

ON PERNICIOUS RECURSION¹

PETER H. FRIES

University of Wisconsin

In the introduction to Robert Longacre's *Grammar discovery procedures* a generative interpretation of tagmemic formulae is developed by which an infinite number of constructions (at any level) may be produced along with a tree describing the various constituents of each construction and the grammatical functions each constituent manifests within that construction. In the light of this generative interpretation of tagmemic formulae it would seem useful to explore possible systems of tagmemic rules with a view to eliminating those systems which do not properly describe the constructions they generate.

A proper description of the constructions generated by tagmemic grammar must include the specification of the immediate constituents of those constructions, and since transformational grammarians have discussed extensively the proper form of immediate constituent (or phrase structure) grammars, it seems only natural to begin by examining some of their points to see if they might be applicable to the phrase structure aspect of tagmemic grammars. Some points are clearly not applicable, since they refer specifically to phrase structure grammars which are an integral part of transformational grammar. Such phrase structure grammars emphasize binary cuts and rely on transformational rules for permutations and readjustments of the forms of morphemes and sequences of morphemes. These phrase structure grammars avoid the overt specification of grammatical function, and as a direct result of this, no rule in the phrase structure component of a transformational grammar may permute elements. Tagmemic rules, on the other hand, do specify functional relations overtly and as a result may be allowed to permute elements. In this paper I would like to examine some of the properties of rules

¹ I would like to thank Kenneth Pike and Bruce Stark for their criticisms of an earlier version of this paper. Though they did not always agree with my argument, their comments were most helpful.

which I believe do hold for any grammar which describes constituent structure; in particular, I would like to examine certain properties of recursive systems of rules².

In the original model of transformational grammar no recursive rules or systems of rules were allowed. Postal (1964: 10 - 13) states this clearly when he presents the following restrictions on phrase structure rules.

"condition (2) if $U \rightarrow W$. Then:

- a. $U = XAY$ and $W = XZY$
- b. Z is not null...
- c. Z is not identical to A " (10)

After presenting the derivation

"(v) S, AB, AB, ACD, \dots " (13)

Postal then continues

"In (v) the procedure of P marker construction yields a tree but this would be identical with that for the quite different derivation containing no repetition. Hence if the claim that a labelled tree represents the set of strings in an equivalence class of derivations is to be maintained, expansions of A into A cannot be allowed because these do not affect the set of strings in the full class." (13)

He then adds in footnote 21:

"Furthermore, if condition (2)c is not met there will be an unbounded number of lines possible in the derivation of any string and an infinite number of derivations for any terminal string. This would make the languages enumerated not recursive." (84)

Postal clearly intends to prevent rules such as 4 in the following set³.

1. $S \rightarrow NP + VP$
2. $VP \rightarrow Aux + be + Pred$
3. $Pred \rightarrow Adj$

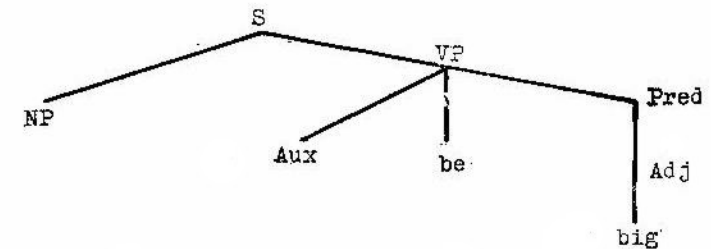
² Note that my definition of recursion is not the same as Longacre's (1965). For me, any system of rules in which a unit occurs to the left of a rewrite symbol and later to the right of a rewrite symbol in such a way that the system of rules generates an infinite number of sequences, and also infinitely long sequences is a recursive system of rules. For Longacre, recursion is more intimately tied to the notion of hierarchy. A system of rules which rewrites a sentence as part of a sentence, a clause as part of a clause, etc. is a recursive system. This concept of recursion differs from mine in that 1. a recursive system in Longacre's sense is not necessarily infinite. (Usually such cases do involve infinite systems, however). 2. Longacre would not say that a clause playing a role within a phrase constitutes recursion. He would call it "Back Looping". My use of the term makes no distinction as to the level of relevance of the included constituent.

