

The American Revolution (in Grammar) and the Declaration of Independence (of Syntax)

By

John M. Dienhart

Odense University

In 1775 the British Parliament declared that the State of Massachusetts was in rebellion.

By a strange quirk of history, that same state has become the center of a new rebellion, with the Boston Tea Party replaced by the M.I.T. Party.

The Father of this new revolution is Noam Chomsky, and the manifesto a slim volume entitled *Syntactic Structures*, which appeared in 1957.

This book raised the study of syntax to new heights (some critics have suggested the opposite direction), and has provided the impetus for countless articles, books, and doctoral dissertations in the twenty years since its publication.

To understand something of Chomsky's concept of grammar, let us begin in a rather conventional fashion. Consider the sentence:

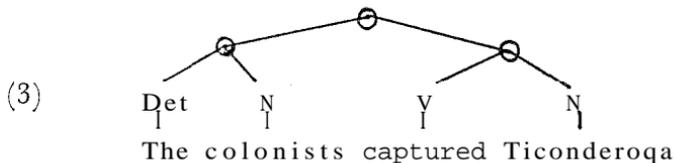
(1) The colonists captured Ticonderoga.

To each of the words in this sentence we can assign traditional word class labels (lexical categories) as follows:

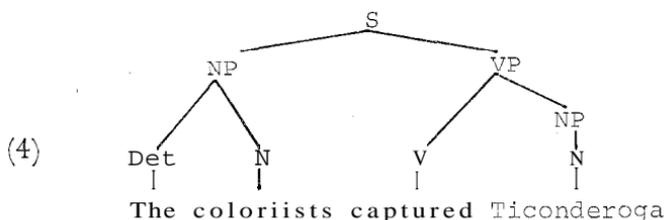
Det N V N

(2) The colonists captured Ticonderoga

where Det = Determiner, N = Noun, and V = Verb. Furthermore, we can provide this sentence with a rather traditional constituent analysis, which can then be displayed in the form of a tree diagram:



If we now label the new nodes (see the circles in (3)) with the (non-lexical) categories NP (for Noun Phrase), VP (for Verb Phrase), and S (for Sentence), we get the tree, or Phrase Marker, shown in (4):



In addition to such notions as word class membership and constituent structure, traditional grammar has also provided us with the concept of "grammatical relation." Thus we might say that *the coloriists* is "subject" of the sentence in question, and *Ticonderoga* is "direct object" of the verb *captured*. These relations, says Chomsky, need not be labeled directly, since they are definable in terms of the tree diagram we have drawn in (4): the "subject" will be that NP right under (i.e., immediately dominated by) S; the "direct object" is that NP immediately dominated by VP.

So far, then, our analysis differs little from that of traditional grammar. At this point, however, we can introduce one of the new features proposed by Chomsky: the concept of a *formal rule* of grammar.

The three (phrase structure) rules in (5) allow us to construct, or generate, a simple phrase marker (like that in (4)), down to the level of lexical category:

$$(5a) \quad S \rightarrow NP + VP$$

$$(5b) \quad VP \rightarrow V + NP$$

$$(5c) \quad NP \rightarrow (Det) + N \quad \text{where } () = \text{optional choice}$$

Ignoring, for the present, the question of just how lexical items are to be inserted, we see that these phrase structure rules allow us to generate structural descriptions for such sentences as:

(6) Washington crossed the Delaware

(7) Cornwallis surrendered his sword

(8) The colonists dumped the tea

Returning now to our sentence (1)—which I reproduce here, for convenience, as (9)—we observe that there are numerous other

sentences which, though syntactically quite different, yet make use of the same lexical items and are (nearly) equivalent in meaning:

- (9) The colonists captured Ticonderoga
- (10) It was the colonists that captured Ticonderoga
- (11) It was Ticonderoga that the colonists captured
- (12) What the colonists captured was Ticonderoga
- (13) Ticonderoga was captured by the colonists
- (14) It was Ticonderoga that was captured by the colonists
- (15) What was captured by the colonists was Ticonderoga
- (16) It was the colonists that Ticonderoga was captured by

With some patience and effort we could add to and emend phrase structure rules (5a, b, c) so that sentences (10) through (16) could also be generated directly, but Chomsky has argued that to do so we must pay the price of redundancy, complexity, and inelegance (see *Syntactic Structures*, ch. 5).

Chomsky chooses instead to introduce a second kind of formal rule called a transformation. Unlike phrase structure rules, transformations do not build phrase markers from scratch, but rather convert one phrase marker into another one.

To see how this solution helps us generate sentences (10) through (16), assume first that they, like (9), start out with the structural description displayed in (4) (I shall return, in a bit, to the question of what "start out" may mean). We can next introduce three transformational rules which may be formulated roughly as follows:

- (17a) Passive
 $NP_1 + V + NP_2 \Rightarrow NP_2 + BE_f + V_{\text{past part}} + \text{by} + NP_1$
- (17b) Cleft Sentence
 $X + NP + Y \Rightarrow \text{it} + BE_f + NP + \text{that} + X + Y$
- (17c) *Pseudo-Cleft Sentence*
 $X + NP + Y \Rightarrow \text{what} + X + Y + BE_f + NP$

where RE_f represents the (appropriate) finite form of the verb be; and X and Y cover any structures to the left and right, respectively, of NP.

Let me now try to demystify these rather formidable looking rules:

The passive transformation says, in plain English: to convert an active sentence into the corresponding passive one, interchange the subject and object, add the preposition *by*, convert the verb to a past participle, and throw in the proper form of the verb be;

The cleft sentence transformation says: to focus, initially, on a noun phrase, move it to the front of the sentence, preceded by it and the proper form of be, and followed by the particle that; nothing else changes;

The pseudo-cleft transformation performs a similar function: to focus, finally, on a noun phrase, move it to the end of the sentence, preceded by the proper form of be; put what at the beginning of the sentence.

To see how these three transformations function and interact, let us return to sentences (9) through (16):

- (9) none of these transformations apply
- (10) only CLEFT applies (to the colonists)
- (11) only CLEFT applies (to Ticonderoga)
- (12) only PSEUDO-CLEFT applies (to Ticonderoga)
- (13) only PASSIVE applies
- (14) PASSIVE followed by CLEFT (*Ticonderoga*)
- (15) PASSIVE followed by PSEUDO-CLEFT (Ticonderoga)
- (16) PASSIVE followed by CLEFT (the colonists)

Observe that we could even get a sentence such as (18) where all three transformations apply, in the order: PASSIVE, CLEFT, PSEUDO-CLEFT:

- (18) What it was that was captured by the colonists was Ticonderoga.

In Chomsky's system of syntax, then, there are two kinds of rules:

- (a) Phrase Structure rules, and
- (b) Transformational rules.

