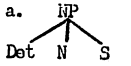
Restrictive Relatives and Surface Constituent Structure

James D. McCawley
University of Chicago

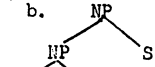
Restrictive Relatives and Surface Constituent Structure

James D. McCawley
University of Chicago

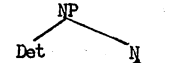
In the existing literature on restrictive relative clauses, literally all of the logically possible surface groupings of 'determiner', noun, and relative clause have been seriously advocated:

(1) a. 

(Smith 1964)

b. 

(Ross 1967, Thompson 1971)

c. 

(Stockwell et al 1972, Partee 1975, McCawley 1977a)

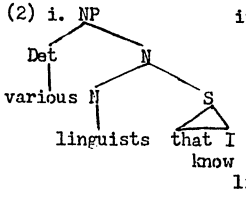
In a recently completed paper (McCawley 1977a), I argue for (1c) against the more popular (1a) and (1b) as the correct surface grouping. Later in this paper I will sketch the analysis that I argued for there, in which restrictive relative clauses are derived from an underlying structure structure involving coordination (though not, NB, the well-known coordinate source that is proposed in Thompson 1971, which I reject for reasons that I will touch on below); I show there that the constituent structure in (1c) arises naturally from a plausible semantic structure, in the sense that (1c) is what you get from the most simple-minded rules that yield the surface word order and are consistent with certain general assumptions such as that transformations apply cyclically, that their operands are constituents, and that Chomsky-adjunction is the unmarked kind of adjunction.

The structures given in (1) constitute all of the logically possible groupings of the elements (or at least, all the logically possible groupings that involve only continuous constituents), but they obviously do not include all the logically possible labelings of the nodes, nor all the possibilities for extra non-branching nodes. In the absence of a theory of what is a possible syntactic category, it of course makes no sense to speak of 'all logically possible labelings' of the nodes. It may be helpful if I clarify here my general position on syntactic categories, so as to give the reader some perspective on what should be taken seriously about various node labels that appear below and how those labels are related to the analyses sketched here. I take the position, which I elaborate on and argue for in McCawley 1977b, that there really are no such things as syntactic categories, in the sense that when linguists ostensibly have been talking about syntactic

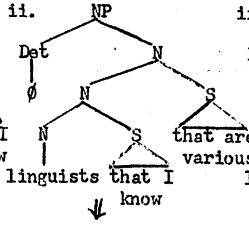
categories they really have been talking about other things (or perhaps, that the only way to make sense out of what they are saying is to interpret them as having been talking about other things). These 'other things' include the logical category of the corresponding constituent of logical structure, the morphological category of the item in question or of its head, dependency relations (such as 'head' and 'modifier'), and other topological relations (e.g. where the item is relative to the head of the construction). The node labels that appear below are to be interpreted as merely informal abbreviations for combinations of information of these various types. In particular, the occurrence of a particular node label should not be taken as implying that there is a corresponding node of 'deep structure' having the same label: since not all information of the types that are reflected in the informal node labels will be present at the deepest levels of derivations (in particular, morphological categories will not be distinguished at that level), derivational steps (such as lexical insertion) can set up the basis for differences in these informal labels.[1]

One cautionary point should be made at the outset about the structures that appear in (1), namely that what any proposal embodying one of those structures will imply about a particular NP will depend on how one applies the informal category labels given there. For example, the claim embodied in (1c) has quite different implications about the surface constituent structure of various linguists that I know, depending on whether one treats various as a determiner or as a reduced relative clause, e.g.[2]

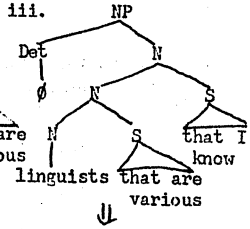
(2) i.

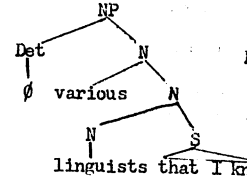


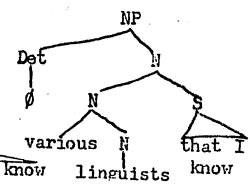
ii.



iii.







