising, without subsequent application of Passive:

(51) *Everyone said John to have stolen the pie.
(52) John was said by everyone to have stolen the pie.

Perhaps we can supply whatever is missing from this argument to render the conclusion valid, and to allow P to maintain his analysis of say as parallel to believe. But premise 50a is still suspect. I shall now try to undermine that premise, by showing that the sentences with verbs like begin are quite unlike P’s other instances with Raising verbs. First, verbs like believe or seem can occur either with have + EN or with be + ing in their complements. Not so keep on or begin:

(53) John was believed to have eaten the caviar.
(54) John was believed to be secretly visiting the monk.
(55) *John kept on having eaten the caviar.
(56) *John began being secretly visiting the monk.

Second, verbs like begin appear to form a much more cohesive link with their complements in -ing. Sentences in which these verbs are separated from the complement are not acceptable:

(57) ?Mary began last year running a mile every morning.
(58) ?There kept on last week being riots in the Sudan.
(59) There seem nowadays to be many riots in the Sudan.
(60) Eleven-speed bikes are believed by many to cause congestion.

Third, true Raising verbs of the B-type select stative complements. Verbs like begin exclude them:

(61) John is believed to own a Cadillac.
(62) *John began owning a Cadillac.

But this may depend on the verb (try end up in 62).

Most important, the uniform treatment of these aspectual verbs presupposes that gerundive structures are essentially the same as the structures associated with verbs like believe and seem. But several recent studies have cast considerable doubt on this assumption. Schachter 1976 and Horn 1975 argue on independent grounds that the gerundives do not arise from a sentential source at all. From this it would follow that they could not undergo Extraposition; and Rosenbaum’s rule, either in its original form or as given above, would not apply to them. Hence, they would have to be handled by a different rule. Since P provides no rule, it is not possible to say what effect the re-analysis of gerundives would have on his conclusions.

5.2. COMPLEMENTIZERS AGAIN. As noted above, P assumes without argument that corresponding sentences with that and raised structures have the same source. That is, although P purports to follow a theory in which identical meaning requires identical remote structures (and the converse), and is himself quite cognizant of the meaning differences between raised and non-raised structures (cf. Chapter 11), nevertheless he maintains that 26–27 (repeated here) must have the same source (359):

(63) I found Julius Caesar to be boring.
(64) I found that Julius Caesar was boring.
I don’t know why P has chosen to ignore Bresnan’s analyses of the various sorts of raised structures discussed in his book (Bresnan 1970, 1972—the latter not cited). Perhaps it is because these analyses are carried out within a framework rejected by him (cf. the quotations in §1.4 above). But parts of Bresnan’s analysis are quite independent of that framework, and are relevant to P’s own analysis. One such aspect is the question of complementizers.

Bresnan argues for a deep-structure source for complementizers, partly on the basis of semantic contrasts, partly on purely syntactic grounds. Within the bounds of the types of constructions considered by P, she distinguishes several types and argues specifically against Rosenbaum. The three classes which Bresnan distinguishes within P’s A-type verbs are shown in the deep-structure configurations of Figures 7–9.

\[
\begin{array}{c}
\text{S} \\
\text{NP} \\
\Delta \quad \text{V} \\
\quad \quad \quad \quad \text{S} \\
\end{array}
\]

\textbf{FIGURE 7.} happen, be certain, begin.

\[
\begin{array}{c}
\text{S} \\
\text{NP} \\
\Delta \quad \text{V} \\
\quad \quad \quad \quad \text{S} \\
\quad \quad \quad \quad \quad \text{COMP} \\
\quad \quad \quad \text{that} \\
\quad \quad \quad \quad \quad \text{S} \\
\end{array}
\]

\textbf{FIGURE 8.} happen, be certain.

\[
\begin{array}{c}
\text{S} \\
\text{NP} \\
(\text{it}) \quad \bar{\text{S}} \\
\quad \quad \quad \quad \text{V} \\
\quad \quad \quad \quad \quad \text{COMP} \\
\quad \quad \quad \quad \quad \text{that} \\
\quad \quad \quad \quad \quad \text{S} \\
\end{array}
\]

\textbf{FIGURE 9.} matter, be certain.
Sentences like 65 then arise by an obligatory rule that shifts the subject of the complement S (which has no complementizer for Bresnan) into the position of the 'empty' subject:

(65) John happens to be sick.

This analysis correctly predicts the following array of facts:

(66) There happened to be a riot in the kitchen.
(67) *That there was a riot in the kitchen happened.
(68) *It happened frequently for John to be sick.
(69) *What happened was for John to be sick.
(70) *It began that John was sick.

Presumably, P would reject these structures because of the (for him) unmotivated 'empty' subjects. But what could be retained, even in his framework, is the distinction between 'bare' S's and sentences with complementizers like for or that. Bresnan devotes a good deal of effort to showing that these elements are not logically empty markers; and although she herself does not provide a semantic analysis, it seems quite plausible that an explicit semantics would ultimately justify the presence of these items in 'logical structure'.

The same distinction turns up in Bresnan's analysis of S-to-O Raising. P's B-verbs like believe (which never show up with for) have bare S complements, while his W-verbs (want) have for + S complements. A good deal of motivation is given for this analysis by Bresnan.

Bresnan gives convincing arguments against the assumption that for occurs in the complements of B-verbs. She distinguishes three types of superficially similar sentences, with underlying structures as in Figures 10–12.

Note that expect permits all three structures; hence, in Bresnan's words, 'if we had taken a verb like expect as paradigmatic, nothing would be revealed as systematic, since expect has properties of want, believe, and challenge' (1972:162; she leaves more or less unresolved the question of whether a Raising rule applies to structures like Figure 8.) A variant of Bresnan's first type of structure is provided by verbs like hope with prepositional complements. The for of

---

**Figure 10.** want, expect.

**Figure 11.** believe, expect.
I hope for you to be here is the complementizer, since the prepositional for is removed by an otherwise needed deletion rule. But if Bresnan is right, then it appears that Rosenbaum's analysis cannot be maintained, since Extrapo- sition applies only to S's, not S's (alternatively, to COMP + S structures).\(^8\)

If we accept the essentials of Bresnan's analysis, a number of facts follow:

(71) Complements of believe will never appear with for; e.g.,
    a. *I believe very much for John to have done that.
    b. *What I believe is for John to have done that. [cf. want]

(72) If Pseudo-cleft and Passive are defined for NP or S, and not S, we will get these contrasts:
    a. *What I believe is John to have done that.
    b. *John to have done that was believed by everyone.
    c. What I expect is for John to do that.
    d. For John to do that was expected by no one.

