
Unbounded Dependencies and Coordinate Structure

Author(s): Gerald Gazdar

Source: *Linguistic Inquiry*, Vol. 12, No. 2 (Spring, 1981), pp. 155-184

Published by: The MIT Press

Stable URL: <http://www.jstor.org/stable/4178215>

Accessed: 07/02/2009 16:23

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=mitpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



The MIT Press is collaborating with JSTOR to digitize, preserve and extend access to *Linguistic Inquiry*.

Unbounded Dependencies and Coordinate Structure

Consider eliminating the transformational component of a generative grammar. In particular, consider the elimination of all movement rules, whether bounded or unbounded, and all rules making reference to identity of indices. Suppose, in fact, that the permitted class of generative grammars constituted a subset of those phrase structure grammars capable only of generating context-free languages. Such a move would have two important metatheoretical consequences, one having to do with learnability, the other with processability. In the first place, we would be imposing a rather dramatic restriction on the class of grammars that the language acquisition device needs to consider as candidates for the language being learned. And in the second place, we would have the beginnings of an explanation for the obvious, but largely ignored, fact that humans process the utterances they hear very rapidly.¹ Sentences of a context-free language are provably parsable in a time which is proportional to the cube of the length of the sentence or less (Younger (1967), Earley (1970)). But no such restrictive result holds for the recursive or recursively enumerable sets potentially generable by grammars which include a transformational component.

My strategy in this article will be to assume, rather than argue, that there are no transformations, and then to show that purely phrase structure (PS) treatments of coordination and unbounded dependencies can offer explanations for facts which are unexplained, or inadequately explained, within the transformational paradigm.

In section 1, I shall briefly outline a phrase structure treatment of coordination that eliminates the need for a rule of Coordination Reduction (CR). This proposal is but a minor variant of a schema that has become standard in recent years. Section 2 shows how the syntax and semantics of unbounded dependencies can be handled in a PS grammar which employs complex symbols. English relative clauses and constituent questions are used to exemplify the apparatus. Section 3 then demonstrates that Ross's

I am indebted, for comments, conversation, and criticism, to Emmon Bach, Lee Baker, Anne Cutler, Elisabet Engdahl, Janet Fodor, Paloma Garcia-Bellido, Paul Hirschbühler, Dick Hudson, Aravind Joshi, Ewan Klein, Robert May, Jim McCloskey, Barbara Partee, Stan Peters, Geoff Pullum, Andrew Radford, Ivan Sag, Lisa Selkirk, Neil Smith, Jean-Roger Vergnaud, Edwin Williams, and two anonymous LI referees. This research was supported by grant HR 5767 from the SSRC (UK).

¹ See, for example, Marslen-Wilson (1973).

Coordinate Structure Constraint (CSC) and the “across-the-board” (ATB) violations of it follow as theorems from the grammar fragments given in the previous two sections. Detailed consideration is given to the data, and to an alternative analysis, given in Williams (1977; 1978). Finally, section 4 argues that the apparent boundedness of rightward dependencies is a consequence of independently motivated constraints on parsing and that the syntax can thus treat such dependencies as unbounded. A general schema is given for rightward dependencies, and it is shown that this interacts with the coordination schema to generate all those sentences produced by the transformation known as Right Node Raising (RNR), together with some that such a rule ought to produce but cannot.

I shall follow McCawley (1968) and interpret phrase structure rules as node admissibility conditions rather than as string-to-string mapping rules. Accordingly, I shall not use the familiar rewrite arrow notation for PS rules, but shall instead use a notation which reflects more directly the relation such rules bear to the (sub)trees that they admit. Instead of (1), then,

- (1) $S \rightarrow NP VP$

I shall write (2),

- (2) $[_S NP VP]$

and analogously for all other rules.

I shall assume that each syntactic rule in the grammar should be associated with a semantic rule which gives the meaning of the constituent created by the syntactic rule as a function of the meaning of the latter's parts.² I assume further that the semantic rules should take the form of rules of translation into intensional logic.

I take a rule of grammar to be a triple of which the first member is an arbitrary integer (the number of the rule), the second is a PS rule, and the third is a semantic rule showing how the intensional logic representation of the expression created by the PS rule is built up from the intensional logic representations of its immediate constituents. I shall use a Montague-like prime convention in the semantic rules: NP' stands for the (complex) expression of intensional logic which is the translation of the subtree dominated by NP, run' is the constant of intensional logic which translates the word *run* in English, etc.

Within this framework, the first rule of a grammar might be this:

- (3) $\langle 1, [_S NP VP], VP'(\text{'NP'}) \rangle$

I shall use *rule* to refer both to the triple and to its second and third members (sometimes

² This assumption commits us to what Bach (1976, 2) has called the *rule-to-rule* hypothesis concerning the semantic translation relation. Nothing to be said below hinges on our adopting this hypothesis, and the syntax given is entirely compatible with a less constrained view of the syntax-semantics relation, for example one in which the syntax simply admits structures, and a set of completely unrelated semantic rules interprets those structures.

qualifying the latter with *syntactic* and *semantic*, respectively), but this should not lead to any confusion.

1. Coordination

Consider (4):

$$(4) \quad \alpha \rightarrow \alpha_1 \dots \left\{ \begin{array}{c} \text{or} \\ \text{and} \end{array} \right\} \alpha_n$$

where α is any syntactic category

Phrase structure rule schemata more or less like (4) are nothing new. They date back at least as far as Dougherty (1970) and can be found in many more recent works (e.g. Jackendoff (1977, 51)). But they have never amounted to a sufficient theory of constituent coordination in transformational grammar, for two reasons. First, as long as one retains a rule like Passive as an operation which maps into sentences rather than into verb phrases, it is impossible, as Schachter (1976, 236–237) points out, to use the schema to generate the conjoined VPs in a sentence like (5):

(5) The Dodgers beat the Red Sox and were beaten by the Giants.