Note that in (iii) the grouping is like that in (1b), even though it strictly speaking conforms to (1c). It will in fact turn out that some of the conclusions of McCawley 1977a are questionable precisely because I uncritically assumed that all of the items that have traditionally been referred to as 'determiners' fit in the same way into the constituent structure of restrictive relative clauses.

Leaving aside the question of whether the item preceding the N + Relative combination is appropriately called a 'determiner', a constituent structure like (1c) can be justified on the basis of facts about the interpretation of anaphoric devices such as one and the zero noun:

(3) a. I want the book that I lent you that contains two stories by Maugham, not the one that has only 19th century stuff in it.
b. Tom has two cats that once belonged to George and Steve has three Ø.
c. Many linguists that work in syntax are also interested in phonology, but few Ø are interested in onomastics.

In each case, the preferred interpretation for one or the zero pronoun is the combination of noun and relative clause (book that I lent you, etc.), which argues that such a constituent is available as antecedent at the stage of derivations to which the anaphora rules apply, whether those rules are deletion transformations, as I will henceforth assume, or are something else. The constituent consisting of noun and restrictive relative clause can also be conjoined, as in

(4) { a. Several / b. Nearly a dozen } linguists that play chess and philosophers that play bridge were present.

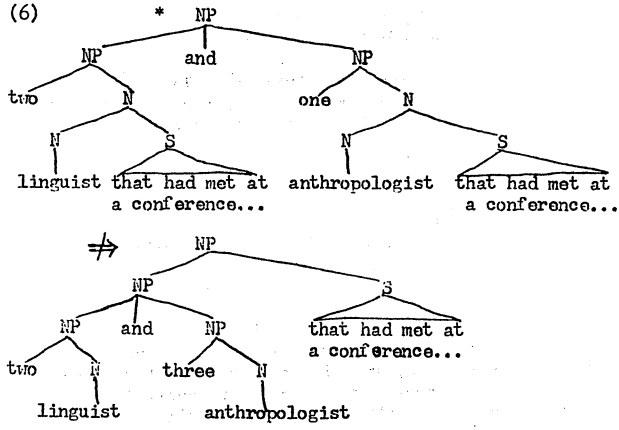
(4a) allows an interpretation in which there is only one semantic occurrence of several: where it does not say that there were several linguists and several philosophers but rather that the linguists who play chess and the philosophers who play bridge jointly were 'several', e.g. it would cover the case where there were three linguists who play chess and three philosophers who play bridge; similarly with (4b).

Consider, however, sentences like (5), in which it appears that constituents consisting of determiner and noun are conjoined, with a single relative clause shared by the conjoined items:

(5) Two linguists and one anthropologist that had met at a conference on language planning were among those arrested.

The relative clause here is chosen so as to rule out any possibility of a derivation from a (1c) structure via Right-node-raising:

(6)



While Right-node-raising would in fact derive structures like (1b) from (1c) structures, such a derivation would make no sense here, since the input structure would be both semantically anomalous and semantically irrelevant, i.e. one anthropologist that had met at a conference on language planning is semantically anomalous, and the meaning of (5) has to do with the whole group of persons meeting at a conference on language planning, not with each of two subgroups meeting there. In addition, the combination of noun and relative clause is not available as antecedent for an anaphoric device in a NP with the sort of relative clause that figures in (5):

(7) ??Mike interviewed three linguists that had met at a conference on language planning and Barbara interviewed two ∅.

Thus, there appears to be exactly as good a case for saying that two linguists that had met at a conference on language planning has a (1b) surface structure as there is for saying that many linguists who work in syntax has a (1c) surface structure.