³ These rules are excerpted from Chomsky (1962). While at no time were the rules in that grammar presented as a definite proposal for the description of English, the fact that he uses rule 4 illustrates that the temptation to use such rules is real.

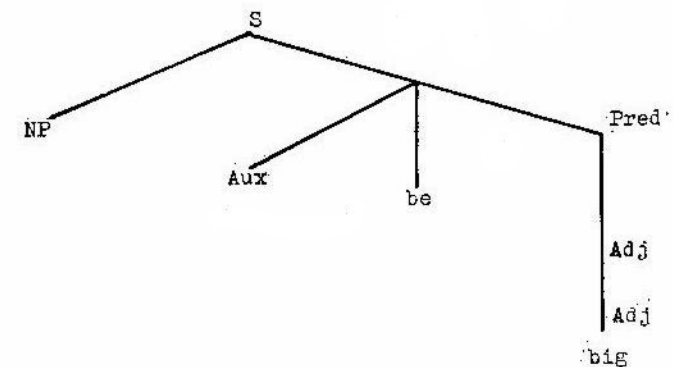
4. $Adj \rightarrow (\text{very}) Adj$

5. $Adj \rightarrow \text{big, old, } \dots$

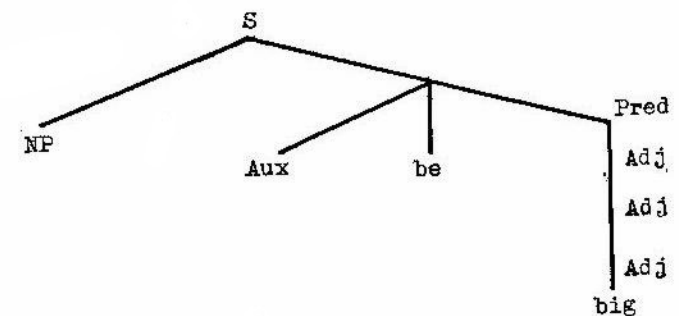
These rules generate the desired sequences, but the trees for any given sequence produced by such rules may be infinite, for rule 4 may be applied an indefinite number of times for any given sequence of words. That is, to produce the sentence *He is big one* may apply rules 1 - 3 and 5, and obtain the following tree:



or rules 1 - 5 may be applied, resulting in



Similarly rules 1 - 3, 4, and 5:



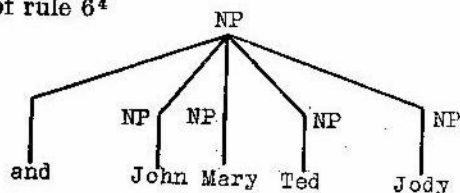
etc. Rule 4 may be applied any number of times (from none to infinity) during the derivation of this sequence, therefore the trees which are produced using this rule are indeterminate. The indeterminacy of trees involving this rule is, in addition, of an uninteresting sort, since it is not relatable to any linguistic ambiguity. That is, we cannot say that the sentence *He is big* is ambiguous in a way that may be explained by varying number of nodes labelled *adjective* which dominate *big*. We must conclude that the indeterminacy of the tree is a result solely of rule 4. Since one of the major purposes of a phrase structure grammar is to assign a tree structure to each sequence producible by the grammar, rules such as 4 are not allowed within the grammar.