To handle such sentences, one has to postulate a rule of Coordination Reduction (CR) which applies to coordinate sentences, deletes material under identity, and regroups the remainder. Second, no transformational grammarian has ever indicated what the semantics for the rule schema might be. This point should not be taken lightly: one advantage of the CR theory was that it treated (almost) all coordination as underlyingly sentential, and a fully explicit semantics for sentential coordination was to be found in propositional logic. The CR theory consequently made strong explicit predictions about what reduced coordinate sentences meant. But in the absence of any semantics, the schema in (4) makes no semantic predictions. For example, it does not even predict that (5) is synonymous with (6):

(6) The Dodgers beat the Red Sox and the Dodgers were beaten by the Giants.

However, if (as we are assuming) there are no transformations, then there is no Passive transformation and the syntactic problem does not arise. And the semantic problem is solved by Cooper (forthcoming), Gazdar (1980a), and Keenan and Faltz (1978), who have independently provided a model-theoretic semantics for crosscategorical coordination (without invoking an interpretive rule of CR-in-Reverse).

In the classical analysis of a sentence like (5), we begin with a sentence like (7):

(7) The Dodgers beat the Red Sox and the Giants beat the Dodgers.

Passive applies in the second conjunct to give us (6), and then CR applies to give us (5). This analysis is hopeless for a sentence like (8) (cf. Jackendoff (1977, 193–194)):

(8) Different teams beat the Red Sox and were beaten by the Giants.

The classical analysis can handle the syntax, of course, but only at the cost of making nonsense of the semantics.

There is one remaining problem with schema (4): it assigns coordinate expressions the wrong surface structure. Ross (1967, 90–91) shows that there are phonological and syntactic reasons for believing that the coordinating word forms a constituent with the immediately following node and is not simply a sister of all the conjuncts. Suppose then that we allow the names of coordinating morphemes to appear as features on categories, and eliminate the feature by means of a rule schema which expands such categories as the named coordinating morpheme, followed by the category (as in (10), below). This will in turn allow us to revise (4) in such a way as to capture Ross's observations regarding surface constituent structure:

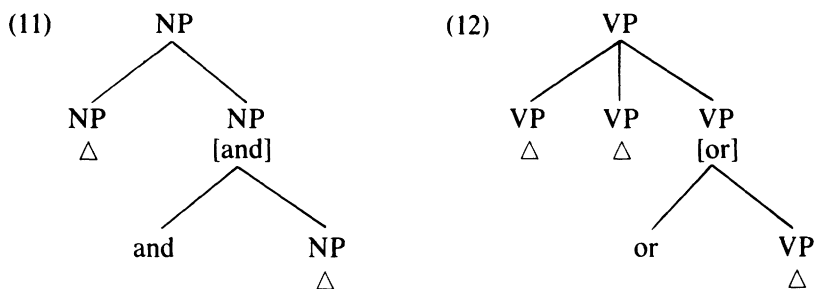
$$(9) \langle 2, [\alpha \alpha_1, \dots, \alpha_n]_{[\beta]}, \beta'(\alpha_1', \dots, \alpha_n') \rangle$$

where $\beta \in \{\text{and, or}\}$ and α is any syntactic category

$$(10)^3 \langle 3, [\alpha \beta \alpha]_{[\beta]}, \alpha' \rangle$$

where $\beta \in \{\text{and, or, } \dots\}$ and α is any syntactic category

Schemata (9) and (10) then combine to give us subtrees like these:



Note that (9) is only intended to handle constituent coordination and thus make the CR transformation otiose. There is no claim that, on its own, it will do the work associated with such transformations as Right Node Raising (see section 4 below) or Gapping (see Stump (1978) for a nontransformational treatment). In particular, we must assume that whatever mechanism is responsible for “gapped” VPs (i.e. VPs lacking the head verb and possibly contiguous constituents) is responsible for Williams's (1978, 38) example (13), rather than CR, as Williams assumes.

³ We can allow β in (10) also to range over the comparative complementizers *as* and *than*, but I shall not pursue this possibility here (see Gazdar (1980b) for discussion). Given schema (10), we can provide a phrase structure rule for the idiomatic *go and VP*, *come and VP* construction:

$$(i) \begin{bmatrix} {}_{VP} V (PP) VP \\ [and] \end{bmatrix}$$

Following Ross (1967, 94) and Schachter (1977, 100), I am assuming that this construction does not involve genuine coordinate structure. But, pace Bever, Carroll, and Hurtig (1976, 170–173), I do assume that the construction should be characterized by the grammar of English.

(13) John gave the books to Mary and the records to Sue.

The problem here, as Williams notes, is that the direct and indirect objects of *give* do not form a constituent and cannot therefore be conjoined in virtue of the schema in (9).

2. Unbounded Dependencies

Phrase structure grammars can handle unbounded dependencies in an elegant and general way, provided that we exploit the resources offered by a complex symbol system and by the possibility of making statements about the set of rules that the grammar may employ.⁴ The paragraphs that follow will, unavoidably, be somewhat technical.