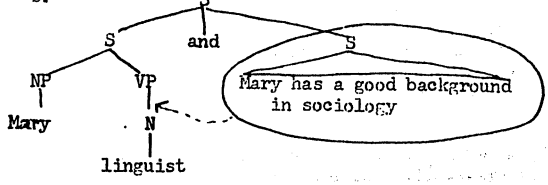
Sentences like (5) present an additional puzzle whose solution will in fact provide a major step towards accomodating such sentences in a coherent treatment of restrictive relatives, namely that the determiners that appear must be 'numerals' in the sense that includes several, a couple, and a few, but not most or few:

(8) *Two linguists and most anthropologists that had met at a ...
*Two linguists and the anthropologists that had met at a ...
A few linguists and several anthropologists that had met at ...
*If two linguists and any anthropologists that had met at the conference had been among those arrested, the university would have fired them.

I propose that this distribution reflects a duality of function on the part of numerals: a numeral, when apparently used as a 'determiner', fulfills both the function of a quantifier (thus conforming to the other things that are classed as 'determiners') and that of an attributive adjective, e.g. Five linguists were arrested involves both an existential quantifier binding a set variable and an adjective predicated of that set variable, as in the paraphrase 'There is a 5-member set of linguists such that they were arrested'. I note in passing that some such reanalysis of numerals is called for in view of the possibility of combining them with definite determiners. If all the things that have commonly been called 'determiners' played the same syntactic role, then none of the three structures in (1) could accomodate NP's such as the two boys or your many publications: each of those structures will accomodate only one determiner per NP unless by fiat more than one 'determiner' is allowed to fill the 'determiner' position.

I have now reached the point where I must recapitulate the analysis of restrictive relative clauses given in McCawley 1977a. I proposed an analysis of restrictive relatives that centers on those that appear on predicate nouns; specifically, a restrictive relative such as that in (9a) is derived from a coordinate structure that also underlies (9b), via a transformation that adjoins one conjunct of a coordinate S to a predicate noun in the other conjunct, provided that some NP in the adjoined S is coreferential to the subject of the host clause:

(9) a. Mary is a linguist who has a good background in sociology
b.

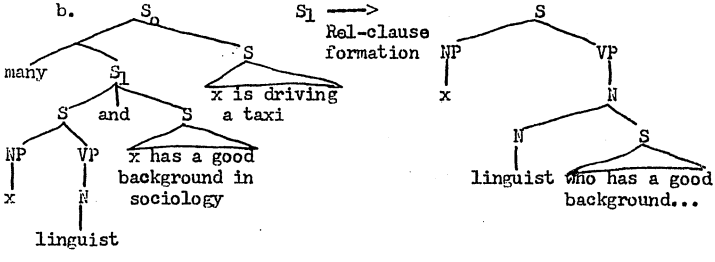


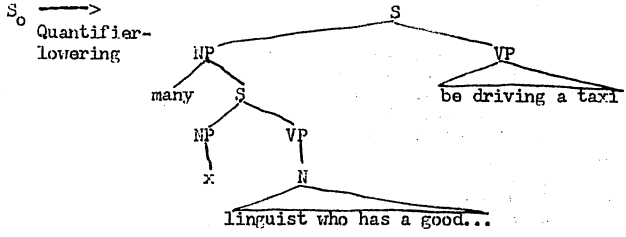
I will henceforth refer to this analysis as the predicate-conjunct analysis. I note parenthetically that the familiar Ross-Thompson analysis of restrictive relatives, in which there is a coordinate

underlying structure and an adjunction that is contingent on coreferentiality between the host NP and something in the adjoined clause (e.g. I caught a fish$_i$ and John ate the fish$_i$ -> John ate the fish that I caught), is incapable of dealing with restrictive relatives on predicate nouns, since predicate nouns are normally non-referential and thus cannot satisfy the coreferentiality condition that Ross and Thompson must impose. In the case of restrictive relatives that are not on predicate nouns, I propose deriving them from structures in which the relevant noun does appear in predicate position, e.g. a quantified NP is derived from an underlying structure in which a S giving the domain of the bound variable appears external to the matrix S, and within that S the relative clause formation sketched in (9) takes place:

(10) a. Many linguists who have good backgrounds in sociology are are driving taxis.

b.