Later revisions of transformational grammar loosen Postal's restrictions considerably. Thus in 1966 we find Lakoff and Peters (1969: 114) using rule schemata such as

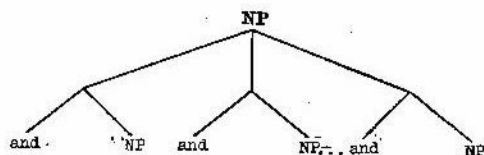
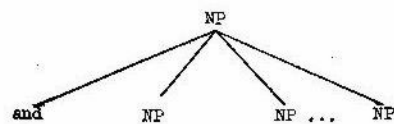
6. $NP \rightarrow \text{and } (NP)^n$. $n \geq 2$

Such rule schemata clearly violate Postal's restrictions in that they introduce NP as a constituent of NP. They do not, however, produce indeterminate trees as rule 4 did above. That is, each different application of rule 5 will correspond to a difference in the tree which, if it does not actually result in different sequences of morphemes, does correlate with actual ambiguities of the sequence. For example, the sequence *John and Mary and Ted and Jody* may be produced using

a) one application of rule 6⁴



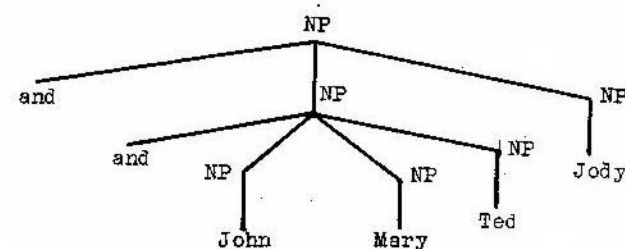
⁴ The form of this rule assumes a general principle which transforms a tree of the form



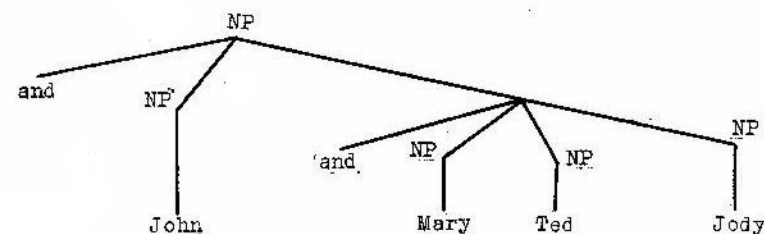
Later transformational rules may delete certain of the conjunctions (see Lakoff and Peters, 1966 fn. 2). These readjustments and deletions do not affect the recursiveness of the rule and therefore will be ignored.

b) two applications of rule 6

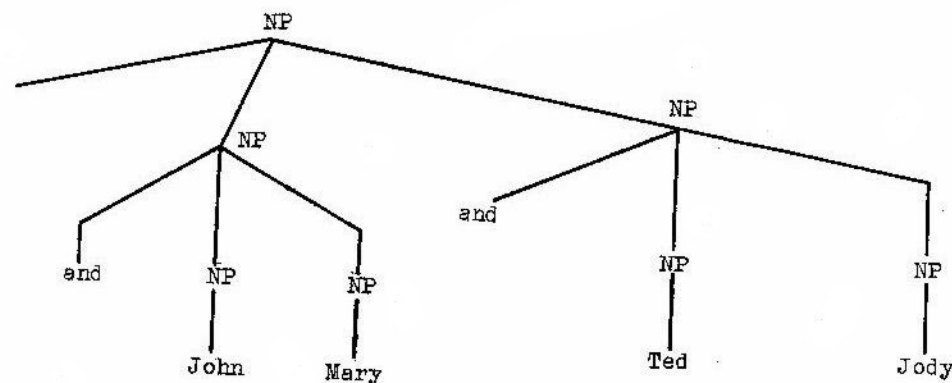
i)



or ii)



or c) three applications of rule 6



These four trees correspond to four different possible semantic interpretations of the sequence and are therefore necessary to the proper description of that sequence. The status of rule 6 therefore differs significantly from that of rule 4 since the various trees describing any given sequence of morphemes derivable by the applications of rule 4 do not correspond to any ambiguity

of the resulting sequence of morphemes. Both rule 4 and rule 6 are recursive rules, but since rule 4 produces vacuously indeterminate trees for all sequences it generates, let us call it (and any rules or systems of rules like it) *perniciously recursive*.