Let V_N be the set of *basic* category symbols (i.e. the set of all nonterminal symbols standardly used). Then we define a set $D(V_N)$ of *derived* categories as follows:

$$(14) D(V_N) = \{\alpha/\beta: \alpha, \beta \in V_N\}$$

Suppose, counterfactually, that S and NP were the only basic categories; then the set of derived categories would consist of S/S, S/NP, NP/NP, and NP/S. This notation is reminiscent of categorial grammar but, despite a tenuous conceptual link, these derived categories are not to be interpreted in the way categorial grammar prescribes. The intended interpretation is as follows: a node labeled α/β will dominate subtrees identical to those that can be dominated by α , except that somewhere in every canonical subtree of the α/β type there will occur a node of the form β/β dominating a resumptive pronoun, a phonologically null dummy element, or the empty string, and every node linking α/β and β/β will be of the σ/β form. Intuitively, then, α/β labels a node of type α which dominates material containing a hole of type β (i.e. an extraction site on a movement analysis). So, for example, S/NP is a sentence which has an NP missing somewhere.

Of course, defining a new set of syntactic categories is not of itself sufficient to ensure that the trees in which they figure have the property just described: we need, in addition, a set of rules to employ them.

What we have to do is define a set of rules each of which expands a derived category just as the corresponding basic rule would have done for the basic category, except that exactly one of the dominated categories is now paired with the same hole-indicating category as is the dominating category. The set of such rules will consequently allow the hole information to be “carried down” the tree.⁵

Let G be the set of *basic* rules (i.e. the set of rules that a grammar not handling unbounded dependencies would require). For any syntactic category β , there will be some subset of the set of the nonterminal symbols V_N , each of which can dominate β according to the rules in G . Let us call this set V_β ($V_\beta \subseteq V_N$). Now, for any category

⁴ Harman (1963) deserves the credit for first seeing the potential of PS grammars incorporating a complex symbol system. The idea of employing a grammar to induce the grammar is due originally to Van Wijngaarden (1969). For some discussion of the properties of systems which exploit a grammar to generate a grammar, see Langendoen (1976) and Gazdar (forthcoming).

⁵ Cf. Baker (1978, 113), Hellan (1977, 128).

β ($\beta \in V_N$) we can define a (finite) set of *derived* rules $D(\beta, G)$ as follows:

$$(15) \quad D(\beta, G) = \{ [\alpha/\beta \sigma_1 \dots \sigma_i/\beta \dots \sigma_n] : [\alpha \sigma_1 \dots \sigma_i \dots \sigma_n] \in G \text{ \& } 1 \leq i \leq n \text{ \& } \alpha, \sigma_i \in V_\beta \}$$

An example of the application of (15) should make this clearer. Suppose that the set G of basic rules looks like this:

- (16) a. $\{[{}_S \text{ NP VP}],$
 b. $[{}_{VP} \text{ V } \overline{\text{VP}}],$
 c. $[{}_{VP} \text{ V NP}],$
 d. $[{}_{PP} \text{ P NP}],$
 e. $[\bar{s} \text{ that S}],$
 f. $[{}_{VP} \text{ V } \bar{S}],$
 g. $[\overline{\text{VP}} \text{ to VP}],$
 h. $[{}_{VP} \text{ V NP PP}],$
 i. $[{}_{NP} \text{ NP PP}]\}$

Then the set $D(NP, G)$ will look like this:

- (17) a. $\{[S/NP \text{ NP/NP VP}], \quad [S/NP \text{ NP VP/NP}],$
 b. $[VP/NP \text{ V } \overline{VP/NP}],$
 c. $[VP/NP \text{ V NP/NP}],$
 d. $[PP/NP \text{ P NP/NP}],$
 e. $[\hat{S}/NP \text{ that S/NP}],$
 f. $[VP/NP \text{ V } \hat{S}/NP],$
 g. $[\overline{VP}/NP \text{ to VP/NP}],$
 h. $[VP/NP \text{ V NP/NP PP}], \quad [VP/NP \text{ V NP PP/NP}],$
 i. $[NP/NP \text{ NP/NP PP}], \quad [NP/NP \text{ NP PP/NP}]\}$

The set $D(PP,G)$ will look like this:

- (18) a. {[_{S/PP} NP/PP VP], [_{S/PP} NP VP/PP],
b. [_{VP/PP} V $\overline{\text{VP/PP}}$],
c. [_{VP/PP} V NP/PP],
d. [_{PP/PP} P NP/PP],
e. [$\tilde{\text{S/PP}}$ *that* S/PP],
f. [_{VP/PP} V $\tilde{\text{S/PP}}$],
g. [$\overline{\text{VP/PP}}$ *to* VP/PP],
h. [_{VP/PP} V NP/PP PP], [_{VP/PP} V NP PP/PP],
i. [_{NP/PP} NP/PP PP], [_{NP/PP} NP PP/PP]}

Derived rules have no special lexical or semantic properties. Thus, all derived rules will have the same rule numbers, the same subcategorization properties, and the same semantic translations as the basic rules from which they derive. Consequently, they do

not need to be separately listed or separately specified; everything about them can be predicted from (15), taken together with the basic rules.⁶

We can formalize certain island constraints, if we wish, simply by stipulating that certain types of derived rule are not employed by a language (or by any language, if the constraint suggested is intended as a universal).⁷ Suppose we wanted to impose the A-over-A Constraint. Then we could add a condition to (15) that $\alpha \neq \beta$. This would have the effect of preventing the creation of any derived rules of the form shown in (19):

$$(19) [\alpha/\alpha \dots]$$

Adding a condition that $\alpha \neq \text{NP}$ to (15) would have the effect of imposing Horn's (1974) NP Constraint. And language-particular constraints, such as one against stranding prepositions, can also be imposed by prohibiting certain classes of derived rule (all rules of the form $[\text{PP/NP} \dots]$ in the case of impermissible preposition stranding).