I wish now to propose logical structures for (11a) and (11b) that will provide a reason for the difference in surface position of the numeral, i.e. that will show why (11a) has a surface structure of the shape (1c) and (11b) one of the shape (1b).

(11) a. two cats that once belonged to George
b. two linguists that met at a conference on language planning

In the case of (11a), I wish to say that there is a constituent 'x is a cat and x once belonged to George'; that constituent undergoes Relative-clause formation, yielding 'x is a cat that once belonged to George'; as long as the numeral originates outside of that clause, a surface structure will result in which two is outside of cat that once belonged to George, i.e. a constituent structure of the general shape of (1c). Since I am taking numerals to be predicated of sets, the clause of logical structure containing two will have to be outside of 'x is a cat and x once belonged to George', as otherwise (given the assumption that transformations apply cyclically) Relative-clause formation will not be applicable: Relative-clause formation is contingent on identity between the subject of a predicate noun and something in a conjoined clause, and the variable x is not coreferential with the set of which two is is predicated. Thus, the relative positions of two, cat, and once belonged to George must be as in (12):

(12) $(\exists M: [(\forall x: x \in M)(x \text{ is cat} \wedge x \text{ once belonged to George})] \wedge M \text{ is 2})$
$(\forall x: x \in M)(\text{Tom have } x)$

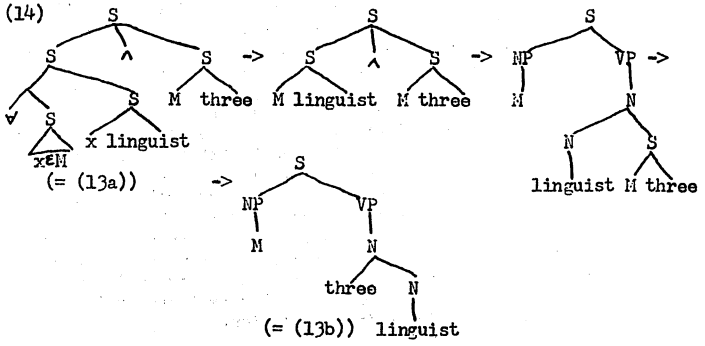
In the case of (11b), there is coreferentiality between the subject of met at the conference (which is the set of individuals, rather than a variable ranging over that set) and the item of which three is predicated. Three, however, is not a noun, and regardless of whether one takes the differences between noun, adjective, and verb to be relevant to logical well-formedness, this distinction is relevant to the surface well-formedness of the ultimate result of Relative-clause formation: head noun position must be filled by a noun, not a verb or adjective. Since I have been assuming that individual properties are predicated of individuals rather than of sets, my Relative-clause formation rule provides no way of adjoining who met at the conference to linguist. Does it provide a way of adjoining it to anything? There is only one other potential constituent of (11b) that could plausibly be regarded as a noun predicated of the set of linguists, namely the complex expression three linguists. Since I have taken both three and linguist to be predicates, no logical structure that I might assign to (11b) will contain three linguists as a constituent; however, a clause having three linguists as a derived predicate could very well arise in the course of the derivation, i.e. if I can justify derivational steps that take me from (13a) to (13b), then I can treat (11b) as having the logical structure (13c) and will have an explanation of why the constituents would have to be grouped together that way in logical structure if a surface structure is to result in which the NP's have nouns as their heads:

(13) a. $(\forall x: x \in M)(x \text{ linguist}) \wedge (M \text{ three})$
b. M (three linguist) [i.e. 'M is three linguists']

c. $(\exists M:(((\forall x:x\varepsilon M)(M \text{ linguist}) \wedge M \text{ three}) \wedge M \text{ met at the conference})\ [(\forall x:x\varepsilon M)(\Delta \text{ arrested } M)]$