What characterizes a set of perniciously recursive rules? We have already seen one instance in rule 4. In this case pernicious recursion occurs because an Adjective may in some cases be the sole manifestation of an including Adjective. In fact, pernicious recursion results every time a system of rules allows a construction to be realized entirely by another identical construction. That is to say, a system of rules such as the following is perniciously recursive.

7. $A \rightarrow (a) b$

8. $b \rightarrow (c) d$

9. $d \rightarrow (e) A$

For A may dominate solely A. No matter what other alternative expansions there are for A, it is possible to apply and reapply rules 7 - 9 any number of times and still result with the same terminal string.

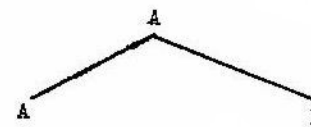
(α)	A		b		d		A
	b						
			d				
	d						
	A						
			b				
			d				
			A				

The terminal strings of α and β are identical, yet α results from applying each of the rules 7 - 9 once, while β results from applying each of the rules 7 - 9 twice.

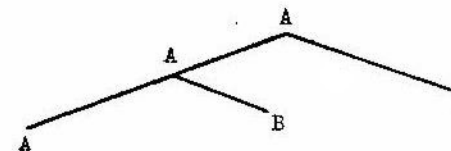
A system of rules is not perniciously recursive as long as a construction may never dominate solely itself. A rule such as 10 is not perniciously recursive, since the strings produced by applying it n times is not identical to the string produced by applying it $n+1$ times.

10. $A \rightarrow (A)+B$

Applying rule 10 once results in the tree:



with the resulting string AB. Applying rule 10 twice results in the tree:



with the resulting string ABB. The situation described by rule 10 is exemplified by the recursion involved in the description of English noun phrases like *the boy's father* where the noun phrase *the boy's father* includes within it another noun phrase, *the boy*, but this included noun phrase can never be the sole manifestation of the including noun phrase since in addition to the determiner tagmeme noun phrase must contain a Head Noun. Note that so long as B is obligatory in rule 10, the rule is not perniciously recursive no matter whether or not A is optional. That is, neither rule 10 nor rule 10a is perniciously recursive

10a. $A \rightarrow A+B$

The bad effects of pernicious recursion might be avoided by the use of an ad hoc rule stating that whenever a node labelled A dominates one and only one node also labelled A (the two nodes labelled A may be separated by any number of other nonbranching nodes) then the dominating occurrence of A and all intervening nodes are erased, and the lower node A is attached to the tree where the upper node A used to be. On the assumption that a grammar which introduces nodes only to erase them at some later date is not maximally simple, transformational grammarians whenever possible avoid the use of this convention. That is to say, they restrict the systems of rules they use to nonperniciously recursive systems.

Should tagmemists also avoid perniciously recursive systems of rules? I believe they should, but there exist certain types of data which may only be described by means of such rules.

We should note first that the overt specification of functional relations does not affect pernicious recursion. That is, rules like rule 4, or rules 7, 8 and 9, which allow a given construction to dominate solely itself, will be perniciously recursive no matter what functional relations the various constituents manifest. Thus even if we specify both function and form instead of only form in rule 4, it would still remain a perniciously recursive rule.

4a. $AP \rightarrow \pm$ Intensifier: very+Head: AP/adjective

of these constructions entails the setting up of a construction type, nominal phrase, which may be a constituent of a noun phrase. Thus the noun phrase formula (with only a partial list of potential filler classes) is:

Noun Phrase = \pm Limiter <only> \pm Determiner 1 <all>

\pm Determiner 2: article/genitive phrase \pm Determiner 3: cardinal numerals

\pm Loose Knit Modifier: Adjective \pm Close Knit Modifier: Nominal phrase+

\pm Head: noun \pm Restrictive Modifier: relative clause \pm Non-Restrictive

Modifier: relative clause

A first approximation to the nominal phrase formula is:

Nominal Phrase = \pm Determiner 3: cardinal numeral \pm Loose Knit Modifier:

Adjective \pm Close Knit Modifier: Nominal Phrase + Head: noun

Nominal phrases differ from noun Phrases in that they may not have Determiner 1 or Determiner 2 tagmemes nor the Restrictive or the Non-Restrictive Modifier tagmemes⁷, so that none of the following sequences are grammatical.