Thus, constraints once thought of as constraints on permissible movement can be reconstructed as constraints on permissible rules.⁸

The only island constraint I want to consider here is Ross's Left Branch Condition (1967, 114):

(20) *The Left Branch Condition*

No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule.

This could be approximately reconstructed in terms of derived rules as (21):

$$(21) *[\text{NP/NP NP/NP} \dots]$$

I propose to adopt a generalized version of (21) in the discussion that follows. The Generalized Left Branch Condition (GLBC) will block NP dependencies into left branches of any constituent defined by (15):

(22) *Generalized Left Branch Condition*

$$*[\alpha/\beta \sigma/\beta \dots]$$

where α and σ are any node labels, and $\beta = \text{NP}$

GLBC could be incorporated in (15), if one wished, by stipulating that $\neg(i = 1 \wedge \beta$

⁶ Definitions (14) and (15) allow only one hole to be "carried down" through any node in the tree. This is desirable for English, which, by and large, does not allow multiple dependencies. However, (14) and (15) would need to be generalized to handle, for example, Scandinavian languages. See Engdahl (1980), Gazdar (forthcoming), Maling and Zaenen (forthcoming) for discussion of some of the issues involved.

⁷ The question of how we formalize universal constraints is, of course, distinct from the question of how we motivate them. The discussion here addresses only the question of formalization, a question that is rarely taken up in the literature.

⁸ A constraint like Subjacency cannot be reconstructed this way. However, if one wanted to impose it, then one could recast it as a tree filter and throw out all trees that involved a violation. Use of tree filters of this type would not allow the overall theory to generate any non-CF languages. This can be proved straightforwardly by defining a tree automaton of the appropriate kind, but I will not pursue this here. See Thatcher (1973) and Levy (forthcoming) for discussion of tree automata.

= NP). I will consider the motivation for this implausible-looking island constraint later in the article.⁹

In addition to derived rules, we need *linking* rules (these will be a subset of the basic rules) to introduce and eliminate derived categories. For the majority dialect of English (British or American), we need only the following rule schema to eliminate derived categories:

- (23) $\langle 4, [\alpha/\alpha \ t], h_\alpha \rangle$
 where $\alpha \in V_N$

Here h_α (mnemonic for *hole*) is a distinguished variable ranging over denotations of type α (i.e. NP denotations if $\alpha = \text{NP}$, PP denotations if $\alpha = \text{PP}$, etc.). t is a dummy element postulated solely for phonological reasons (that is, it will serve to block contraction). It serves no semantic function (h_α is the variable, not t), and for other dialects or languages we could replace t with the empty string \bar{e} (which would have no phonological effects) or with a proform. It will become apparent in what follows that t is placed precisely in those complement subject positions where contraction-inhibiting phonological effects have been noted, and consequently the analysis faces none of the difficulties besetting the analyses criticized in Postal and Pullum (1978).

The apparatus developed above can be used to handle all constructions involving an unbounded dependency.¹⁰ However, since exactly the same principle is involved in every case, it will suffice here to illustrate its application by reference to just two constructions, namely English restrictive relative clauses and constituent questions.

I will assume that relative clauses are dominated by a sentential category R: if one is in the business of generating surface structures directly by means of context-free PS rules, then one cannot identify R with \bar{S} for obvious reasons. We will distinguish sentential categories by means of the following features: $\pm C$ (complement), $\pm R$ (relative), and $\pm Q$ (interrogative). In this system, S is $[-C, -R, -Q]$, R is $[+C, +R, -Q]$, root interrogatives ("Q") are $[-C, -R, +Q]$, and embedded interrogatives (" \bar{Q} ") are $[+C, -R, +Q]$.¹¹ If we assume that $[+C, -R, -Q]$ can expand only as *that*-S (i.e. the *that* is not optional), then we can use \bar{S} as an abbreviation for $[\pm C, -R, -Q]$ (thus regaining the optionality of *that* in most environments). Some verbs and adjectives subcategorize for $[+C, -R, -Q]$ rather than \bar{S} ; consequently, the *that* is obligatory in these contexts (see Shir (1977, 62–63) for relevant data). Likewise, sentential subjects and topicalized and extraposed clauses must be $[+C]$.

⁹ The GLBC is claimed to capture a fact about English, rather than a linguistic universal. Following the arguments in Klein (1980), we take *how many* to be an AP in the following sentence (which I owe to Emmon Bach and Barbara Partee):

(i) How many did you buy of those pies at the fair?

Compare the following French sentence (due to Jean-Roger Vergnaud):

(ii) Combien as-tu donné de ces livres à ces gens?
 how many have you given of these books to these people

¹⁰ Rules are given for the various comparative constructions in Gazdar (1980b), and for topicalization and free relatives in Gazdar (forthcoming). Gazdar, Pullum, and Sag (1980) gives a rule for VP fronting.

¹¹ The R and Q features can be motivated by the difference between relative and interrogative pronouns found in many languages (e.g. Albanian, German, Hindi).