There is a fairly simple way in which one could get from (13a) to (13b) using a free ride on the rule of Relative-clause formation. Specifically, if there were a rule taking one from $(\forall x:x\varepsilon M)(x\ f)$ to $(M\ f)$, the result of applying it to (13a) would then meet the conditions for Relative-clause formation, and that plus Relative-clause reduction would yield the structure in (13b):

(14)

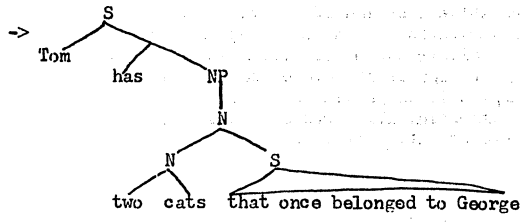


A step deriving 'M linguist' from $(\forall x:x\varepsilon M)(x \text{ linguist})$ is plausible, in that both individual predicates and set predicates act as if they were predicated of the same thing in coordinate structures such as

(15) Those odd chaps in the next room are linguists, are three in number, and met at a conference on language planning.

Moreover, such a step appears to be necessary if the derivation of (11a) is to grind through to completion, since the step of 'Quantifier lowering', which puts the quantified NP into the matrix clause, requires that the bound variable be in the position where the NP is to be inserted, which means that M and not x must be the object of <u>have</u> when Quantifier-lowering applies. My only major worry about this step is that it would give rise to extra derivations for examples such as (11a). For example, there is nothing apparent to exclude derivations such as (16):

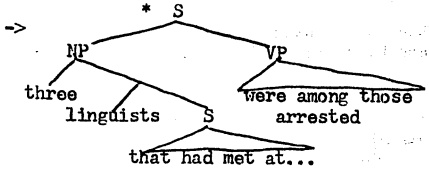
(16) $(\exists M:$ (M two $\wedge$ $(\forall x:x\varepsilon M)(x \text{ cat})) \wedge (\forall x:x\varepsilon M)(x$ once belonged to George)) $(\forall x:x\varepsilon M)$(Tom have x) ->



It is not immediately clear whether it is bad to have such extra derivations and, accordingly, to have alternative surface constituent structures for certain NP's. Note that the argument based on examples (3) only showed that the NP's in question can have a (1c) constituent structure -- it did not rule out the possibility that those NP's might be ambiguous between that constituent structure and some alternative. By contrast, (11b) can only have a constituent structure like (1b): the fact noted in (7) rules out a (1c) structure.

Unfortunately, however, the suggestion of a rule that derives the second tree in (14) from the first one will also give an extra derivation of (11b) yielding the wrong surface structure:

(17) $\exists M$: $(((\forall x{:}x\in M)(x \text{ linguist}) \wedge (M \text{ met at the conference})) \wedge$ M three) (...)



To rule out such a derivation, I can at present offer nothing better than a fiat, specifically, a constraint excluding logical structures in which clauses with a numeral predicate are conjoined with anything other than a clause giving a common property of the set to which the numeral applies, i.e. a constraint that would exclude the logical structure in (17) simply on the grounds that what 'M three' is conjoined with is not of the form $(\forall x{:}x\in M)(x\ f)$.

Given that fiat, there is then only one way that the conjuncts could be grouped in a logical structure for (5). I present that structure in (18), in the process making an arbitrary decision to have the existential quantifier bind variables corresponding to the

two subsets rather than to the whole set of persons:

(18) $(\exists M_1, M_2$: $(\forall x: x \in M_1)$(x linguist) $\wedge$ (M_1 2)) $\wedge$ $(\forall x: x \in M_2)$(x anthrop.) $\wedge$ (M_2 3)) $\wedge$ ($M_1 \cup M_2$ met at the conference)) $(\forall x: x \in M_1 \cup M_2)$(x was among those arrested)

For Relative-clause formation to convert '$M_1 \cup M_2$ met at the conference' into a relative clause, the S to which it is adjoined must have the form '$M_1 \cup M_2$ is N' for some predicate noun N. There is exactly one way that such a structure can be obtained via existing rules from the given logical form: after the steps in (14) have converted the innermost coordinate structure into 'M_1 is two linguists and M_2 is three anthropologists', Conjunction-reduction (in the generalized form that covers the respectively-construction) applies:

(19) (M_1 two linguist) and (M_2 three anthropologist) -> ($M_1 \cup M_2$) (two linguist and three anthropologist)

See McCawley 1968 for discussion of this step, including the point that derived coordinate NP's have as reference the union of the references of the conjuncts. Relative-clause formation can now apply to adjoin '$M_1 \cup M_2$ met at the conference' to 'two linguist and three anthropologist', and Quantifier-lowering then inserts the resulting expression in place of $M_1 \cup M_2$ in the matrix clause, yielding the desired surface structure.

I turn now to the question of what the logical structure and derivation of (4) are. I will lead up to an answer by treating first a similar example that does not involve relative clauses:

(20) Several girls and boys were in the room.

The problem here is to derive the compound noun girls and boys and in the process account for the fact that the conjunction is and, not or. The most obvious proposal for the logical structure is (21), but that would commit one to having an ad-hoc conversion of or to and:

(21) $(\exists M: (\forall x: x \in M)$(x girl v x boy) $\wedge$ (M several)) (...)

If we insist on deriving the and from something that otherwise underlies and, the only possibility that I can see is to invoke again the respectively construction and assign to (20) a logical structure as in (22), with a derivation as indicated:

(22) $(\exists M_1, M_2$: $(\forall x: x \in M_1)$(x girl) $\wedge$ $(\forall x: x \in M_2)$(x boy)) $\wedge$ ($M_1 \cup M_2$ several)) $(\forall x: x \in M_1 \cup M_2)$(x was in the room) ->
$(\exists M_1, M_2$: (($M_1 \cup M_2$)(girl and boy)) $\wedge$ ($M_1 \cup M_2$ several)) $(\forall x$:...) ->
$(\exists M_1, M_2$: (($M_1 \cup M_2$)(several(girl and boy)))) $(\forall x: x \in M_1 \cup M_2)$...

The and will then come from one of its standard sources: set union. The derivation of (4) will parallel that in (22) except that where (22) has atomic clauses 'x girl' and 'x boy', there will be coordinate structures 'x linguist and x play chess' and 'x philosopher and x play bridge', and applications of Relative-clause formation to those structures will precede the steps in (22).

I will conclude this paper by indicating how the logical structures argued for here assist one in making sense of a famous puzzle, namely that of the sentences discovered by Ross and Perlmutter 1970 that apparently involve an extraposed relative clause though there is nowhere that that clause could be extraposed from:

(23) A man entered and a woman left who had met in Vienna.
*A man who had met in Vienna entered and a woman left.
*A man entered and a woman who had met in Vienna left.

These sentences share the peculiarity of relative clause constructions like (5) that the NP's must have numerals for determiners:

(24) a. Two men entered and three women left who had met in Vienna
b. *The men entered and three women left ...
c. *Every man entered and three women left...

The proposals made above are in fact not sufficient to provide a derivation for (23). However, they provide the basis of a derivation involving what Morgan (1972) calls 'patches': devices for extending the grammar in a natural and minimal way to cases that it strictly speaking does not cover. Specifically, I maintain that (23) has a logical structure that normal syntactic rules by themselves do not relate to any surface structure but which can be given a well-formed surface expression if a certain obstruction to normal derivational steps is removed in an inobtrusive fashion, namely if a clause having the semantic role of a relative clause is allowed to bypass normal syntactic steps and be inserted directly in a surface position appropriate to its syntactic structure and semantic role, i.e. the position reserved for extraposed relative clauses. I take (24a) to have the logical form (25):

(25) $(\exists M_1, M_2: (((\forall x: x \in M_1)(x \text{ man}) \wedge (M_1\ 2)) \wedge ((\forall x: x \in M_2)(x \text{ woman}) \wedge M_2\ 3) \wedge (M_1 \cup M_2 \text{ met in Vienna}) (\ldots \wedge \ldots)$

Relative clause formation can derive two men and three women from the first two pairs of conjuncts, but there is no way for it to adjoin '$M_1 \cup M_2$ met in Vienna' to any NP unless Conjunction-reduction is carried out, in which case what results is not (24a) but (26):

(26) 2 men and 3 women who had met in Vienna entered and left resp.
2 men and 3 women entered and left (resp.) who had met in Vienna.