(73) If there is a Raising rule which applies only to structures like Figure 11, we will get contrasts like these:
    a. John was believed to have done that.
    b. *John was wanted to do that.

However, this conclusion will follow only if we assume (as does Bresnan) that for is deleted after the operation of Passive, or that something like P's Clause-mate Condition prevents Passive from dipping into a clause.

Let's contrast the two approaches used by Bresnan and by P. Bresnan sets up a number of different underlying structures, with specific rules applicable to them. Differences in the syntax of the lexical items follow from the joint assumption of rules and structures (and some general principles). In support of her analyses, she shows differences in the interpretations of the different structures, and correlations between these semantic differences and the various surface structures; e.g. expect, as noted, goes with all three structures. Thus a sentence like 74 has three distinguishable senses:

(74) I expect you to be here at three.

One sense (correlated with Figure 10) is weakly intentional (like want). And indeed this sense is selected when we contrive to retain the for:

(75) What I expect is for you to be here at three.

The second sense (like challenge, Figure 9) is compulsive; the third merely predictive. Both of the latter have corresponding passives:

(76) You are expected to be a good boy.
(77) There was expected to be a fire.

In contrast, P minimizes the number of underlying categories, structures, and rules; he relies on devices such as rule features and global constraints to account for the facts.

---

\(^8\) Bresnan argues against Rosenbaum because his analysis would require a violation of her fixed-subject constraint; but I am not sure whether her revised version of that constraint (Bresnan ms) would apply to a rule like Rosenbaum's, which would probably require reference to a COMP in the structural conditions.
Furthermore, P holds to a theory under which, in some sense, logical structure determines (is) remote structure; but when undeniable differences in logical structure show up in different surface manifestations of a single remote structure, he appeals to a distinction between ‘core meaning’ and ‘assumptions’ (presuppositions or entailments?), as well as to global and other principles which determine, e.g., scope relationships. Bresnan specifically denies the proposition that meaning determines underlying structure; however, her analysis is quite consistent, in regard to the facts covered, with a theory in which deep structure determines logical structure (though I do not mean to attribute this view to her).

‘Semantic facts’ play a role in both frameworks, but in quite different ways. For Bresnan, the postulation of a difference between two structures leads to no predictions about meaning differences, since the same logical structure can be manifested by quite disparate deep (and surface) structures. On the other hand, a semantic difference between two surface structures must be attributed to a difference in deep structure (on the ‘standard theory’ view that I am temporarily, and probably counter-factually, attributing to her). For P, synonymous surface structures must be derived from the same underlying structure (give or take the other principles alluded to above).

Finally, as noted before, P seems to believe that we have direct access to ‘independently needed’ representations of logical structure that are known to exist.

5.3. THE BEST ANALYSIS. Let’s assume a theory in which underlying structure reflects logical structure in some way that we want to specify more closely. We can place the following requirements on our theory:

(78) a. Logical structure determines truth conditions and entailments.
   b. Two sentences have the same underlying structure if and only if they have the same logical structure.

On such a theory, that + S and for + S complements cannot be derived from the same source. Note the differences between these pairs:

(79) It offends him that men act weak. [example from Bresnan]
(80) It offends him for men to act weak.

A continuation such as but fortunately, they don’t renders 79 contradictory or truth-valueless, but not 80. Note also:

(81) What I like is that you are here.
(82) What I like is for you to be here.

Again, the difference is clear if we continue (and I wish you were). If we don’t want to bifurcate lexical items like offend and wish, but want to maintain 78a–b, then we must assume different sources for the two complement types.

Now let’s note a few differences in the privileges of occurrence of the two types (some of the following are from Bresnan):

(83) That the earth is flat is true.
(84) *For the earth to be flat is true.
(85) *That you are here is imperative.
(86) For you to be here is imperative.
(87) *That people love their children is common.
(88) For people to love their children is common.
(89) *That people love their children would be crazy.
(90) For people to love their children would be crazy.

These feel like semantic differences, but the proof of the pudding is in the analysis. What is needed is a semantic characterization of the difference between complements with that and for which will account for these and other facts.
Let's follow a time-honored tradition and say that *that + S* is the name of a proposition, which is a function from possible worlds and times to truth values. Hence 83 is OK, but 87 is not. A proposition is a kind of thing that can be true, not one that can be common. Why not? Because it is a kind of thing that is unique and constant across possible worlds. Thus 87 is funny for the same reason as 91, and 88 is not funny for the same reason that 92 is not:

(91) *Fido is common.*
(92)  

Barbara Partee has pointed out to me that not all *for*-clauses sound good with adjectives like *common*:

(93) *For John to love his wife is common.*
(94) *For Algernon to have left work at 6 on the 28th of February, 1776, was common.*

I think this is because we think of *love* as a more or less fixed relation holding between individuals, not something that stops and starts; and we think of the world in such a way that a single event-at-a-time can't happen more than once (and hence can't be common or uncommon).

So what kind of a thing is denoted by a *for*-clause? We can get a clue by examining the distribution and interpretations of various nominals. Consider first abstract nominals:

(95) Murder is common, undesirable, imperative etc.
(96) *That murder was common.*

Although I’ve starred 96, you can almost impose an interpretation on it by thinking of *that murder* as a *kind* of murder.

The examples below show different interpretations for count-noun phrases:

(97) An elephant in the living room would be odd. [Bresnan]
(98) *The elephant in the living room would be odd.*
(99) Cats with only two feet are odd.
(100) Cats with only two feet would be odd.

Jespersen (quoted by Bresnan) writes: ‘in nearly all sentences the combination *for* and an infinitive denotes some vague possibility or something imagined’ (1940:304). Bresnan characterizes *that*-clauses as being more specific and definite than *for*-clauses, and draws parallels (like those illustrated in 97–98) with definite and indefinite NP’s. Many of the properties shown above are also reminiscent of the interpretations of indefinite plurals as names for kinds (Carlson, ms), and of gerunds without determiners as names for kinds of actions, states of affairs etc. (cf. Schachter).