The rule which introduces R is simply this:¹²

$$(24) \langle 5, [_{NP} NP R], \lambda R[NP'](R') \rangle$$

The semantics I am assuming is essentially the NP–S semantics introduced by Cooper (1975) and Bach and Cooper (1978), and developed by McCloskey (1978; 1979), from which further details should be sought. All NPs are assumed to contain a free set variable *R*. Rule (24) abstracts on this variable and applies the resulting function to the set denoted by the relative clause. The relative clause meaning is thus quantified into the NP meaning. The following two rules expand R,

$$(25) \langle 6, [_{R} (NP) S/NP], \lambda n[\lambda h_{NP}[(S/NP)'](NP') \wedge R(n)] \rangle$$

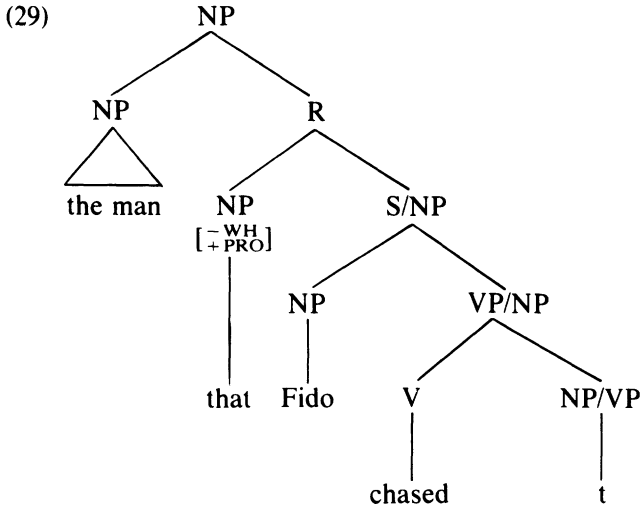
$$(26) \langle 7, [_{R} PP S/PP], \lambda n[\lambda h_{PP}[(S/PP)'](PP') \wedge R(n)] \rangle$$

where¹³

$$(27) NP[\pm R, +WH, +PRO] \rightarrow \textit{who}, \text{ where } \textit{who}' = \lambda PP(n)$$

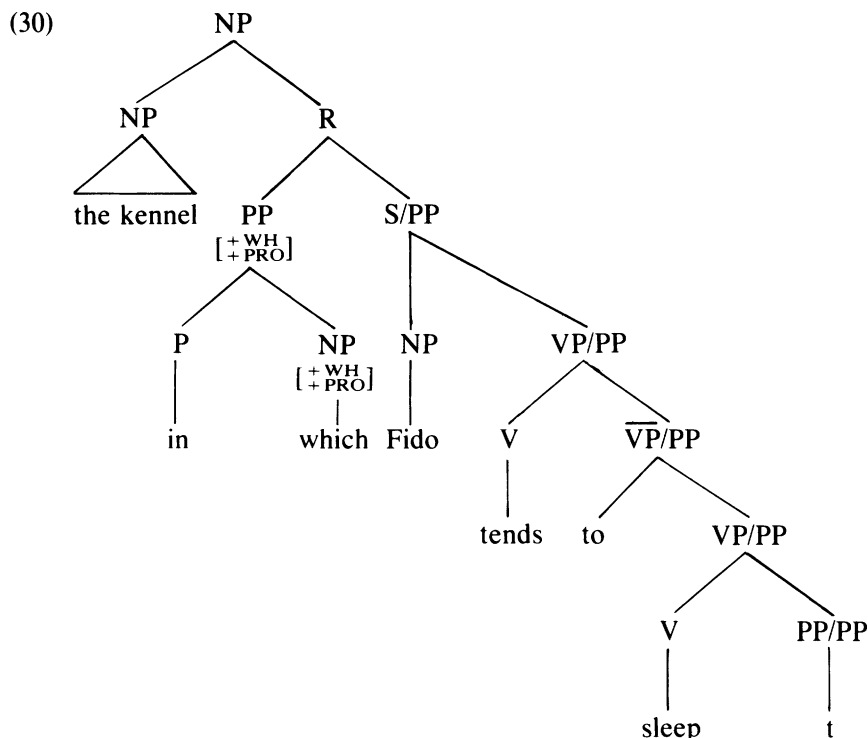
$$(28) NP[+R, -WH, +PRO] \rightarrow \textit{that}, \text{ where } \textit{that}' = \lambda PP(n)$$

These rules induce trees such as (29), (30):



¹² In view of the arguments in Andrews (1975), I have assumed an NP–S syntax in (24), but nothing in the present framework prevents one from adopting a NOM–S syntax: one would just replace (24) with $[_{NOM} NOM R]$ and change the semantics in appropriate ways. Likewise, the rule given allows relatives to stack (pace Jackendoff (1977, 185–190)), but this is also an entirely incidental aspect of the analysis.

¹³ We assume that the R feature, introduced originally by Andrews (1975, 16), is carried down onto the head of the relative clause in virtue of the Head Feature Convention discussed by Gazdar, Pullum, and Sag (1980) and Klein (1980). We also assume that the features mentioned in (27) and (28) can trickle through NP and PP onto embedded NPs to give us such relative heads as *in which*, *for whom*, *whose book*, *in whose book*, etc.



The syntax here is fairly straightforward. Relative clauses consist of a pronominal NP or PP followed by an S with a hole of the appropriate sort in it. In Chaucerian English, this would have been an \bar{S} rather than an S (see Bresnan (1976, 357)). The rules given capture the following facts without any need for separate statements (i.e. conditions, filters, etc.): (i) no pied-piping in *that*-relatives, (ii) optional pied-piping in *wh*-relatives, (iii) obligatory absence of a *that*-complementizer in Modern English *wh*-relatives, (iv) obligatory *c*-command position of the *wh*-phrase, (v) at most one hole in the complement S (except in ATB cases), and (vi) at least one hole in the complement S.

The semantics given in (25) and (26) binds the free variable h_{XP} in the translation of S/XP and applies the resulting function to XP' , which will itself contain a free occurrence of the designated variable n . This variable is then bound by abstraction to form an expression denoting a set.¹⁴ An additional complication is the reintroduction of the R variable in order to allow stacked relatives and contextual binding (see McCloskey (1979, 217–221) on the need for the latter).