The former generate the initial phrase marker, or DEEP STRUCTURE, and the latter convert this, via a series of intermediate stages, to a final phrase marker, or SURFACE STRUCTURE. The number of intermediate stages (each of which is represented by a phrase marker) is dependent upon the number of transformations which apply in any given instance. Thus if three transformations apply, the derivation involves four phrase markers (the general rule being, of course, that the number of phrase markers in a derivation is one more than the number of transformations which apply).

We can note further, with Chomsky, that (somewhat surprisingly) all transformations appear to be structure dependent. If we

examine the three transformations we have met so far (17a, b, c), we observe that reference is made not to lexical items *per se*, nor to their relative position in the sentence (first, last, third, etc.), nor to, say, their length, but rather to the syntactic category to which they belong. That category may be either lexical (such as V), or non-lexical (such as NP).

It is noteworthy, for example, that we do not form a passive sentence from an active one by simply interchanging the first two words, which would be a structurally independent operation.

Chomsky concludes that structural dependency is a universal condition on all transformations, in the grammar of any language. To the extent that transformational rules actually reflect mental operations on the part of native speakers (and as you might guess, a great deal of ink has been spilt on just this question), this condition suggests an interesting property of human language — perhaps, some would say, of the human mind itself.

The fact that transformations seem to be structure dependent suggests an interesting practical use for (some of) them: namely, to help determine constituent structure and, in some cases, even the identity of the constituent.

Consider, for example, the class of *-ing* forms commonly called “gerunds.” Rosenbaum (1967) and Emonds (1972) have pointed out that the Cleft Sentence Transformation helps us distinguish between true gerunds (i.e., those which function as NP’s) and false gerunds (what Emonds labels “reduced progressive infinitives”). To see how the test works, compare the following two sentences:

- (19a) Washington liked crossing the Delaware
- (19b) Washington began crossing the Delaware.

Superficially, these sentences look very much alike. They even behave alike in allowing an infinitive to replace the *-ing* form:

- (20a) Washington liked to cross the Delaware.
- (20b) Washington began to cross the Delaware.

But the similarity stops when we try clefting:

- (21a) It was crossing the Delaware that Washington liked
- (21b) *It was crossing the Delaware that Washington began.

(It is customary these days to use an asterisk to mark ungrammatical sentences.) In other words, application of the Cleft Sentence Transformation suggests that *crossing the Delaware* is an NP in

(19a) but not in (19b); that is, that it is a gerund phrase in the former, but not in the latter.

However, care must be taken in drawing conclusions based on such "trial by transformation." There are several pitfalls for the unwary. Here are two of them:

- (a) The test identifies constituents only at that stage in the derivation at which the rule applies. Thus, the cleft sentence test doesn't necessarily have anything to say about gerund status at earlier (or later) stages in the derivation.
- (b) The transformational rule itself may prove to be improperly formulated. It is sobering to note that there is hardly a transformation alive today that has not been to the doctor at least once.

Our own Cleft Sentence Transformation, as formulated in (17b), needs to be doctored, as the following data make clear:

- (22a) Washington crossed the Delaware *in the winter*
- (22b) It was *in the winter* that Washington crossed the Delaware
- (23a) Washington crossed the Delaware *because he wanted to surprise the British*
- (23b) It was *because he wanted to surprise the British* that Washington crossed the Delaware.

It appears that we can use clefting to focus not only on NP's, but also on prepositional phrases (22a,b) and subordinate clauses (23a,b). In other words, it seems to be a necessary but not a sufficient condition for NP-hood that a constituent can be focused by clefting.

Another, even more basic, problem is due to what we might call the TRADE-OFF RELATION. We have already observed that a transformational grammar contains two types of rules: phrase structure rules (PS-rules) and transformational rules (T-rules). Now the problem is this: how does one decide when a given constituent is to be generated by a transformational rule, when by a phrase structure rule? The problem is a very real one, since it is generally, if not always, the case that both options are available.

Recall that in analyzing sentences (9) through (16) we chose to generate (9) directly by PS-rules (strictly speaking, this is not quite true; no sentence ever surfaces without at least one T-rule applying); sentences (10) through (16) were then derived by the application of (optional) transformations. I suggested at that time that

such a solution was somehow the "simplest" and most "elegant" one, though no one knows just exactly how to measure "simplicity" and "elegance," particularly for partial grammars (Chomsky has stressed that it is the complete grammar that must be evaluated, not a grammar fragment, since complicating one rule may result in simpler rules elsewhere in the grammar, and vice versa; unfortunately, no one has yet, nor is anyone about to, come up with a complete grammar for any natural language). Still, we can, in theory if not in practice, offer one criterion for dealing with the TRADE-OFF relation:

- (a) the SIMPLICITY PRINCIPLE: evaluate most highly that combination of PS-rules and T-rules which results in the simplest grammar.

Another intriguing criterion that has been suggested is

- (b) the UNIVERSAL BASE HYPOTHESIS: all natural languages make use of the same, or highly similar, base (= PS) rules.

This hypothesis is appealing since, if true, it would not only go a long way toward solving the TRADE-OFF problem (we would still have to find that universal base), it would also help account for the fact that children the world over, regardless of social background or intellectual ability, learn their native language in an amazingly short time (amazing, at least, from the grammarian's point of view; contrast the child's feat with the fact that grammarians, working in concert for more than two thousand years, have not yet been able to construct a complete grammar for a single natural language).

Unfortunately, there is a major problem with the UNIVERSAL BASE HYPOTHESIS: it can not be falsified. In fact, it has been demonstrated (see Peters, 1970; Peters and Ritchie, 1969, 1971, 1973) that there is an infinite number of such universal bases, and even that the number of base rules can be reduced to two. This unhappy result is due to the fact that transformations, in their present form, are much too powerful. Just how transformations are to be constrained is one central area of research (see, for example, Ross, 1967; Emonds, 1970, 1972; Chomsky, 1973). One of the more general operating principles (formulated after the publication of *Syntactic Structures*), and one which has a direct bearing on the TRADE-OFF relation is

- (c) the MEANING-CONSERVING PRINCIPLE: Transformations must not alter meaning.

There has been considerable debate about the appropriateness of this hypothesis (see, for example, Partee, 1971). Some critics have even argued that no one could suggest such a hypothesis unless he were highly insensitive to the nuances of language—and such insensitivity, in a grammarian, is as disastrous as a tin ear in a phonetician. Such critics hold the not unreasonable view that there are no cases of true synonymy, i.e. that no two sentences ever have exactly the same meaning. If this be true, then a principle such as (c) would outlaw transformations altogether.