Suppose we say that a *for*-clause denotes the kind of thing that a *that*-clause names an instance of. Then we would have a proportion like this:

(101) Fido : dogs :: that John loves Mary : for John to love Mary

The relationship between Fido and dogs is that of a member to a class, or an instance to a kind. Just as kinds of things can be only virtual, so kinds of events or states of affairs can be virtual, hypothetical, non-existent. The following syllogisms are valid:
(102) Dogs have four legs.
   Fido is a dog.
   Fido has four legs.

(103) For John to love Mary would be nice.
   (That John loves Mary is for John to love Mary.)
   That John loves Mary is nice.

The trouble is that we don't have any comfortable way to state the second premise, and our tenses and moods mess things up. Try this:

(104) Flying horses would have wings.
   Pegasus is a flying horse.
   Pegasus has wings.

We have no good names for virtual objects, but we can make them up:

(105) Round squares would be impossible.

When we use common nouns as names for kinds in the indefinite plural, we always sound as if we don't doubt their existence—unless we're specifically denying it. For-clauses seem to be just the opposite. Maybe that's a difference between virtual states of affairs and virtual objects. We have a common linguistic means to differentiate virtual from 'actual' propositions, but we have none for common NP's.

The trouble is that we don't have any comfortable way to state the second premise, and our tenses and moods mess things up. Try this:

(106) The man that you see is here.
(107) The man for you to see is here.

Of course, we don't know whether the for in such sentences is the preposition or the complementizer. But there seems to be something here very much like the difference in semantic force that we have been struggling to characterize. Note the following:

(108) The man that she will marry doesn't exist.
(109) The man for her to marry doesn't exist.
(110) I'm looking for a man for us to hire, but he may not exist.
(111) I'm looking for a man that we will hire, but he may not exist.

The sentences with that sound odd, but the others don't, and for the same reason (I think) that the contrasts we have noted occur. The for-clause expresses an eventuality.

Verbs and adjectives that can be construed with that-clauses are those dealing with propositions, relations among them, and relations between individuals and propositions (in some cases, sentences in place of propositions): say, know, believe, be true, prove, entail, contradict, deny, mean, be odd, regret. Verbs and
adjectives that go with for-clauses are those appropriate for talking about eventualities, and are often volitional: desire, wish, want, intend, mean (in another sense), be imperative, and many predicates when used with the subjunctive. Bresnan notes contrasts like these:

(112) *For him to eat cabbage means that he will be sick.
(113) For him to eat cabbage would mean that he was of low birth.

Why is it odd to use that-clauses with the second class of predicates? Because you can’t want (intend, desire) propositions or sentences. What you CAN want is for an eventuality to be actualized, for the world to be such that a proposition be true. Thus we find that abstract nominals line up in the right way with names for propositions (or sets of them) other than that-clauses:

(114) She \{ believes \{ proved \} \} that proposition
proved
that fact.
the Pythagorean theorem.

*that state of affairs.
*that contingency.
*this eventuality.
*the Ten Commandments.

(115) He \{ desires \{ wants \} \} *that proposition.
wishes
*that fact.
*the Pythagorean theorem.

that state of affairs.
that kind of a game.
that opportunity.
this eventuality.

(I don’t mean that all the collocations are equally felicitous. Clearly, a lot of different types of predicates are lumped together in this classification.)

Now what about the underlying structure for verbs like believe that occur with Raising? So far I have shown that a theory that accepts 78a-b must distinguish between for-clauses and that-clauses. The underlying structures for raised sentences might then be identified with either of the two types so far established, or they could instantiate some third type of underlying clausal element. I’ll argue that they should be identified with neither of the established types.

First, as to for-clauses. Because analyses like Bresnan’s incorporate a rule of for-deletion (triggered by an immediately preceding verb), it is difficult—but, I hope to show, not impossible—to pin down differences in the interpretation of superficially similar sentences. Let’s revert to the example of expect. First, to exclude the case in which expect behaves like persuade, we choose an inanimate or contentless subject for the complement sentence, e.g. there or the ‘weather’ it.

Two senses of expect remain. One, very close to the persuade type, is volitional (like want); this is the type that we will correlate with the underlying for-complement. The other is merely predictive (this is the type that Bresnan has analysed as going with a bare S, and we will correlate this with the raised structure). The two senses are exemplified as follows:

(116) I expect there to be no talking while I’m out of the room.
(117) I expect it to rain when we have our picnic.

If so, we should be able to choose a single sense when for is present. This seems to be correct:

(118) What I expect is for there to be no talking while I am out of the room.
(119) ?What I expect is for it to rain when we have our picnic.
Admittedly, this judgment is tenuous. But in a moment we will see another clear difference that can be explained only if we distinguish between the two types (and assume that there is no Raising with W-verbs).

If underlying structures are determined by logical structure, then we would expect raised structures to show a similarity, not with for-clauses, but with that-clauses. After all, it was precisely because of the closeness in meaning between these two types that linguists were led to assume a basic identity of underlying structure in the first place. It is very difficult to discern any difference in meaning between raised structures and that-clauses in many pairs of sentences:

(120) John believes that the earth is flat.
(121) John believes the earth to be flat.

But there are differences. P judges sentences like the following to differ in their scope characteristics:

(122) I believe that a unicorn is eating my radishes.
(123) I believe a unicorn to be eating my radishes.

Here 122 appears to support two interpretations, but 123 only one (wide scope). On the other hand, P feels that 124 can support two interpretations again:

(124) A unicorn is believed (by many) to be eating my radishes.

Clearly, the truth conditions for these different interpretations are different, but P invokes a principle referring to derived structure to explain the difference between 122 and 123 (thus departing from 78a-b). A consequential application of 78a-b would lead us to conclude that 122 and 124 could be derived from the same sources, but not 123.

It seems to me that these matters are not as clear as P may think, and that we cannot take unclear cases like 123 as crucial. If narrow scope is to be ruled out for that sentence, we must also rule it out for the following:

(125) No one believes a gold mountain to exist.