* *a big the apple corer handle*

* *an old my record shelf*

* *an old all men's quartet*

* *a University that was recently founded faculty*

The two formulas just given account for constructions such as:

the men's shoes department manager

the tree surgeons' association president

an apple corer handle marker

The following paraphrases⁸ show that these are all left branching constructions (some of the sequences marked with * may be grammatical, but they are never paraphrases of the original construction)

the manager of the men's shoes department

⁷ Recent discussions have thrown doubt on whether the nominal phrase is to be considered truly distinct from the noun phrase, since it contains solely a sub-set of the tagmemes which are contained in the noun phrase. Thus, given a sequence such as *radar installation*, one cannot tell whether it is a noun phrase or a nominal phrase unless he can find it in contexts such as *the cost of radar installation contributes 10 per cent of the price* or *a radar installation supervisor*. If the two constructions are to be considered variants of the same construction type the nominal phrase will merely be considered a restricted noun phrase, (restricted both in the possible expansions it may accept, and in its distribution). The decision as to the status of the nominal phrase will not, however, affect the basic argument about recursion which is presented here, though it may necessitate some rephrasing of that argument.

⁸ These paraphrases are intended only as non-formal support for the immediate constituent cuts made in my analysis for these constructions; they play no role in their formal description. For this reason I have allowed myself considerable latitude in supplying appropriate words (mostly prepositions) where useful.

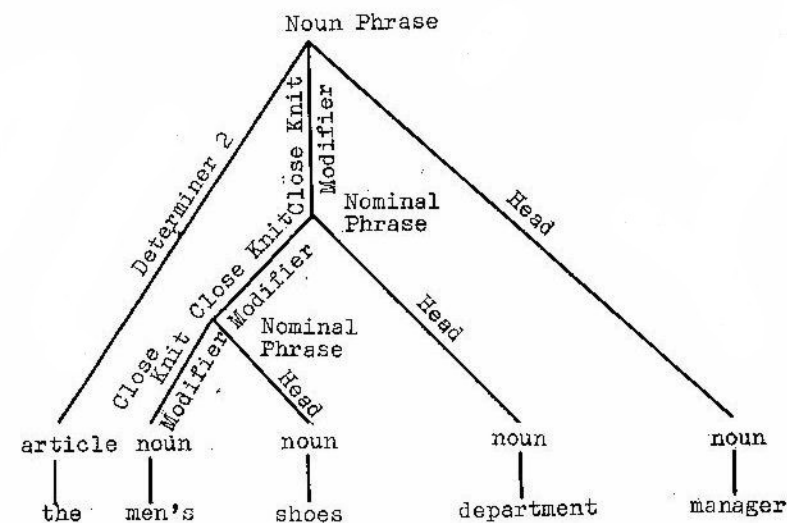
* *the department manager of the men's shoes*

the manager of the department { *of* / *for* } *men's shoes*

* *the manager of the shoes department of men's*

the manager of the department { *with* / *selling* } *shoes for men*

The constituent structure tree produced by the formulas for this example would be:



Once it is made clear that the Loose Knit Modifier and Close Knit Modifier tagmemes of the noun phrase and the nominal phrase are repeatable, a number of other constructions may be described by these formulae. e.g.