Now, of course such an argument is not quite valid since in Chomsky's theory transformations do not actually change one sentence into another sentence, but rather one phrase marker into another, or, if you like, one abstract string into another (less abstract) string. In other words, hypothesis (c) is based on the assumption that all *phrase markers in any given derivation* are synonymous. Of course, it would have to follow, if true sentence synonymy does not exist, that no one deep structure could ever give rise to more than one surface structure, or, equivalently, that there are no *optional* transformations (such as our CLEFT and PSEUDO-CLEFT), but only obligatory ones. Admittedly, this would still call into question one of the basic tenets of transformational grammar (a major attraction of TG has been its elegant accounting of alleged cases of synonymy and ambiguity). On the other hand, it is no coincidence, I think, that many transformations which once were held to be optional have since been reformulated as obligatory. This point I shall examine in some detail, after proposing a fourth principle which has played a role in the TRADE-OFF relation:

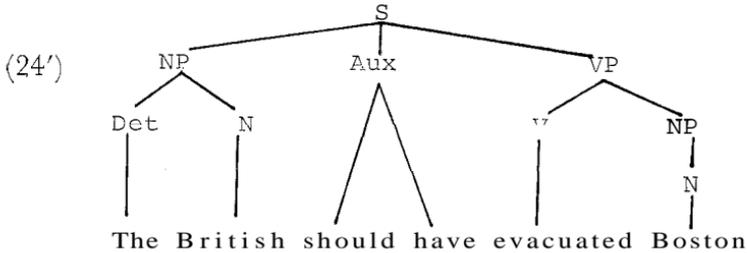
- (d) the CONTINUITY PRINCIPLE: phrase structure rules must not introduce discontinuous constituents.

Consider now such sentences as:

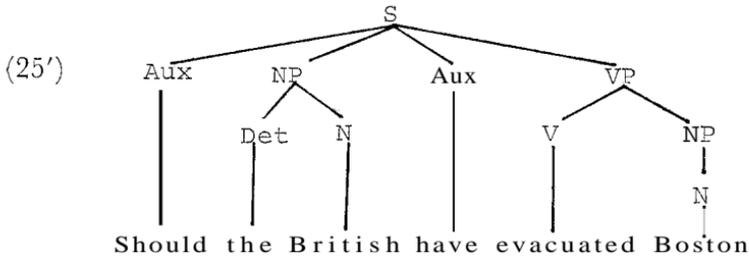
- (24) The British should *have* evacuated Boston
(25) *Should* the British *have* evacuated Boston?
(26) The British *should not have* evacuated Boston.

If we assume, with Chomsky, that helping (or auxiliary) verbs form a constituent in deep structure, then the CONTINUITY PRINCIPLE (d) dictates that (24) must, in some sense, underlie

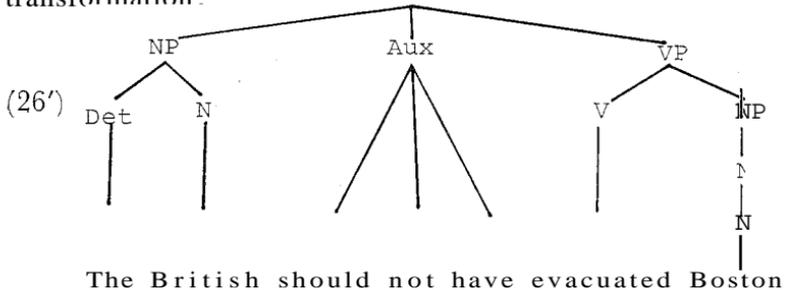
(25) and (26), since only in (24) do the auxiliary verbs should and have form a continuous constituent. We may conclude, therefore, that the deep structure for (24), at least, is something like this:



In Syntactic Structures, Chomsky used two optional transformations, QUESTION and NEGATION, to create the necessary discontinuities. Thus (25') would result from (24') by the QUESTION transformation:



and (26') is derived from (24') by means of the NEGATION transformation:



It can be seen that the QUESTION transformation creates a discontinuity by fronting the first auxiliary verb (call it Aux.), and the NEGATION transformation does so by inserting the word not after the first auxiliary. These two rules can be formulated as follows:

(27) QUESTION

$NP + Aux_1 + X \Rightarrow Aux_1 + NP + X$

(28) NEGATION

$X + Aux_1 + Y \Rightarrow X + Aux_1 + not + Y$

We have here, incidentally, an embarrassing problem, one which has pestered transformationalists since 1957: as (27) and (28) make clear, it is not the category *Aux*, but the subcategory *Aux₁* which is crucial in such operations as negation and interrogation. But while *Aux* can easily be introduced by a PS-rule such as (29):

(29) $S \rightarrow NP + Aux + VP$

it is not so easy to incorporate the variable category *Aux*. However, this is not the issue I wish to consider here. Rather, I want to point out that Chomsky's theory underwent a major change between 1957 and 1965 with the addition of the MEANING-CONSERVING principle.

It is clear that (24), (25), and (26) have much in common: the same subject, the same direct object, the same lexical items (except for *not*). But it is also clear that they do not mean the same thing. So here are two cases where transformations alter meaning. However, on the basis of work done by Lees, Klima, Katz and Postal, Chomsky in 1965 (*Aspects*, p. 132) proposed the following alteration: introduce, in the base, a special kind of "marker" to label a string as (optionally) negative and/or interrogative. As a result of this addition, the NEGATION and QUESTION transformations are converted into obligatory transformations triggered by the presence of these markers. This is a classic example of what I have called the TRADE-OFF.

A similar fate was in store for the PASSIVE transformation. Whereas it had originally been an optional rule in 1957, by 1965 (*Aspects*, pp. 103–106) it had become obligatory, dependent on the presence, in deep structure, of a special passive marker. The reformulation seemed to be necessary because of the existence of English verbs which, though they behave like ordinary transitive verbs in taking a "direct object," yet misbehave in that they fail to passivize. Examples of such verbs, which Lees has called "middle verbs," are given in sentences (30) through (32):

(30a) Washington resembles Cornwallis

(30b) *Cornwall is resembled by Washington

(31a) Washington weighed 180 pounds

(31b) *180 pounds was weighed by Washington

- (32a) The British had a tough time at Bunker Hill
 (32b) *A tough time was had by the British at Bunker Hill.

(Note, incidentally, that if we replace *tough* by *good*, and substitute all for the British, we get the surprisingly grammatical (33):

- (33) A good time was had by all at Bunker Hill

This anomaly is generally relegated to the "idiom" bag.) The problem now is this: how do we allow our grammar to generate the (a) sentences above, while blocking the ungrammatical (b) sentences?