So I am inclined to say that the preference of some speakers for P's reading of 123 results from non-grammatical factors; hence I agree with P that 122–24 could have the same source. (On the other hand, the fact that W-verbs do not share this property could be taken as a further argument against deriving raised structures from for-clauses.)

P also notes meaning differences between that-clauses and raised structures in sentences like those given above as 26–27 and 63–64:

(126) I found Julius Caesar to be boring.
(127) I found that Julius Caesar was boring.

He attributes this to a difference in presuppositions (‘assumptions’), to be left for some future analysis to explain.

But another clear contrast between sentences with and without Raising cannot be so conveniently swept under the rug of future advances. Note these sentences:

(128) I believe that John ate so much that he got sick.
(129) I believe that so much was eaten by John that he got sick.
(130) I believe John to have eaten so much that he got sick.
(131) I believe so much to have been eaten by John that he got sick.
(132) So much was believed to have been eaten by John that he got sick.

Here 131–32 are acceptable sentences, but seem to require the odd reading that John's troubles were not a function of how much he ate, but of how much someone believed he ate.

W-verbs do not share this property. Although there are some problems about mood and tense, sentences like these are ambiguous:

(133) I want you to eat so much that you'll get sick.
(134) She wants to marry so many men that she'll get rich.

P's A-verbs act like believe in this respect:

(135) It seems that so many people came to the party that it was a brawl.
(136) John seems to have invited so many people to the party that it will be a brawl.
(137) So many people seem to have come to the party that it will be a brawl.
Here 135–36 are ambiguous as to the scope of the *that*-clause, but 137 is unambiguous. The connection of cause and result is drawn between the number of people that *seem* to have come and the *that*-clause; the *that*-clause cannot be in the scope of *seem*, as it can in the first two.

With respect to such facts, the Raising verbs *seem* and *believe* act like traditional Equi verbs such as *try* and *persuade*:

(138) John tried to eat so much that he got fat.
(139) John tried to eat so much that he would get fat.
(140) So many people tried to come to the party that it was (*would be*) a brawl.
(141) Mary invited so many people to come that it got boring.

The meaning of all *so–that* sentences can be roughly characterized as follows. There is a presupposition: if $p$ to a certain degree $x$, then $q$. If the sentence is declarative, it conveys an assertion: $p$ to the degree $x$, and (hence) $q$. The differing interpretations of the *so–that* sentences with and without Raising are associated both with different presuppositions and different assertions. Hence it would seem that a theory embodying 78a–b would either have to admit different structures for them, or somehow ensure that some manifestations of certain underlying structures would be blocked.

The facts are puzzling on any account. They may be described as follows (restricting ourselves to the NP-to-PRED type):

(142) *believe*: If so is in the NP object position or in the subject NP (with Passive), the result clause is interpreted as outside the scope of the matrix verb. If so is in the PRED, either reading is possible (give or take mood, tense etc.)
(143) *persuade*: Same as *believe*.
(144) *want*: If so is in the NP object position (to beg the question, temporarily), then the result clause can be interpreted as within or outside the matrix verb. If so is in the PRED, again either reading is possible. (There is no Passive.)

In the next section, I will suggest an explanation for these facts within a non-standard theory. For the moment, let’s consider the alternatives available within either a Postalian ‘natural’ grammar or an interpretive theory like Chomsky’s.

In a theory following 78a–b, we would have to set up different underlying structures for the different interpretations. To ensure the right pairings of surface structures and interpretations, some restrictions (hopefully of a general sort) would have to block some derivations. Presumably P would treat sentences like those we have looked at by positing higher structures with abstract verbs, much like those used in his analysis of comparative sentences (Postal 1974). To say that John ate so much that he got sick is to say something about the amount that John ate in much the same way as a comparative sentence like the following:

(145) John ate more food than you can imagine.

The scope ambiguities we have noted could then be treated by positing underlying structures where the requisite predicates were below or above the relevant verbs. But then the contrasts we have noted would force us to posit some difference in underlying structure between *want* and *believe* (to say nothing of the various structures we would have to associate with the same verb).

And as far as I can tell, no known principles of extractibility or islandhood cover precisely the cases we are concerned with.

Would an interpretive theory be any better here? Let’s assume that *so–that* constituents are generated in the DET (much like Bresnan’s comparatives), and that there is a rule of Result Clause Extraposition (RCE). The rule is obligatory, or else there is a surface filter marking as ungrammatical any sentence with such a clause remaining in the determiner (Williams 1975). An interpretive rule would then provide different logical forms for such sentences. In the case of a *so* (plus trace) in the PRED of a sentence under *believe*, *persuade*, *want*, *seem*, *try*, it would somehow stuff the interpretation back down into the scope of the matrix verb, as one option—or give it an interpretation in which it was outside the scope of the matrix. The former option would have to be blocked in case the *so* was in the DET of the NP object of *persuade* and *believe* (but not *want*) or in any subject NP. Again, it would be hoped that some general principles might be invoked; but I know of none that would apply. In fact, given the assumptions implicit in the Chomsky–Bresnan analysis, several of Chomsky’s general conditions seem to give just the
wrong results. In any case, it seems that an interpretive theory will have to recognize a three-way contrast in the transitive verbs involved. In these respects believe acts like persuade and unlike want; but in other cases (There-Insertion etc.), believe and want should be treated as alike.\(^9\)

Before turning to an alternative theory, let's consider whether the contrast between bare S's and \(\bar{S}\) would be of any help here. Williams argues that RCE is a rule applying only at the \(\bar{S}\) level. As far as I can tell, no prediction whatever about these facts arises from this assumption and Bresnan's analysis. RCE will not apply within the S cycle, but only at a later stage, when we reach the \(\bar{S}\) containing the matrix. But on Williams' assumption, Bresnan's analysis predicts that the clause will be extraposed (from any position) past all the lower-level elements of the matrix clause, and this seems to be incorrect:

\begin{align*}
(146) \text{The doctor believed John to have taken so many pills last night that he got sick before taking his temperature.}
\end{align*}

\begin{align*}
(147) \text{The doctor believed John to have taken so many pills before going to sleep that he got sick.}
\end{align*}

\begin{align*}
(148) \text{The doctor believed John to have taken so many pills before going to sleep that he got sick before taking his temperature.}
\end{align*}

Example 148 is pretty clumsy, but I think OK; and it shows that Extraposition must have happened 'downstairs', on the Williams account.\(^{10}\)

5.4. THE BEST THEORY. A principled explanation for the facts we have just noted can be found within a modified form of classical transformational grammar which includes generalized transformations that not only embed sentences into sentences, but embed NP's into sentences. I have argued elsewhere (Bach 1976, ms, b) that this theory is more restrictive than the standard theory (and all its progeny), and that it allows a much more elegant framework for treating matters of scope and quantification—since it is able to incorporate such distinctions in a way that does not lead to the ad-hoc postulation of a distinct configuration, whether 'surface' or underlying, for every different sense of a sentence with quantified NP's.