My proposed patch moves '$M_1 \cup M_2$ met in Vienna' directly into the extraposed relative position of the matrix clause 'M_1 entered and M_2 left', circumventing the need to adjoin it to a head noun the way a relative clause would normally have to be. The contribution of the system of logical structure sketched above to the analysis of these examples is to provide something for a patch to patch: a structure accomodating all parts of the sentence and allowing one to identify the semantic roles of all of those parts.[4]

FOOTNOTES

[1] I now interpret the many arguments (Lakoff 1970, Bach 1968, McCawley 1972) allegedly showing that syntactic categories are identical to logical categories as merely having shown that a lot of syntactic phenomena are sensitive to logical category or reflect the combinatorics of logical structure. The arguments are neutral as to whether syntactic rules are sensitive to other factors, and a major failing of the tradition in which they were presented was a failure to attempt to identify the full range of factors that play a role in syntax. This failure can be attributed in part to the authors' retention of a conception of syntax in which syntactic category membership determined the applicability of transformations and in which the syntactic categories that figured in the base rules were all the syntactic categories there were.

[2] Here, as at many places in this paper, it is immaterial whether the 'deepest' trees exhibited are derived from anything still deeper or not.

[3] For discussion of the contributions of these quantifiers to semantic structure, see McCawley 1977c and Peterson 1977.

[4] In McCawley 1977a I argue for a patch analysis of another well-known case of relativization, namely NP's such as the aspersions that he cast on my character, in which the head noun is part of an idiom, the remainder of which is in the relative clause. I demonstrate that Brame's proposal of an underlying structure in which there is no head noun, with a copying transformation creating a head noun from material inside the relative clause, is not tenable as a general analysis of restrictive relatives but can be interpreted as a patch that allows the formation of relative clauses in a class of cases that normal Relative-clause formation does not cover.

REFERENCES

Bach, Emmon. 1968. Nouns and noun phrases. Bach and Harms, Universals in linguistic theory, New York: Holt, 90-122.

Lakoff, George. 1970. Irregularity in Syntax. New York: Holt.

McCawley, James D. 1968. The role of semantics in a grammar. Bach and Harms, Universals in linguistic theory, New York: Holt, 124-169.

---. 1972. A program for logic. Davidson and Harman, Semantics of natural language, Dordrecht: Reidel, 498-544.

---. 1977a. Relative and relative-like clauses. Paper read at 1977 Nijmegen conference on semantics.

---. 1977b. The nonexistence of syntactic categories. Proceedings of 1977 Michigan State conference on linguistic metatheory.

---. 1977c. Lexicographic notes on English quantifiers. Papers from the 13th regional meeting, CLS, 372-383.

Morgan, J. L. 1972. Verb agreement as a rule of English. Papers from the 8th regional meeting, CLS, 278-286.

Partee, B, H. 1975. Montague grammar and transformational grammar. Linguistic Inquiry 6.203-300.

Peterson, P. 1977. On the logic of 'few', 'many', and 'most'. Notre Dame Journal of Formal Logic

Ross, J. R. 1967. Constraints on variables in syntax. MIT thesis.

Ross, J. R., and D. M. Perlmutter. 1970. Relative clauses with split antecedents. Linguistic Inquiry 1.350.

Smith, C. 1964. Determiners and relative clauses in a generative grammar of English. Language 40.37-52.

Stockwell, R. P., Partee, B. H., and Schachter, P. 1972. The major syntactic structures of English. New York: Holt.

Thompson, S. A. 1971. The deep structure of relative clauses. Fillmore and Langendoen, Studies in linguistic semantics, New York: Holt, 78-94.