One approach might be to adopt an innovation proposed by one of Chomsky's former students, now ardent critic, George Laliöff. We could add the rule feature [-Passive] to such lexical items as *resemble*, *weigh* and *have*. But such a "solution" not only seems *ad hoc*, it has the disadvantage of increasing the power of a grammar which is already overly powerful (as noted by Bach, 1974: 215). For by adding the notion of "rule feature" we (a) widen the concept of "lexical item," (b) allow the lexicon to participate in the TRADE-OFF relation, and (c) give to transformations the added power of peeking "inside" lexical items.

Although Lakoff's innovation may indeed prove to be a necessary addition to transformational theory (and Lakoff is not alone in believing that the proof has already been provided, in Lakoff 1970), Chomsky managed in 1965 to solve, without recourse to rule features, the problem posed by "middle verbs." His solution is based on the observation by Lees that "middle verbs," unlike true transitive verbs, do not generally collocate with manner adverbials. Contrast, for example, (34) with (35):

- (34) Washington slowly crossed the Delaware
 (35) *Washington slowly resembled Cornwallis.

Consequently, rather than introduce the rule feature [-Passive], we can subcategorize verbs in the lexicon according to whether or not they can co-occur with manner adverbials. And by adding rules of subcategorization and co-occurrence to the base, we can put more stringent controls on lexical insertion. To understand more clearly the point being made here, we must now face squarely an issue which I have hitherto skirted: just how do lexical items get put into phrase markers?

Basically, there are three competing views, which we might for convenience label as follows:

- (36a) the CONTEXT-FREE approach
- (36b) the CONTEXT-SENSITIVE approach
- (36c) the LEXICAL DECOMPOSITION approach.

The first view (36a) characterizes Chomsky's early work (*Syntactic Structures*, for example), and it still has a few adherents, though Chomsky himself has since abandoned it in favor of the second approach (36b). The third technique, that of lexical decomposition (36c), is one which Chomsky opposes, though it is held by an increasing number of his own students and colleagues.

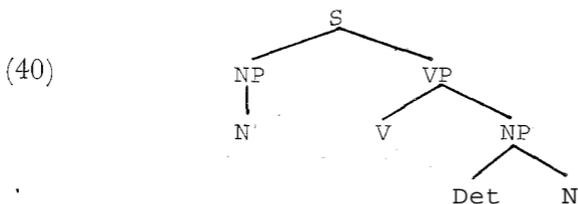
The CONTEXT-FREE approach assumes that lexical items are inserted, as unanalyzable units, by simple context-free phrase structure rules such as:

(37) Det → a, the, etc.

(38) N → Washington, Cornwallis, Delaware, fort, sword, etc.

(39) V → capture, cross, evacuate, drink, etc.

This implies that given a phrase marker such as (40):



we may apply the lexical insertion rules (37) through (39) to obtain, for example, not only:

(41a) Washington crossed the Delaware

(41b) Cornwallis captured the fort

but also such questionable strings as:

(41c) Washington drank the fort

(41d) Cornwallis evacuated the sword.

(I ignore here the question of tense, since this is not directly relevant.)

Since a grammar, according to Chomsky, must be capable of generating all the "grammatical" sentences and none of the "ungrammatical" ones, the issue here is whether or not strings like (41c) and (41d) are "grammatical." Originally Chomsky held that they were, i.e. that such strings were semantically anomalous, but syntactically well-formed (see *Syntactic Structures*, pp. 15-16).

By 1965, however, he had changed his mind, and his book *Aspects of the Theory of Syntax* describes the apparatus he feels is necessary to block strings like (41c) and (41d). Without going into details, what Chomsky's new proposal does, basically, is make lexical insertion CONTEXT-SENSITIVE. Through the use of features, he allows for a more delicate subcategorization of lexical items, and for selectional restrictions to obtain between nouns and verbs (see *Aspects*, ch. 2). For example, the verb *drink* might be entered in the lexicon as follows:

(42)	(a)	drink	
	(b)	+V	
	(c)	+ — NP	
	(d)	+[+animate] — —	
		+ — [+liquid]	
		etc.	

This matrix contains the information (a) that *drink* is a verb, (b) that it is a transitive verb (takes a direct object NP), and (c) that it requires an animate subject as well as (d) a liquid object. Since *fort* will not be marked as [+liquid], (41c) will not be generated.

Chomsky claims, furthermore, that there is a distinction between features such as (42a, b) on the one hand, and (42c, d) on the other. The former, which involve only category symbols, he calls strict subcategorization features; the latter, which incorporate labels like *animate* and *liquid*, he calls selectional restriction (i.e. co-occurrence) features. The co-occurrence of *drink* and *fort*, then, would violate a selectional restriction feature.

Not every transformationalist is convinced that selection restrictions are part of syntax. A feature like [+liquid] looks suspiciously semantic. Consider an analogous example, which I adapt from an observation of James McCawley's. Upon introspection, any native speaker of English understands that the verb *try* does not normally collocate with any inanimate subject. Yet, if George Washington had suddenly begun shouting:

(43) My toothbrush is trying to kill me
 he would not have been advised to brush up on his English grammar, but would probably have been relieved of his command instead.

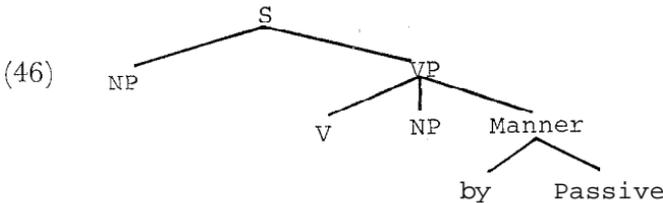
The use of features of strict subcategorization, on the other hand, seems to be less suspect, though the difference between these two types of features is not, I think, as clear-cut as Chomsky's *Aspects*

seems to imply. Since the distinction depends on what we are willing to regard as a category symbol, considerable latitude exists until firm restrictions are put on the type and number of different categories which PS-rules are permitted to introduce. Thus, the problem posed by "middle verbs" is handled by Chomsky in terms of strict subcategorization, since he feels free to add to the base two new categories, namely Manner and Passive. These can be inserted by incorporating PS-rules such as:

(44) $VP \rightarrow V + (NP) + (\text{Manner})$

(45) $\text{Manner} \rightarrow \text{by} + \text{Passive}$

which permit the generation of a phrase marker like (46):



The verb in this tree will automatically be subcategorized as (47):

(47) $\left[\begin{array}{l} +V \\ + \text{---} NP + \text{Manner} \end{array} \right]$

In other words, no verb can be inserted in this tree unless it is marked in the lexicon as permitting both a "direct object" NP and a Manner adverbial. Since a verb like *resemble* will not be so marked, it cannot be inserted in such a tree. Consequently, the verb *resemble* will never be subject to passivization, and our grammar will no longer generate the undesirable (30b) **Cornwallis is resembled by Washington*.