Briefly, the main idea is that the syntactic and semantic build-ups of a sentence go hand in hand. As each rule is applied, a corresponding translation rule is applied to the already built-up 'semantic' representation. Different derivations can then be

\(^9\) I am grateful to Roger Higgins for pointing out to me that the facts about 'peculiar' interpretations of result clauses were documented in an unpublished paper by him (Higgins, ms), and for allowing me to read his paper. He not only notes the phenomena in sentences with S-to-O Raising, but also in sentences with fronting of the so-phrase: So many pictures did John intend to paint that we were surprised (where it is the number of pictures that he intended to paint that is at issue). He suggests an explanation on the basis of an interpretive principle which makes the result contingent on the meaning of the sentence to which the result clause is attached. Hence his argument supports the general conclusion that there is a difference in the structures associated with W-verbs and B-verbs. Since so-phrases in the predicate of both W-structures and B-structures can be associated either with the inner sentence or the outer sentence, the interpretive rule must either be framed in terms of a reference to S, or else both types must be construable with \(\bar{S}\). I consider the question of the correct explanation still open, and will continue to do so until someone has constructed a syntax and semantics for result clauses (in line with my remarks at the end of this essay).

\(^{10}\) In general, Williams' examples purporting to show that RCE can't apply in gerundives and nominals are convincing to me only if the so is in the subject NP:

\begin{itemize}
  \item [(a)] *so many people's coming to the party that it was a bore ...
  \item [(b)] *so many senators' approval of the motion that it was surprising ...
  \item [(c)] Germany's bombing of so many cities that British morale increased ...
  \item [(d)] the destruction of so many cities by the Germans that British resistance was strengthened ...
\end{itemize}
associated with different interpretations (much as in the theory of Montague grammar). For example, the two senses of 149 are associated with two different derivations:

(149) John is looking for a unicorn.

In one (narrow scope), the sentence is built up directly in the kernel rules. In the second, an open sentence with a variable is generated by the kernel rules, and a rule of Quantification inserts the already built-up NP in place of the variable.

The necessary ingredients for an explanation of the scope facts noted above are these:

(i) A rule that embeds sentences into the object position of sentences with verbs like believe. Without evidence to the contrary, I'll assume that the rule has the effect of P's Raising rule, rather than operating on already embedded structures. It looks like this:

(150) R-embedding

\[
\begin{array}{cccc}
S_1: & X, & V-R, & that_i, & Y \\
1 & 2 & 3 & 4 & \Rightarrow 1, 2, 5 + to + 7, 4 \\
\end{array}
\]

\[
\begin{array}{ccc}
S_2: & NP, & TNS, & Pred \\
5 & 6 & 7 \\
\end{array}
\]

\(\text{(That}_i\text{ is a propositional variable.)}\)

(ii) A similar rule for A-verbs like seem.

(iii) A rule for embedding sentences into the object position for W-verbs like want:

(151) \(\begin{array}{cccc}
S_1: & X, & V-W, & that_i, & Y \\
1 & 2 & 3 & 4 & \Rightarrow 1, 2, for + 5 + to + 7, 4 \\
\end{array}\)

\[
\begin{array}{ccc}
S_2: & NP, & TNS, & Pred \\
5 & 6 & 7 \\
\end{array}
\]

The associated translation rules say to form the lambda abstract\(^{11}\) over that\(_i\) from the translation of S\(_1\), and apply it to the intension of the translation of S\(_2\); this is the general form of the translation rule for sentence-embedding rules. (I am ignoring the problem of supplying have + EN for past tenses.)\(^{12}\)

(iv) A rule for forming sentences with so-that clauses in them. I won't try to state the rule, since I don't understand the details; but I assume that the rule gets so-that S into NP's, and that there is an obligatory singulary transformation that flips the that-clause to the end of the sentence.

\(^{11}\text{I.e., an expression denoting the set of propositions } p_i \text{ such that } \varphi, \text{ where } \varphi \text{ translates the first sentence.}\)

\(^{12}\text{These rules are actually not formulated in a way precise enough to indicate the difference in derived structure between the embedding under a B-verb and that under a W-verb, and it is not clear to me at present exactly how the rules should work. A brute-force solution would be to substitute the sequence NP + to + Pred for the entire NP dominating that\(_i\) in the case of B-verb embedding, and to substitute the entire constituent sentence for that\(_i\) in the case of W-verbs.}\)
A freezing principle of some kind which says (at least): An NP from which an element has been extracted may not be moved by any other rule.\textsuperscript{13}

Now we can show why the scope facts line up the way they do. This analysis (offered only as a first approximation) predicts that wide and narrow scope interpretations of result clauses associated with embedded sentences will be possible just in case the NP with \emph{so} does not move from its original position. This will be true for all NP's in the PRED, and for the NP immediately following \emph{want}-verbs. On the other hand, there will be only one possible derivation for raised structures—namely, one in which the embedded sentence has a variable in subject position, which is then raised, has a NP 'quantified in', and undergoes the \emph{so–that} rules. A sketch of a derivation with \emph{believe} runs like this:

\begin{enumerate}
\item[(152)] Kernel: \begin{align*}
S_1: & \quad \text{John believes that,} \\
S_2: & \quad x \text{ came to the party} \\
S_3: & \quad \text{it was a brawl} \\
\text{NP:} & \quad \text{\textemdash-many people}
\end{align*} \\
\begin{align*}
S_1 + S_2: & \quad \text{John believes } x \text{ to have come to the party (} S_1' \text{)} \\
\text{NP + } S_3: & \quad \text{so many that it was a brawl people (} \text{NP'} \text{)} \\
\text{NP'} + S_1': & \quad \text{John believes so many that it was a brawl people to have} \\
& \quad \text{come to the party} \\
\text{RCE:} & \quad \text{John believes so many people to have come to the party that} \\
& \quad \text{it was a brawl.}
\end{align*}
\end{enumerate}

The freezing principle will block a derivation in which the \emph{so–that} rules have already applied in the constituent sentence.