Notice that, given our adoption of the GLBC, we at present have no way of generating relatives with subject dependencies. Thus, we cannot derive any of the examples in (31):

¹⁴ The semantic rule for (25) when the NP is missing is as given except that NP' is replaced by $\lambda PP(n)$.

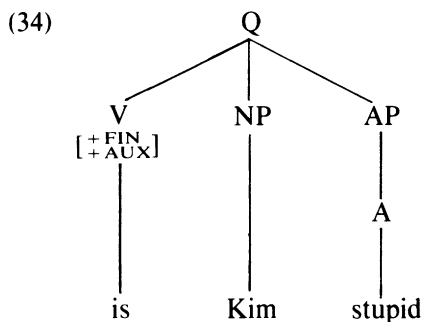
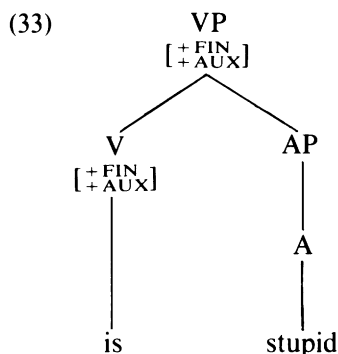
- (31) a. *The man chased Fido returned.
 b. The man that chased Fido returned.
 c. The man (who) I think chased Fido returned.
 d. *The man (who) I think that chased Fido returned.

I shall return to this issue when we have considered interrogatives.

We can achieve Subject–Auxiliary Inversion in a PS grammar for the English auxiliary system by means of the following metarule:¹⁵

$$(32) \quad \begin{array}{c} [_{VP} \ V \ X] \\ \left[\begin{array}{c} +FIN \\ +AUX \end{array} \right] \end{array} \Rightarrow \begin{array}{c} [_{Q} \quad V \quad NP \ X] \\ \left[\begin{array}{c} +FIN \\ +AUX \end{array} \right] \end{array}$$

This says that for every VP rule which introduces a tensed auxiliary verb, there is also to be a corresponding rule expanding the sentential category Q as the auxiliary verb, followed by NP, followed by whatever the auxiliary verb subcategorizes for. Thus, the VP rule responsible for a subtree like (33) will be mapped by the metarule into a Q rule which in turn will induce subtrees like (34):



Given (32), the rule for root constituent questions follows straightforwardly:

$$(35) \quad \langle 8, [_{Q} \quad \alpha \quad Q/\alpha], \lambda p \exists n [\lambda h_{\alpha} [[(Q/\alpha)'] (p)] (\alpha')] \rangle$$

where $\alpha \in \{NP, PP, AP, AdvP\}$

The semantics assumed in (35) is due, essentially, to Karttunen (1977). This rule then allows us to generate such sentences as (36a–d):

¹⁵ Metarule-like operations are proposed in Vergnaud (1973) and Roeper and Siegel (1978). For full discussion of the properties of metarules, see Gazdar (forthcoming). The metarule given in (32) oversimplifies matters in a number of respects that are tangential to the main themes of this article. In particular, we need to distinguish the category which immediately dominates the subject–auxiliary construction from the one that dominates root constituent questions. See Gazdar, Pullum, and Sag (1980) for a more satisfactory metarule theory of Subject–Auxiliary Inversion.

- (36) a. Who did you think Mary saw?
 b. In which car was the man seen?
 c. How slowly would you say he was driving?
 d. How suspicious was Mary?

But, as it stands, (35) will not allow us to generate the following, perfectly grammatical, sentences of English:

- (37) a. Who saw the man?
 b. Which man drove the car?

There is no auxiliary verb in either (37a) or (37b), but since Q necessarily dominates an auxiliary verb, these examples are inevitably excluded from the scope of (35). Clearly, what we need is a rule like (38):

- (38)
$$\begin{bmatrix} Q & NP & VP \\ [+WH] & [+FIN] \end{bmatrix}$$

Now we could simply add (38) to the rules of the grammar and leave it at that. But it would be more interesting if we could show that the form of (38) followed from some general principle, for example a metarule. Suppose that the grammar of English includes the following metarule:

- (39)
$$\begin{bmatrix} \alpha X & \Sigma / NP \dots \\ [-C] \end{bmatrix} \Rightarrow \begin{bmatrix} \alpha X & VP \dots \\ [+FIN] \end{bmatrix}$$

where X contains at least one major category symbol, where α is anything, and where Σ ranges over sentential categories

Consider what this metarule requires: for every rule in the grammar which introduces some $[-C]$ sentential category with an NP hole in it (i.e. Q/NP, S/NP, and \tilde{S} /NP—since \tilde{S} can be $[-C]$) which has as a left sister at least one major category symbol (i.e. N, NP, V, VP, A, AP, P, etc.), there is to be a corresponding rule which is identical except that the Σ/NP is replaced by a tensed VP.¹⁶ The GLBC blocks all subject dependencies; but, if subject dependencies were permitted, then the residues (i.e. the sentences with *t* as subject) would be indistinguishable from tensed VPs. Metarule (39) claims that these constituents, which look just like tensed VPs, are exactly that.¹⁷ This should become clearer if we give some examples of rules which meet the input conditions of (39) and then show what the rules created from them by (39) look like. The grammar will contain the following rules, all of which are input to (39):

- (40) a.
$$\begin{bmatrix} Q & NP & Q/NP \\ [+WH] \end{bmatrix}$$
 (root constituent questions, (35) above)
 b.
$$\begin{bmatrix} Q & NP & S/NP \\ [+WH] \end{bmatrix}$$
 (embedded constituent questions)

¹⁶ Semantically, we must substitute $(VP'(^h_{NP}))$ for $(\Sigma/NP)'$ in the translation of the original rule.