state tree surgeons association president

a walnut long playing record cabinet sale

an oak apple corer handle marker

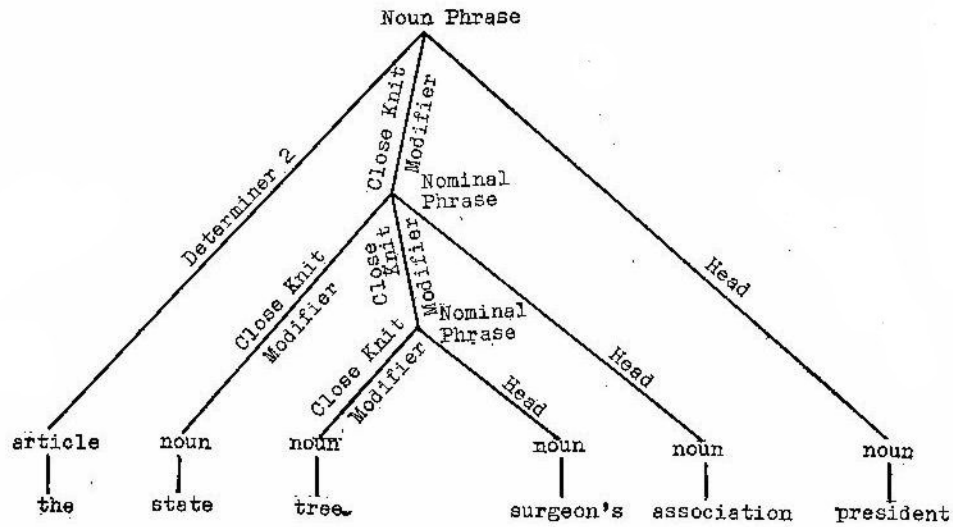
The grammaticality of the two phrases marked with braces in the following set of phrases shows that the noun phrases above contain a nominal phrase with two occurrences of the Close Knit Modifier tagmeme within it.

the president of the state tree surgeons association

the association president of the state tree surgeons

{ *the president of the state association of tree surgeons* / *the president of the tree surgeons association of the state* }

The constituent structure produced by the formulas for this example would be:



There are a few examples, however, which cannot be described by the rules as currently stated. These examples all involve right branching; that is, in each case the noun phrase contains a nominal phrase which itself contains as its head a nominal phrase (not a noun).

the school admissions policy committee

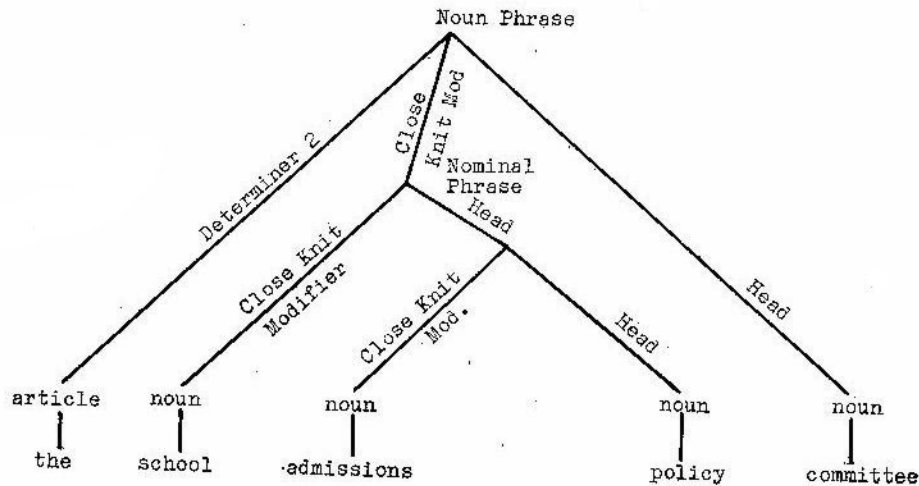
the committee for (forming) the school admissions policy

* *the policy committee for school admissions*

* *the committee for (forming) the policy for school admissions*

the committee for (forming) the admissions policy for the school

The tree for this phrase must be as follows:



Because the nominal phrase *school admissions policy* in this construction contains as its head the nominal phrase *admissions policy*, the formula given above for the nominal phrase will have to be revised to read⁹:

Nominal Phrase = ±Determiner 3: cardinal numeral
 ±Loose Knit Modifier: Adjective ±Close Knit Modifier:
 nominal phrase/noun+Head: nominal phrase/noun