For the explanation to work, it is essential that \emph{W}-verbs not undergo the same rule as \emph{B}-verbs, i.e. that the embedded sentence remain intact. In the next section, we will briefly examine P's arguments for Raising with \emph{W}-verbs.

It remains to be shown that the principle I have used has independent justification. There is actually ample evidence for the need of such a principle. Thus NP's from which a relative clause has been extracted by Extraposition can never be topicalized:

\begin{enumerate}
\item[(153)] *that guy I saw yesterday who sold me this lemon ... (= that guy who sold me this lemon I saw yesterday ...)
\end{enumerate}

Similarly, NP's with \emph{so}:

\begin{enumerate}
\item[(154)] *so many people I saw that you'd never believe me ...
\end{enumerate}

Comparative elements may be topicalized if they are intact, but not if they have had the compared clause extracted from them:

\begin{enumerate}
\item[(155)] *As many things as you like you can put into this closet.
\item[(156)] As many things you can put into this closet as you like.
\end{enumerate}

The freezing principle says that when Extraposition of \emph{it} has occurred, the \emph{it} cannot be moved—and hence that sentences like 157 have the derivation Movement–Extraposition, rather than the opposite:

\begin{enumerate}
\item[(157)] It seems to be true that he's here.
\end{enumerate}

Such sentences were used by Ross to argue for last-cyclic Extraposition. The Extraposition is last-cyclic (in this sentence), but by virtue of our principle rather than as a special property of

\textsuperscript{13} Culicover & Wexler ms have proposed a freezing principle within a more or less standard framework, but I am not sure whether it generalizes to the case at hand.
the rule. It follows that, when Extraposition is obligatory ‘downstairs’, raising the it will produce a bad sentence:

(158) *I believe it to seem that it’s raining.

(159) It’s expected to seem likely that it’s raining.

The principle also predicts that sentences like 160 arise by a derivation Passive–Extraposition, rather than the reverse:

(160) It’s believed by many that the earth is flat.

Many writers have argued for this derivation.

The principle predicts the order of application WH-Movement–Extraposition, rather than the reverse (cf. Ross 1967):

(161) Who do you know that’s from India?

(162) ... a man who I know from India ...

The latter phrase must have a structure derived from an independent PP with an NP, rather than from a man [I know [a man from India]]. Again, this seems right.

If all this is correct, then we have an explanation for the difference between 158 and a parallel sentence with want:

(163) I want it to seem that we’re happy.

This difference can be explained in this way only if there is no Raising in sentences with W-verbs.14

5.5. W-VERBS. Now we must examine P’s evidence for Raising with W-verbs like want, since the above analysis depends crucially on the absence of such an operation in these verbs. For want, at least, I find P’s presentation quite unconvincing. According to his theory, the following sentences should be bad:

(164) I want us to get there on time for once.

(165) I want Bob alone to win.

(166) I want not many people to be there when I arrive.

(167) I want him, to be happy even more than Bob, does.

(168) I wanted but no one else wanted—John to be successful.

(169) Bob wants Ted to win and Bill John to lose.

(170) What he wants is us to be there early.

14 The framework of assumptions sketched here suggests a possible explanation for the facts noted in fn. 9 above. Suppose the underlying structure for gerundives and nominal phrases is actually one in which the agent phrase comes after the verbal or deverbal element (this analysis is defended in Stockwell, Schachter & Partee 1973). Then the freezing principle would block derivations like this:

(a) the destruction of so many that S cities by NP
   RCE the destruction of so many cities by NP that S
   NP-Preposing *so many cities’ destruction by NP that S.

What remains unexplained on this account is why it would not be possible to have a derivation with x in place of the NP containing so, quantification of an NP with so, and RCE operating in the matrix. The status of sentences like the following is unclear to me:

(b) So many people’s views on Raising have been published that I’m sick of the whole subject.

(c) So many people’s discussing Raising surprises the editor that you wouldn’t believe it. I find (b) fine and (c) very bad; but there may be a different reason, a constraint against having indefinite plural subjects for gerundives, action nominals, and deverbal nominals (when interpreted ‘verbally’):

(d) *Children’s being liberated would surprise me.

(e) *Children’s berating of their parents 

(f) *Women’s liberation in this century ... (in the sense of the action of liberating women)
I find the first three perfect, and the rest too cloudy to hang a theory on. Sentence 170 is pretty bad, but there's a perfectly good explanation for it in terms of for-deletion: for-deletion occurs only in the immediate context of verbs like want; therefore, under a number of reasonable alternatives, the only good pseudo-cleft is going to be one with for. A similar possibility might account for the relative unacceptability of 168–69. Unfortunately, without more careful analysis than P provides (i.e. more detailed structures and rules), it is difficult to evaluate much of the evidence.

5.6. The best argument for the hypothesis that verbs like believe occurring in raised structures start out with bare S complements (reverting now to a standard framework) is the fact that they never occur with for. This is a very simple argument. It depends on the methodological principle that you don't posit more structure than you need. This kind of consideration rules out infinitely many hypotheses that no linguist (or child) would ever think of. For example, no one would seriously propose that every sentence contained in its deep structure the NP My God! Whenever obligatorily deleted items are hypothesized, it is considered necessary to motivate the choice. The early assumption that all infinitivals contained for was motivated by the existence of other sentences with other verbs which did contain for; but this hypothesis has been shown to be incorrect.