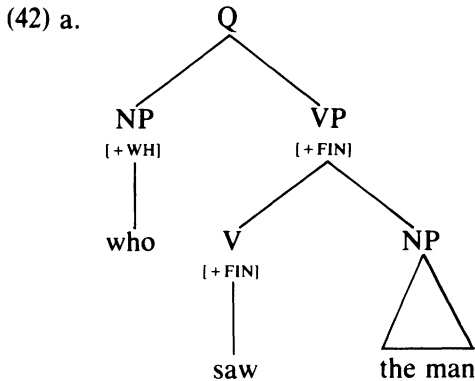
¹⁷ This makes sense from a language processing point of view; when the parser encounters a constituent that looks like a tensed VP, it can decide then and there that that is what it is, without needing to allow for the possibility of subsequently having to reanalyze it as a sentence with a missing subject.

- c. $\begin{bmatrix} \text{R} & \text{NP} & \text{S/NP} \\ & \begin{bmatrix} \pm \text{WH} \\ + \text{PRO} \end{bmatrix} & \end{bmatrix}$ (relative clauses, (25) above)
- d. $\begin{bmatrix} \text{S} & \text{NP} & \text{S/NP} \end{bmatrix}$ (NP topicalization)
- e. $\begin{bmatrix} \text{VP/NP} & \text{V} & \tilde{\text{S/NP}} \end{bmatrix}$
- f. $\begin{bmatrix} \text{VP/NP} & \text{V} & \text{NP} & \tilde{\text{S/NP}} \end{bmatrix}$
- g. $\begin{bmatrix} \text{VP/NP} & \text{V} & \text{PP} & \tilde{\text{S/NP}} \end{bmatrix}$ } (derived rules)

Metarule (39) applies to (40a–g) to give us (41a–g), respectively:

- (41) a. $\begin{bmatrix} \text{Q} & \text{NP} & \text{VP} \\ & \begin{bmatrix} + \text{WH} \end{bmatrix} \begin{bmatrix} + \text{FIN} \end{bmatrix} & \end{bmatrix}$ (= (38))
- b. $\begin{bmatrix} \tilde{\text{Q}} & \text{NP} & \text{VP} \\ & \begin{bmatrix} + \text{WH} \end{bmatrix} \begin{bmatrix} + \text{FIN} \end{bmatrix} & \end{bmatrix}$
- c. $\begin{bmatrix} \text{R} & \text{NP} & \text{VP} \\ & \begin{bmatrix} \pm \text{WH} \\ + \text{PRO} \end{bmatrix} \begin{bmatrix} + \text{FIN} \end{bmatrix} & \end{bmatrix}$
- d.¹⁸ $\begin{bmatrix} \text{S} & \text{NP} & \text{VP} \\ & \begin{bmatrix} + \text{FIN} \end{bmatrix} & \end{bmatrix}$
- e. $\begin{bmatrix} \text{VP/NP} & \text{V} & \text{NP} & \text{VP} \\ & & \begin{bmatrix} + \text{FIN} \end{bmatrix} & \end{bmatrix}$
- f. $\begin{bmatrix} \text{VP/NP} & \text{V} & \text{NP} & \text{VP} \\ & & \begin{bmatrix} + \text{FIN} \end{bmatrix} & \end{bmatrix}$
- g. $\begin{bmatrix} \text{VP/NP} & \text{V} & \text{PP} & \text{VP} \\ & & \begin{bmatrix} + \text{FIN} \end{bmatrix} & \end{bmatrix}$

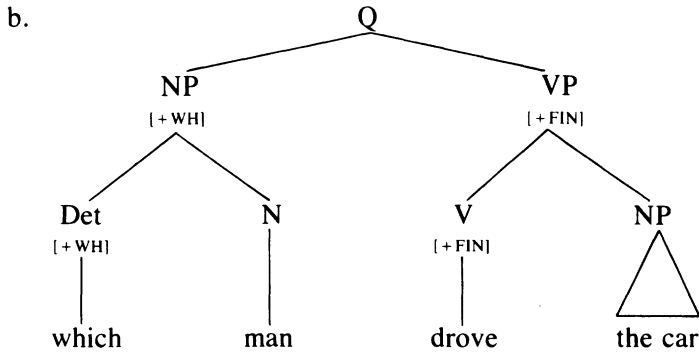
The output root constituent question rule (41a) now allows us to generate the auxiliary-less questions in (37):



¹⁸ As Geoff Pullum has pointed out to me, most current theories commit their proponents to the claim that all sentences like (i) are structurally, though not semantically, ambiguous between the structures shown approximately in (ii) and (iii):

- (i) Kim loves Sandy.
 (ii) $\begin{bmatrix} \text{S}[\text{NP Kim}] & \text{VP loves Sandy} \end{bmatrix}$
 (iii) $\begin{bmatrix} \text{S}[\text{NP Kim}] & \text{S}[\text{NP } t] & \text{VP loves Sandy} \end{bmatrix}$

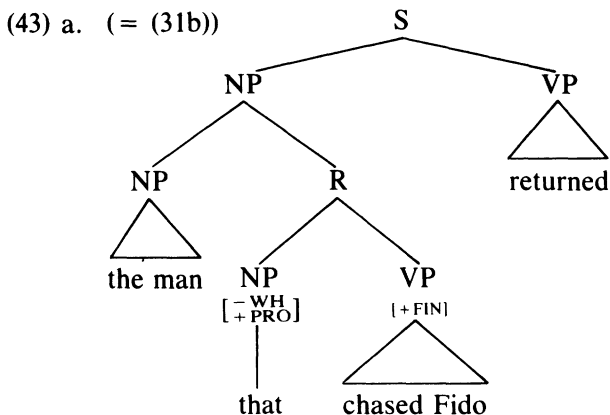
The present analysis only allows such sentences to have structure (ii), since the GLBC eliminates structure (iii).