But this formula is perniciously recursive, since all tagmemes other than the head are optional; thus a nominal phrase node may directly dominate a nominal phrase node with no other branching coming from the dominating node. In this case, however, we encounter an added difficulty; the two alternate fillers of the head tagmeme are emically the same construction. That is to say, a minimal nominal phrase is a noun. The phrases *a record cabinet marker* and *a cabinet marker* differ only in that the first contains a slightly expanded nominal phrase (*record cabinet*) while the second contains a minimal nominal phrase (*cabinet*). Since only emically contrasting construction types may be listed as alternative fillers of a tagmeme, and since a minimal nominal phrase is a noun, the expanded and minimal forms of the nominal phrase should be treated as emically the same: both should be called nominal phrases. (This is analogous to saying that both *old* and *very old* are adjective phrases). This means that the formula for the nominal phrase ought to be:

⁹ Additional support for this analysis comes from the fact that we often find sequences of noun-adjective-noun, e.g.

Bendix automated radar installation

the oxygen partial pressure sensor

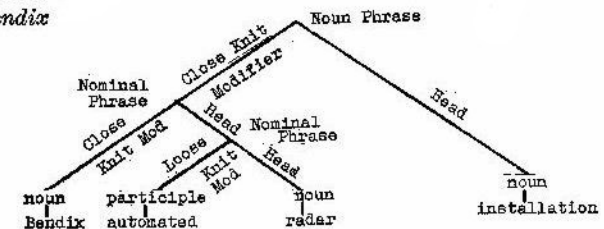
The immediate constituent structure of these constructions can be demonstrated by the use of paraphrases:

an installation with Bendix automated radar

a radar installation with Bendix automated

radar made by Bendix automated

automated radar made by Bendix



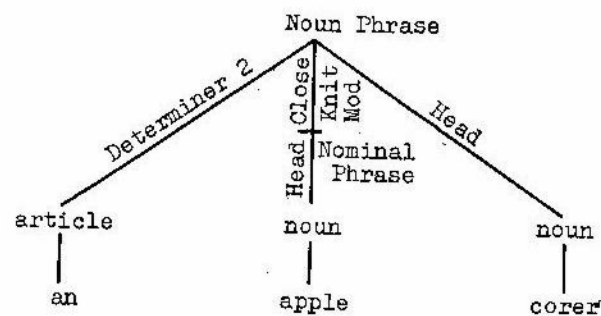
Normally adjectives and participles occur only before nouns since the Loose Knit Modifier tagmeme occurs before the Close Knit Modifier tagmeme. But if a nominal phrase may be a head of a nominal phrase, the included nominal phrase may contain a Loose Knit or Close Knit Modifier tagmeme of the including nominal phrase.

Nominal Phrase = \pm Determiner 3: cardinal numeral

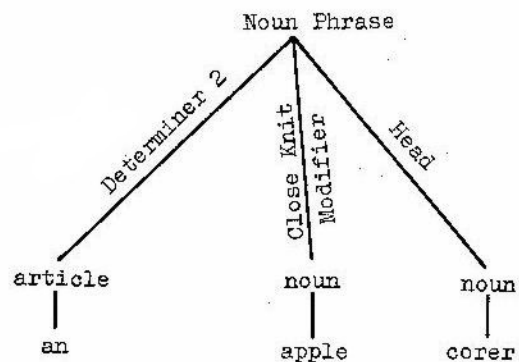
\pm Loose Knit Modifier: Adjective \pm Close Knit Modifier:
nominal phrase + Head: nominal phrase

But this introduces pernicious recursion with a vengeance, for there is now no way whatever to exit from this rule. Since the head tagmeme of the nominal phrase is obligatory and filled *only* by nominal phrase, and the head of that nominal phrase is likewise filled only by a nominal phrase, etc., the tree of any construction containing a nominal phrase is not merely indeterminate, but *necessarily* infinite. The convention proposed earlier (p. 42), the deletion of nodes introduced through pernicious recursion, is little help here, since there is no way of exiting from the circle even if we can limit the number of nodes in the tree. The question is no longer, can we describe the English nominal phrase without using systems which involve perniciously recursive rules (thus making necessary the use of the node deleting convention) but rather can these nominal phrases be described at all? Let us rephrase this in more concrete terms. Given the noun phrase, *an apple corer*, what is the proper structure to be assigned to it? Should it be a) or b)?