Although little attention is paid to the question, Bresnan assumes that (possibly unlike S) bare S's do not come from NP's. This is supported by the fact that bare S's do not undergo typical NP rules like Passive, Pseudo-cleft, or Topicalization. Further support for this analysis comes from a number of predictions that it correctly makes:

(a) Bare S's cannot be extrapolosed.
(b) No verb subcategorized for any complementizers will ever occur with raised structures. For example, verbs taking indirect questions will never trigger Raising.
(c) S-domain transformations and S-dependent categories will not occur in raised structures except by virtue of the fact that they occur in the superordinate structure.

What support is there for the bare S analysis, as opposed to the No-Rule Analysis entertained and rejected above? Most of P's arguments are against the Chomskyan assumption that the S remains intact. Many of them do not distinguish the Raising analysis from an analysis that assumes the same structure for persuade and believe. Here I believe the classical arguments are the best:

(i) The S analysis predicts that S-domain transformations and structures will occur in raised structures (there, S-adverbs etc.)
(ii) The S analysis, without any special assumptions, predicts the differences between persuade and believe with respect to semantic interpretations and relative independence of the NP object (there, tabs etc.)
(iii) It is likely that the scope differences noted by Ioup 1975 will provide further support for the difference; but this evidence must await a satisfactory explicit account of the interpretations of quantifier scope. Ioup shows that surface information is inadequate. If believe and persuade end up (as I believe is correct) with the same surface structures, then some other difference must be drawn upon to explain the differing interpretations of quantified NP's:
I persuaded a student to visit every professor.
I expected an egg to be eaten by every customer.

5.7. **You too can formulate a rule of raising.** How you do it will depend on what you think rules can do. If you are a structure-preserver, you will probably change the scheme just argued for, assume that believe is categorized to occur with an empty object and an S, and write the rule like this (or just say ‘Move NP’):

(173) W, NP, X, [s NP, Y], Z
1 2 3 4 5 6 ⇒
1 4 3 0 5 6

Note that this one rule suffices for S-to-O and S-to-S raising.

If you are content to have two rules, then S-to-O raising might look like this:

(174) X, V, [s NP, Y], Z
1 2 3 4 5 ⇒
1 2+3 0 4 5

where 2+3 signifies right sister adjunction.

The extent to which rules of such simple forms are possible will be a direct function of the extent to which you are able to discover independently motivated general principles that will rule out misapplication of the rules.

5.8. **You can do without a rule of raising, if you are willing to generate structures like V NP to VP directly for both persuade and believe.** What you will need then is some ‘interpretive’ rule that tells you about the difference between the two verbs. For example, we might adopt a rule like this:

(175) Given a structure V235 NP to VP,
interpret it as V235 (Ax[NP]),
where V235 = persuade, force etc., and VP’ etc. are the translations of VP etc.

Whenever the NP in question is an empty one like there, tabs, or an inappropriate one like stone, the rule would give you an anomalous or empty reading. This rule would not apply to verbs like believe; for such verbs, we could supply a meaning postulate telling us that the truth conditions for structures like A believes B to X are the same as those for A believes that B X’s. Note that it is the latter rather than the former that is problematical, since it requires a rather powerful theory of meaning postulates. In terms of something like the extended standard theory, using interprettive principles, it requires something like a transderivational interpretive principle. It is also obviously oversimplified, since an adequate rule would have to deal with the problem of linking 176 to 177 as well as 178:

(176) I believe John to have said that.
(177) I believe that John said that.
(178) I believe that John has said that.

6. **Frameworks, frameworks.** In §5.4 I tried to argue that consistent application of P’s assumptions about the relationship between logical and syntactic structure leads to a system like that proposed by Bresnan, where a number of superficially similar structures have different remote structures. Of course we don’t need quite as strong a hypothesis to get the same result. We only need half of 78b:
Two sentences have the same underlying structure only if they have the same logical structure. This is, of course, something like our old friend, the Katz–Postal hypothesis. Nowadays, of course, all bets are off. Nobody seems to have a worked-out theory that predicts anything about the relationship between underlying structure and logical structure. P certainly doesn’t, since he lumps together all kinds of disparate semantic classes of verbs and gets the correlations by appeal to a large number of inexplicit devices: surface filters, global constraints etc. And, as I have noted, he doesn’t provide anything approaching a characterization of the nature of logical structure. But the same is true of his principal opponent. In Chomsky’s latest work (1975, ms), there is assumed to be a level of representation called ‘logical form’, defined as the level which results from the application of all rules of ‘sentence grammar’ (transformations and so-called interpretive rules). It is a kind of mixed representation consisting of brackets, quasi-logical operators, traces, and English morphemes. There is no independent characterization of the notion ‘possible logical form’, and hence no predictions can be made about the relationship between underlying structure and logical form; in fact, it is not at all clear that the notion of underlying form plays any essential role in the system at all (cf. Bach, ms; Lightfoot, ms). Finally, those linguists and philosophers who have worked within the tradition of formal semantics (e.g. Montague and followers) have for the most part been uninterested in the question of what is a possible natural language. To the extent that their systems make any predictions at all about the connections between syntax and semantics, they are so general that they fail to distinguish between natural languages and arbitrary formal systems. (An exception is Partee ms.)

The moral of all this is that, in the current state of our art or science, there are simply too many alternatives available—both in the choices for analyses within a fixed framework, and in the choice of frameworks. I believe there is a reason for this situation, and my next section is devoted to a discussion of it.

6.1. WHATEVER HAPPENED TO TRANSFORMATIONAL THEORY? In the opening section of this review, I noted that in P’s entire book there is no formulation of the rule to which the book is devoted. This is not an untypical situation in present-day ‘transformational’ theory.