Note that, in Chomsky's solution to the "middle verb" problem, even though lexical insertion is now feature-dependent, the Passive transformation is not. It is still governed solely by categorial information in the phrase marker. Thus, this solution differs in principle from one that would make the Passive transformation dependent on a rule feature such as $[\pm \text{Passive}]$.

Can all transformations be so constrained that they depend only on configurations of category symbols, and never on lexical features of any kind? This question is part of a larger issue which we might label and define as in (48):

(48) TRANSFORMATIONAL PRIVILEGE: How much information, and what kind of information, ought transformations be privy to?

Since transformations are already too powerful, we must, according to Chomsky, view with skepticism any proposal which would increase that power. Consequently, the burden of proof rests with anyone who would propose to make transformations privy to information other than that provided by configurations of category symbols.

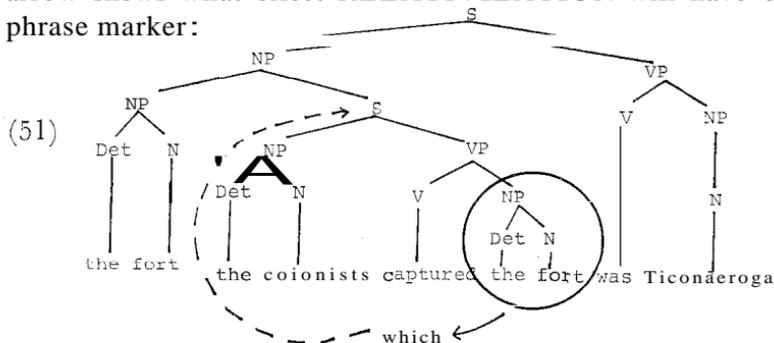
Actually, Chomsky himself has made use of transformations employing such lexical features as [&Animate] and [\pm Human], and this as early as Syntactic Structures (see p. 69; but note the escape clause in footnote 2). Consider, for example, the transformational analysis of relative clauses. A sentence like (49) has generally been assumed to be derived from an underlying string such as (50):

- (49) The fort which the colonists captured was Ticonderoga
 (50) The fort—the colonists captured the fort—was Ticonderoga.

The crucial transformation here is RELATIVIZATION, which accomplishes two things:

- (a) it replaces the second instance of the *fort* with the relative pronoun *which*, and
 (b) it moves that pronoun to the front of the embedded sentence.

Diagram (51) may be helpful in envisioning these operations. The arrow shows what effect RELATIVIZATION will have on this phrase marker:



The point to note here is that if indeed it is the RELATIVIZATION transformation which is to be responsible for inserting relative pronouns, then that transformation must be privy to the following information: is the deleted noun "human" or not? If human, then the rule puts in *who* or *whom*; if non-human, it inserts *which*. This will account not only for the grammaticality of (49) (since fort is non-human), but also for such data as:

- (52) *The fort *whom* the colonists captured was Ticonderoga
 (53) The man *who* captured Ticonderoga was Ethan Allan
 (54) *The man *which* captured Ticonderoga was Ethan Allan.

This analysis suggests that categorial information alone will not do. The RELATIVIZATION transformation must have access to the lexical feature [*&Human*].

Note further that this transformation must also be privy to information regarding the identity and coreferentiality of constituents. Relativization could not, for example, apply to a string such as (55), even though this is a string that could be generated by Chomsky's (1965) base rules:

- (55) *The fort*—the colonists captured *the town*—was Ticonderoga

In order to block RELATIVIZATION in this case (and, ultimately, to prevent a string like (55) from surfacing as an English sentence), we need to constrain the transformation. Paradoxically, this must be done by giving it access to *more* information. We can formulate the condition, broadly, as a constraint on pronominalization in general:

- (56) the IDENTITY/COREFERENTIAL CONDITION: A transformation may only pronominalize a NP if that NP is identical to and coreferential with some other NP in the sentence.

Though this is a necessary condition, it is not a sufficient one. We can add that, in general, Pronominalization is forwards and not backwards—that is, it is generally the second, not the first, of the two NP's that is pronominalized. This is illustrated by examples (57) through (GO), where the underlying strings are given in (a), forward pronominalization in (b), and (the ungrammatical) backward pronominalization in (c). Identity of subscripts indicates coreferentiality:

- (57) Relative Pronouns:
 (a) *The fort*—the colonists captured *the fort*—was Ticonderoga
 (b) *The fort which* the colonists captured was Ticonderoga
 (c) **Which* the colonists captured *the fort* was Ticonderoga
 (58) Reflexive Pronouns:
 (a) *The ragged soldiers* defended *the ragged soldiers* well
 (b) *The ragged soldiers* defended *themselves* well
 (c) **Themselves* defended *the ragged soldiers* well
 (59) Possessive Pronouns:

- (a) Washington often despaired over the condition of Washington's army
 - (b) Washington, often despaired over the condition of *his_i* army
 - (c) **He_i* often despaired over the condition of Washington's, army
- (60) Personal Pronouns:
- (a) *Cornwallis* didn't suspect that Cornwallis would be forced to surrender at Yorktown
 - (b) *Cornwallis_i* didn't suspect that he, would be forced to surrender at Yorktown
 - (c) **He_i* didn't suspect that *Cornwallis_i* would be forced to surrender at Yorktown

Although all the (c) strings here are ungrammatical, there are some cases where backward pronominalization is allowed:

- (61a) Though *the* Americans were outnumbered at Bunker Hill, the Americans fought bravely (Underlying string)
- (61b) Though the Americans_i were outnumbered at Bunker Hill, *they_i* fought bravely (Forward pronominalization)
- (61c) Though *they_i* were outnumbered at Bunker Hill, the Americans, fought bravely (Backward pronominalization)

The general constraint seems to be:

- (62) "Pronominalization cannot occur backwards except into a subordinate clause" (Bach, 1974: 274).

Ignoring possessive pronouns, whose deep structure is extremely problematical (one suggestion would derive *Washington's* army from the army which Washington has), note the rather surprising fact that the relative, reflexive, and personal pronouns cited in (57) through (61) are in complementary distribution. This distribution depends on the relationship between the two coreferential NP's. If both NP's are in the same simple sentence, we get a reflexive pronoun. If the two NP's are in different clauses, we get either a relative pronoun (in case one of the clauses is subordinate to one of the NP's), or a personal pronoun. This suggests that the RELATIVIZATION transformation may need to be reformulated so as to integrate it into a more general process of PRONOMINALIZATION.