a)



b)



The implications of solution (a) are: 1. the grammar contains a circular rule which obligatorily produces infinite trees 2. it is claimed that a minimal nominal phrase is a noun; thus the difference between *a record cabinet* and *a twelve inch record cabinet* does not lie in the type of syntagmeme which manifests the Close Knit Modifier tagmeme of the noun phrase but in the particular variant of the syntagmeme chosen 3. nominal phrases may be realized by single nouns. Thus no artificial restrictions as to minimal length of nominal phrases need be imposed (compare 3 under solution b and 4) nominal phrases may occur as potential fillers of the Close Knit Modifier or head tagmemes of the noun phrase or nominal phrase. (Nouns may not so occur: all single nouns occurring in these positions will be analyzed as minimal nominal phrases).

The implications of (b) are 1. we avoid the problem of a circular rule producing necessarily infinite trees 2. noun contrasts with nominal phrase as a possible filler of the Close Knit Modifier tagmeme, thus the fillers of the Close Knit Modifier tagmeme of the following noun phrases are realizations of different syntagmemes.

Determiner 2	Close Knit Modifier	Head
a	record (noun)	cabinet
a	twelve inch record (nominal phrase)	cabinet

3. nominal phrases must have more than one tagmeme present. That is at least one of the optional tagmemes must be present, or, if none is, the head must be filled by a nominal phrase which contains at least one optional tagmeme. (Otherwise the result will be a single noun, which we have said belongs to a contrasting syntagmeme) and finally 4. because we find phrases such as *the university chemistry club*, *the school admissions policy* (in which the constructions *chemistry club*, and *admissions policy* are modified by *the* and *university* and *the* and *school* respectively) one possible filler of the head tagmeme of the noun phrase must be a nominal phrase. In this position it again contrasts with nouns.

We are thus left with a dilemma: Neither solution (a) nor solution (b) is wholly satisfactory. Solution (b) seems to avoid theoretical inconsistency but provides a counter-intuitive analysis of points (2) and (3). Solution (a) is untenable because of its circularity, yet it provides intuitively satisfying answers to points (2) and (3).

To summarize, we began by exploring the implications of the properties of certain recursive rules and found a. that no completely coherent system of rules could allow perniciously recursive rules, and b. that certain constructions of English can only be described using perniciously recursive rules.

REFERENCES

- Chomsky, N. 1962. "A transformational approach to syntax" in A. A. Hill (ed.) *Proceedings of the third Texas conference on problems of linguistic analysis in English, 1958*. Austin: The University of Texas, 124 - 158. Reprinted in Jerry A. Fodor and Jerrold J. Katz (eds.), *The structure of language: readings in the philosophy of language*. Englewood Cliffs: Prentice-Hall, 1964, 211 - 245.
- Lakoff, G. and Stanley P. 1966. "Phrasal conjunction and symmetric predicates". *Mathematical linguistics and automatic language translation*. Harvard Computation laboratory, Report No. NSF-17. VI-1 to VI-49. Reprinted in David A. Reibel and Sanford A. Schane (eds.) 1969. *Modern studies in English*. Englewood Cliffs: Prentice-Hall, 113 - 142.
- Longacre, R. E. 1965. "Some fundamental insights of tagmemics". *Language* 41. 65 - 76.
- Postal, P. 1964. *Constituent structure: A study of contemporary models of syntactic description*. (Indiana University Research Center in Anthropology, Folklore, and Linguistics, Publication No. 30.) Bloomington: Indiana University Press.
- Roey, J. van. 1964. "A note on noun and noun combinations in Modern English". *English Studies* 45. 48 - 52.