When *Syntactic structures* came onto the linguistic scene, now almost twenty years ago, its most exciting idea (for many of us, at least) was the concept of a grammar as a set of explicit rules defining the structures associated with the sentences of a language. This idea, which seems simple and obvious today, was truly revolutionary, at least in the modern linguistic tradition. The way to achieve insights about language was to test out hypotheses about the possible form of grammars, and to construct actual grammars for given languages. Of course, everyone realized that particular attempts to construct grammars were fragmentary, often plainly wrong; still, the conclusions that they were wrong, and conclusions about the nature of language that might be gleaned from these attempts, rested squarely on the ideal of explicitness. A grammar was to be an explicit representation of what native speakers appeared to know about their language. Early discussions of linguistic theory (e.g. Chomsky 1955, 1957, 1962; Lees 1960; Klima 1964) actually included reasonably rich sets of rules.
Besides the commitment to explicitness in providing testable sets of rules, there was a commitment to explicitness in general linguistic theory. Chomsky 1955 is an example. Just as it was intended to be possible to take a set of rules and actually derive structures for sentences to be matched up with intuitions of native speakers, it was supposed to be possible to take a grammar and tell whether or not it instantiated some particular theory of grammars. A lot of early work was devoted to trying to make explicit the fundamental assumptions about the form of grammars that were present in various then-current models (e.g. Chomsky 1956, Postal 1964). P’s book and a good deal of current work under a variety of frameworks show that these ideals have been lost. What happened?

6.2. THEORIES WITHOUT GRAMMARS, GRAMMARS WITHOUT RULES. The basic test of native-speaker intuition, available for early work in generative grammar, was the judgment of grammaticality. This followed from the nature of the early theory. Except for phonology, everything about sentences had to be expressed in the syntax. The lexicon was simply a continuation of the phrase-structure rules. There was no semantic component. Hence, only one kind of judgment was really relevant to testing a set of rules. Judgments about ambiguity and anomaly were a kind of extra, which could be brought in to make the final product more convincing.

As everybody knows, this all changed with the developments up to and around Chomsky’s Aspects, the advent of the so-called ‘standard theory’. The most significant change was the addition of semantics to the concerns of the grammarian; but an almost equally significant change was the development of the lexicon and the associated addition of complex elements to the syntax. These were surely welcome and necessary changes. What was not so desirable was that the ideal of explicitness at both levels began to be eroded. Various theories of the lexicon made it possible to explicate certain facts outside the syntax proper, and the development of a framework for semantics had the same effect. But whereas linguists had earlier felt it incumbent upon themselves to provide reasonably rich fragments for natural language syntaxes, somehow this necessary work was never carried on in the new areas. We were given ‘the structure of a semantic theory’, but no semantic theories for significant portions of a natural language. We have had numerous theories of the lexicon, but no lexica.

I think it was in the late sixties that I began to hear the opinion voiced that there was no point in trying to write rules—since all known rules were so imperfect, and the complexities of language so vast, that it would be premature or impossible actually to capture natural language by any kind of explicit rules. Thus the practice of the ‘semanticists’ infected syntax as well. The culmination of this trend can be seen in the present book: hundreds of facts and arguments, no rules.

In a recent paper, Postal (MS), replying to a similar criticism made by Bresnan 1976b, has actually defended this position by asking for evidence that writing rules provides any insights that cannot be gained without providing explicit rules. But there can’t be any evidence for this concept, because it is simply a postulate as to what linguistics is all about. No rules, no grammars. No grammars, no theory.

I would not exclude a good deal of my own work from the present criticism; but I don’t believe I have ever elevated inexplicitness to the status of an ideal.
6.3. GRAMMARS WITHOUT THEORIES. Lest I be thought to be taking sides, let me hasten to add some diatribical remarks against P's opponents, the proponents of the 'revised extended standard theory'. Curiously, the practical results of the recent work of Chomsky and some of his followers are about the same as those of the ruleless grammarians. Where P has no rules at all, Chomsky (ms) has extremely general rules like 'Move wh' or 'Move NP'. This is in line with the program of restricting the transformational component severely and putting the burden of the syntax into various other conditions, non-transformational rules, interpretive rules, and lexical redundancy rules.

The trouble here is not only that no examples of such extra rules are given (or if given, are given in such a general or vague form as to be untestable or plainly wrong), but also that there is no general theory of possible conditions, non-transformational rules, etc. etc., so that the general adequacy of the theory is untestable. (See Bach, msb, for some remarks on the latest results of this program.)

7. CONCLUSIONS. At one point, P characterizes his book as follows (p. 2):

'Although descriptive and theoretical questions form the chief elements of interest in what follows, this monograph can also [emphasis supplied] be regarded in part as a reference work in which a rather large number of restrictions involving Raising constructions and related areas are documented, even if they are not explained or even adequately described.'

I agree with this characterization, except for the word 'also', which I would replace by 'only'.

As to P's chief claims: I think that there is a rule of Raising into superordinate object position. I do not know whether it is the same rule as the one that raises subjects into subject position. I do not know whether it applies in all the sentences where its applicability is postulated by P, but I suspect not. I do not know at all what conclusions we can draw about the nature of language from the debate about Raising. The reasons I cannot draw any conclusions are:

(a) No explicit syntactic rules are given by Postal.
(b) No explicit 'semantic' rules are given by anyone participating in the debate.
(c) No explicit characterization is given by Postal or his opponents of the notion 'possible grammar'.

Clearly, there is a lot of work left to do.

REFERENCES


BACH, EMMON. 1976. 'The position of embedding transformations in a grammar' revisited. To appear in a volume of papers from the International Summer School of Computational Linguistics (Pisa), ed. by A. Zampolli.


—. msb. Comments on N. Chomsky's paper for the Irvine Conference on Formal Syntax. To appear in Culicover et al. (eds.)


BEVER, T. G.; J. R. LACKNER; and R. KIRK. 1969. The underlying structures of sentences are the primary units of immediate speech processing. Perception and Psychophysics 5.225–31.

BOWERS, JOHN. MS. The theory of grammatical relations. Unpublished (Cornell University).


----. MS. Variables in the theory of transformations. To appear in Culicover et al. (eds.)


----. MS. On wh-movement. To appear in Culicover et al. (eds.)


CULICOVER, PETER; TOM WASOW; and ADRIAN AKMAJIAN (eds.) MS. Formal syntax. (Proceedings of the 1976 conference at Irvine, to appear.)


HIGGINS, ROGER. MS. On the syntax of result clauses in English. MIT.


LIGHTFOOT, GORDON. MS. Trace theory and diachronic syntax. To appear in Culicover et al. (eds.)


—, 1973. The penthouse principle and the order of constituents. You take the high node and I'll take the low node, Papers from the Comparative Syntax Festival, 397–422. Chicago: CLS.

[Received 12 November 1976.]