Be that as it may, one point seems clear: in Chomsky's system, pronominalization transformations need to refer not only to categorial information, but also to notions of formal identity and core-

ference, as well as to features such as [\pm Human] for relative pronouns (not to mention the more traditional "grammatical" categories of number, case and gender, which are needed for the proper insertion of the non-relative pronouns in our examples). Since Chomsky claims that syntax is independent of semantic considerations (but see Aspects, pp. 75–79), he is forced to conclude that aspects such as these are syntactic rather than semantic.

For those who do not agree with Chomsky, those who would argue that such considerations are semantic in nature rather than syntactic, there are at least two paths open:

- (63a) *Option 1*: Refuse to extend TRANSFORMATIONAL PRIVYLEGE to include such notions as coreference and humanness, since they are "semantic";
- (63b) *Option 2*: Extend TRANSFORMATIONAL PRIVYLEGE to include them anyway, even though they may be "semantic."

Adopting the first option allows one to maintain an autonomous syntax. This approach has been taken by a few of Chomsky's more recent students (see, for example, Dougherty 1969 and Jacliendoff 1972). The option is made possible by what I have labeled the TRADE-OFF relation. Instead of treating pronominalization as a transformational process, one may allow PS-rules to insert pronouns directly. In other words, the (b) sentences in (57) through (60) can be generated directly by the base, without positing the underlying (a) strings, and we can simply throw out all pronominalization transformations. This solution, of course, imposes a heavier burden on the PS-rules and/or rules of semantic interpretation, since now it is these rules which must block the (c) strings (as well as the (a) strings). This is but another illustration of the excessive power of the present theory. Without additional constraints on the various components of a transformational grammar, too many alternative solutions are possible.

The second option which I mentioned (63b) essentially abandons Chomsky's claim that syntax is independent of semantics. This view, which has become increasingly popular among transformationalists since the late 1960's, holds that it was a mistake to try to construct (as Chomsky has done and continues to do) purely syntactic deep structures which then can be subjected to semantic interpretation. Instead, the theory ought to allow for the generation of underlying semantic structures, perhaps in the form of phrase

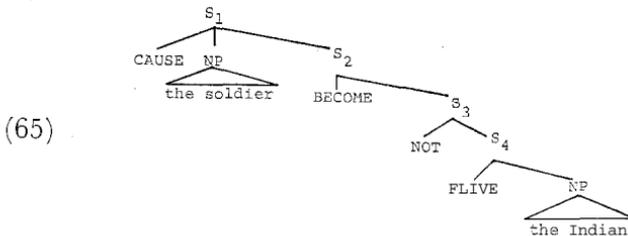
markers. Transformations may then make use of any information available in such phrase markers, be it semantic or syntactic. It has even been suggested that the distinction between syntax and semantics is an artificial one, or at least a very fuzzy one (Chomsky himself has remarked on this possibility; see *Aspects*, p. 77).

This new position is generally referred to as **GENERATIVE SEMANTICS**, as opposed to Chomsky's **INTERPRETIVE SEMANTICS**.

One of the characteristics of this new development is that **LEXICAL DECOMPOSITION** (recall (36c)) is permitted. That is, lexical items may be decomposed into alleged **SEMANTIC PRIMES**, and these primes can appear in underlying phrase markers. One of the new roles of transformations, then, is to assemble these atomic particles of meaning into lexical wholes.

To illustrate, let us consider a proposal by James McCawley (1968) that the English word *kill* can be decomposed into the four semantic primes **CAUSE-BECOME-NOT-ALIVE**. Then a sentence such as (64) might have an underlying structure such as (65):

(64) The soldier killed the Indian



Note first of all that the tree has become much more abstract than any of Chomsky's. In particular, each of the semantic primes is treated here as an underlying (logical) predicate with the result that the superficially simple sentence (64) derives from an underlying structure containing four "sentences."

If we now incorporate a new transformation, **PREDICATE RAISING**, which can optionally move any of the predicates in (65) into the S directly above it, we can create combinations of semantic primes which can then be lexicalized. Some of the options are given in (66) through (69):

No Predicate Raising

(66a) The soldier **CAUSED** the Indian to **BECOME NOT ALIVE**

(66b) =The soldier caused the Indian to become not alive

Raise One Time

- (67a) The soldier *CAUSED* the Indian to *BECOME NOT-ALIVE*
(67b) =The soldier caused the Indian to become dead

Raise Two Times

- (68a) The soldier *CAUSED* the Indian to *BECOME-NOT-ALIVE*
(68b) =The soldier caused the Indian to die

Raise Three Times

- (69a) The soldier *CAUSED-BECOME-NOT-ALIVE* the Indian
(69b) =The soldier killed the Indian

An analogous example is due to Barkai (1972). To fully appreciate his observation, a moment of briefing is required. Chomsky has suggested that most, if not all, adjectival modifiers in English derive from (reduced) relative clauses. Thus (70c) is presumed to derive from (70a) via (the FALSE STEP; see Zwicky, 1974) (70b):

- (70a) The *colonists*—*who were angry*—dumped the tea overboard
(70b) *The *colonists angry* dumped the tea overboard
(70c) The *angry colonists* dumped the tea overboard.

In other words, a transformation (called WHIZ DELETION) optionally removes the relative pronoun *who* and the copula *were* in (70a) to give the intermediate (False Step) (70b); then an obligatory rule of ADJECTIVE PREPOSING converts (70b) to (70c).

Consider now the issue raised by Barkai: the issue of preposed past participles in English. Examples (71) through (74), which rely heavily on Barkai, illustrate the problem:

- (71a) The *murdered* (drowned, stabbed, strangled, etc.) soldier was only one of many casualties
(71b) *The *killed* soldier was only one of many casualties
(72a) Washington's *polished* (scrubbed, sterilized, deoxydized, etc.) sword was saved for posterity
(72b) *Washington's *cleaned* sword was saved for posterity
(73a) The *stolen* (purloined, pilfered, appropriated, etc.) equipment was never recovered
(73b) *The *taken* equipment was never recovered
(74a) The *muttered* (murmured, shouted, whispered, etc.) oath astonished the captain

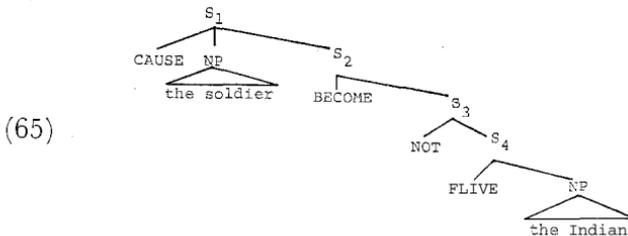
markers. Transformations may then make use of any information available in such phrase markers, be it semantic or syntactic. It has even been suggested that the distinction between syntax and semantics is an artificial one, or at least a very fuzzy one (Chomsky himself has remarked on this possibility; see *Aspects*, p. 77).

This new position is generally referred to as **GENERATIVE SEMANTICS**, as opposed to Chomsky's **INTERPRETIVE SEMANTICS**.

One of the characteristics of this new development is that **LEXICAL DECOMPOSITION** (recall (36c)) is permitted. That is, lexical items may be decomposed into alleged **SEMANTIC PRIMES**, and these primes can appear in underlying phrase markers. One of the new roles of transformations, then, is to assemble these atomic particles of meaning into lexical wholes.

To illustrate, let us consider a proposal by James McCawley (1968) that the English word *kill* can be decomposed into the four semantic primes **CAUSE-BECOME-NOT-ALIVE**. Then a sentence such as (64) might have an underlying structure such as (65):

(64) The soldier killed the Indian



Note first of all that the tree has become much more abstract than any of Chomsky's. In particular, each of the semantic primes is treated here as an underlying (logical) predicate with the result that the superficially simple sentence (64) derives from an underlying structure containing four "sentences."

If we now incorporate a new transformation, **PREDICATE RAISING**, which can optionally move any of the predicates in (65) into the S directly above it, we can create combinations of semantic primes which can then be lexicalized. Some of the options are given in (66) through (69):

No Predicate Raising

(66a) The soldier **CAUSED** the Indian to **BECOME NOT ALIVE**

(66b) =The soldier caused the Indian to become not alive

(74b) *The *said* oath astonished the captain.

What is of interest here is the fact that the (a) sentences are all grammatical, whereas the seemingly parallel (b) sentences are not.

Barkai suggests that an explanation for this discrepancy is perhaps to be found in the semantic content of the participles. More specifically, he argues that the participles in the (a) sentences are more complex, semantically, than those in the corresponding (b) sentences.

Barkai proposes that each of the participles in the (a) sentences is semantically decomposable into two parts: a prime manifested by the participles in the parallel (b) sentences, plus some sort of adverbial modifier. Thus we might analyze the "complex" verbs as follows:

(71c) murder = KILL + Adverb (with MALICIOUS INTENTION)

(72c) polish = CLEAN + Adverb (by RUBBING)

(73c) steal = TAKE + Adverb (in UNLAWFUL manner)

(74c) mutter = SAY + Adverb (in LOW and INDISTINCT manner)

Preposing of past participles could then be made dependent upon such semantic structures: if only a simple prime, without adverbial modifier, is present then preposing is blocked.

Proposals such as these by McCawley and Barkai have far-reaching consequences for transformational theory. Not only are some transformations (such as PREDICATE RAISING) now allowed to operate *before* lexical insertion is complete, others (such as WHIZ DELETION and/or ADJECTIVE PREPOSING) are made openly dependent on semantic structure. In Chomsky's system such situations are strictly outlawed.

Another innovation is that the notion of synonymy (or paraphrase) is considerably extended. Thus, for McCawley, all the (b) sentences in (66) through (69) are derived from the same underlying structure, namely (65). In other words, these sentences are all to be viewed as synonymous provided (and apparently most generative semanticists subscribe to this proviso) that transformations are not allowed to alter meaning. In Chomsky's system, on the other hand, the (b) sentences all have different deep structures, and it is up to the rules of semantic interpretation to account for the obvious similarities in meaning.

But there is a further issue which many of the old guard find

(74b) *The *said* oath astonished the captain.

What is of interest here is the fact that the (a) sentences are all grammatical, whereas the seemingly parallel (b) sentences are not.

Barkai suggests that an explanation for this discrepancy is perhaps to be found in the semantic content of the participles. More specifically, he argues that the participles in the (a) sentences are more complex, semantically, than those in the corresponding (b) sentences.

Barkai proposes that each of the participles in the (a) sentences is semantically decomposable into two parts: a prime manifested by the participles in the parallel (b) sentences, plus some sort of adverbial modifier. Thus we might analyze the "complex" verbs as follows:

(71c) murder = KILL + Adverb (with MALICIOUS INTENTION)

(72c) polish = CLEAN + Adverb (by RUBBING)

(73c) steal = TAKE + Adverb (in UNLAWFUL manner)

(74c) mutter = SAY + Adverb (in LOW and INDISTINCT manner)

Preposing of past participles could then be made dependent upon such semantic structures: if only a simple prime, without adverbial modifier, is present then preposing is blocked.

Proposals such as these by McCawley and Barkai have far-reaching consequences for transformational theory. Not only are some transformations (such as PREDICATE RAISING) now allowed to operate *before* lexical insertion is complete, others (such as WHIZ DELETION and/or ADJECTIVE PREPOSING) are made openly dependent on semantic structure. In Chomsky's system such situations are strictly outlawed.

Another innovation is that the notion of synonymy (or paraphrase) is considerably extended. Thus, for McCawley, all the (b) sentences in (66) through (69) are derived from the same underlying structure, namely (65). In other words, these sentences are all to be viewed as synonymous provided (and apparently most generative semanticists subscribe to this proviso) that transformations are not allowed to alter meaning. In Chomsky's system, on the other hand, the (b) sentences all have different deep structures, and it is up to the rules of semantic interpretation to account for the obvious similarities in meaning.

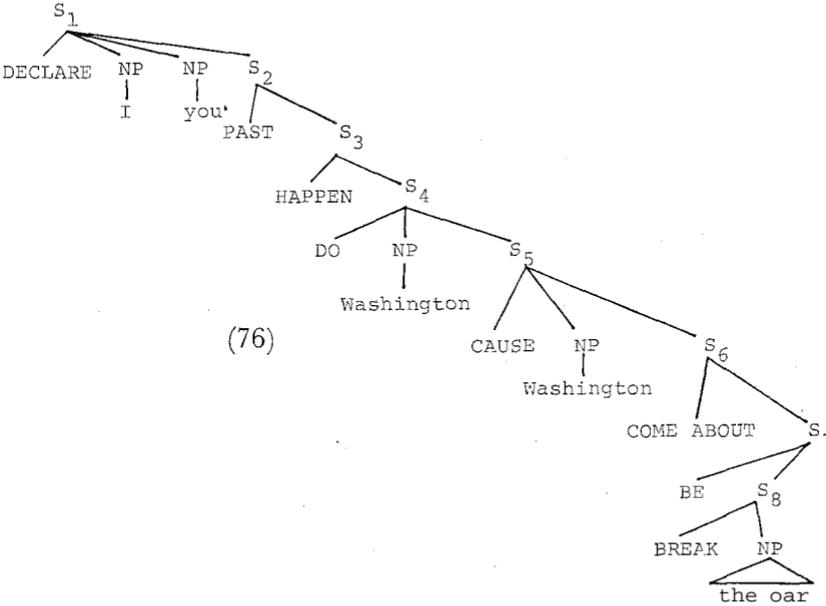
But there is a further issue which many of the old guard find

even more disquieting: generative semantics appears to increase, if that is possible, the power of a theory which is already overly rich in capability. For not only has TRANSFORMATIONAL PRIVY-LEGE been greatly extended, the admission of semantic primes has made underlying phrase markers decidedly more remote.

Unless the concept of "possible semantic prime" can be constrained in some manner, there may be no way to stop the Hydra-like proliferation of branches in underlying trees. For example, it has recently been suggested that an utterance such as

(75) Washington broke the oar

derives from an underlying structure containing at least eight clauses:



Arnold Zwicky has recently proposed (1973: 473) that we limit the Hydra's heads by imposing one or other of the following substantive constraints on semantic primes:

- (77a) "Every semantic prime is realizable as a lexical unit (root, inflection, or derivational affix) in some natural language"
- (77b) "In any language, most semantic primes are realized as lexical units"
- (77c) "In any language, all semantic primes are realized as lexical units."

Zwicky feels that (77c), the strongest of these three restrictions, is "too much to hope for"; we may have to be content with the weakest constraint, (77a).

At any rate, the upshot of Zwicky's proposal is that no semantic prime may be proposed which is not manifested morphemically in some language or other. No purely abstract primes should be permitted.

Though this sounds like a useful restriction, it is sobering to note that none of the alleged semantic primes which we have examined (CAUSE, BECOME, NOT, ALIVE, KILL, CLEAN, TAKE, SAY, DECLARE, PAST, HAPPEN, DO, COME ABOUT, BE, BREAK) violate even Zwicky's strongest condition (77c); every single one of these "primes" is realizable as an English word or morpheme.

Be that as it may, the search for semantic structures, and for these atomic units of meaning, goes on, accompanied often by expectant hope that the isolation of a universal set of semantic primes may provide the key to the structure of human language and of human thought.

Needless to say, there are skeptics, not only within the transformational camp:

"Perhaps the reason that semantic representations have proved to be so elusive is simply that, after all, there aren't any" (Fodor, Fodor, and Garrett, 1975: 530)

but also without. Ian Robinson, in his recent book *The New Grammarians' Funeral* (1975: x), writes:

"Fashions in linguistics come and go with a rapidity which in itself suggests something suspect about the essential claim of linguistics, that it is a science. Had I been writing fifteen years ago my examples would have been different, and I am pretty sure that in another fifteen years they will be out of date."

There are many, no doubt, who hope that Robinson is right, that "unless the world goes finally mad, unless the human mind in our universities loses utterly its grasp of what makes sense" (Robinson, 1975: 185), the theory of transformational grammar, in all its varieties, will be allowed to fade quietly away, so that grammarians can go back to doing whatever it was they were doing before the Revolution.

Personally, I believe such desires are premature as well as unrealistic. Whatever the future of transformational grammar itself may be, I think it is safe to predict that the study of grammar, as a result of the Chomskyan Revolution, will never again be the same.

REFERENCES

- Anderson, Stephen and Paul Kiparsly (eds.) (1973), *A Festschrift for Morris Halle*, New York: Holt, Rinehart and Winston.
- Bach, Emmon (1974), *Syntactic Theory*, New York: Holt, Rinehart and Winston.
- Barlrai, Malachi (1972), "On the Shiftability of Past Participles in English," *Linguistic Inquiry*, vol. 3, no. 3, pp. 377-378.
- Chomsky, Noam (1957), *Syntactic Structures*, The Hague: Mouton.
- Chomsky, Noam (1965), *Aspects of the Theory of Syntax*, Cambridge, Mass.: M.I.T. Press.
- Chomsky, Noam (1973), "Conditions on Transformations," in Anderson and Kiparsky, pp. 232-285.
- Dougherty, Ray (1969), "An Interpretive Theory of Pronominal Reference," *Foundations of Language*: vol. 5, pp. 488-519.
- Emonds, Joseph (1970), *Root and Structure-preserving Transformations*, Ph.D. dissertation, M.I.T.
- Emonds, Joseph (1972), "A Reformulation of Certain Syntactic Transformations," in Peters, 1972, pp. 21-62.
- Fillmore, Charles and Terence Langendoen (eds.) (1971), *Studies in Linguistic Semantics*, New York: Holt, Rinehart and Winston.
- Fodor, J. D., J. A. Fodor, and M. F. Garrett (1975), "The Psychological Unreality of Semantic Representations," *Linguistic Inquiry*, vol. 6, no. 4, pp. 515-531.
- Jackendoff, Ray (1972), *Semantic Interpretation in Generative Grammar*, Cambridge, Mass: M.I.T. Press.
- Lakoff, George (1970), *Irregularity in Syntax*, New York: Holt, Rinehart and Winston.
- McCawley, James (1968), "Lexical Insertion in a Transformational Grammar without Deep Structure," *Papers from the 4th Regional Meeting of the Chicago Linguistic Society*, pp. 71-80.
- Partee, Barbara (1971), "On the Requirement that Transformations Preserve Meaning," in Fillmore and Langendoen, pp. 1-21.
- Peters, Stanley (1970), "Why There Are Many 'Universal' Bases," *Papers in Linguistics*, vol. 2, no. I, pp. 27-43.
- Peters, Stanley (ed.) (1972), *Goals of Linguistic Theory*, Englewood Cliffs, New Jersey: Prentice-Hall.
- Peters, Stanley and Robert Ritchie (1969), "A Note on the Universal Base Hypothesis," *Journal of Linguistics*, vol. 5, pp. 150-2.
- Peters, Stanley and Robert Ritchie (1971), "On Restricting the Base Component of Transformational Grammars," *Information and Control*, vol. 18, pp. 483-51.
- Peters, Stanley and Robert Ritchie (1973), "On the Generative Power of Transformational Grammars," *Information Sciences*, vol. 6, pp. 49-83.
- Robinson, Ian (1975) *The New Grammarians' Funeral: A Critique of Noam Chomsky's Linguistics*, Cambridge: Cambridge University Press.

- Rosenbaum, Peter (1967), *The Grammar of English Predicate Complement Constructions*, Cambridge, Mass.: M.I.T. Press.
- Ross, John Robert (1967), *Constraints on Variables in Syntax*, Ph.D. dissertation, M.I.T.
- Zwicky, Arnold (1973), "Linguistics as Chemistry: the Substance Theory of Semantic Primes," in Anderson and Kiparsky, pp. 467-485.
- Zwicky, Arnold (1974), "Taking a False Step," *Language*, vol. 50, pp. 215-224.