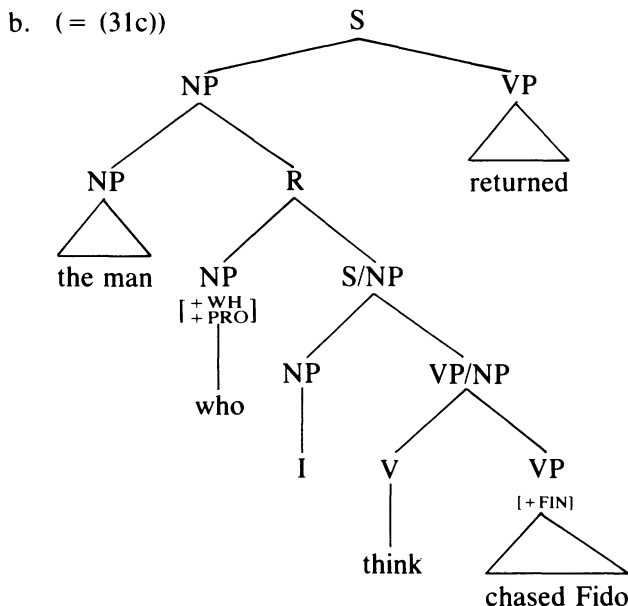


Notice that in the output relative clause rule (41c), the head NP is not optional, as it was in (25), because the metarule requires the X variable to be nonempty in order to be applicable. This means that a relative clause cannot be realized as a bare tensed VP, and we are therefore still unable to generate (31a), repeated here:

(31) a. *The man chased Fido returned.

But we can now generate examples (31b) and (31c), which the GLBC earlier prevented us from generating:





The metarule will not apply to the rule which expands \bar{S} , since the X variable must contain a major category symbol.¹⁹ Therefore, we will not have a rule of the form shown in (44):

$$(44) \quad * [\bar{S}/NP \text{ that } VP]_{[+FIN]}$$

Consequently, our grammar still cannot generate (31d), as we would wish:

(31) d. *The man (who) I think that chased Fido returned.

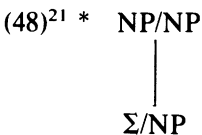
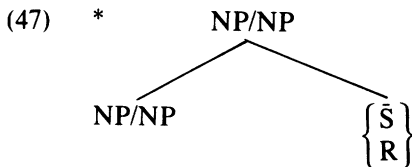
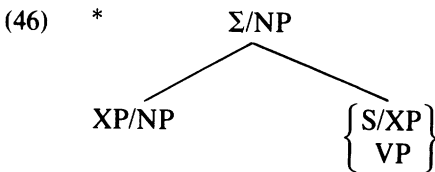
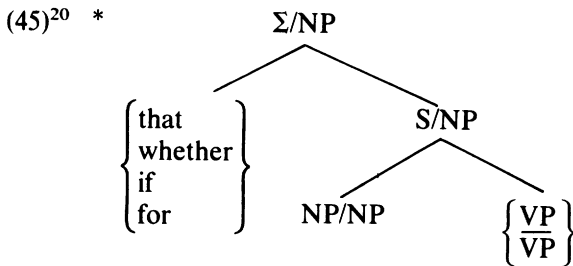
So far we have seen that (39) has two main effects: (i) it allows us to generate all the acceptable examples that the GLBC appeared to block, while not allowing us to generate any of the unacceptable examples that are rightfully blocked by the GLBC. Thus, we can maintain the GLBC. And (ii) it provides us with an NP VP analysis of matrix subject relatives and matrix subject constituent questions. We will see in the next section that the GLBC and the metarule given in (39) interact with the coordination

¹⁹ The use of the familiar “ \bar{S} ” is slightly misleading here, since we are using the symbol to abbreviate $[\pm C, -R, -Q]$ and it is only $[\pm C, -R, -Q]$ which expands as *that-S*. In this connection, notice that (39) will not apply to derived rules which introduce $[\pm C, -R, -Q]/NP$. Consequently, verbs like *regret* which require a $[\pm C, -R, -Q]$ complement *cannot* be followed by a tensed VP, in contrast to verbs like *think* which take $[\pm C, -R, -Q]$. Thus, our analysis predicts the following contrasts:

- (i) a. Who do you think (that) you saw?
- b. Who do you think (*that) saw you?
- (ii) a. Who do you regret *(that) you saw?
- b. *Who do you regret (that) saw you?

schema (9) to make a number of surprising but correct predictions about possible ATB dependencies (see especially the examples in (60)–(61), below).

The GLBC predicts the unacceptability of all the tree configurations shown in (45)–(48):



Thus, the grammar will be unable to generate any of the strings shown in (49):

- | | |
|---|------------------|
| (49) a. *Who did you believe that <i>t</i> came? | (structure (45)) |
| b. *Who did you wonder whether <i>t</i> came? | (structure (45)) |
| c. *Who did you wonder if <i>t</i> came? | (structure (45)) |
| d. *Who did you arrange for <i>t</i> to come? | (structure (45)) |
| e. *Which table did you wonder on <i>t</i> Kim put the book? | (structure (46)) |
| f. *Which did you buy the table on <i>t</i> Kim put the book? | (structure (46)) |

²⁰ For the classic discussion of this configuration, see Bresnan (1977, 170–183).

²¹ I am grateful to Geoff Pullum for drawing to my attention the fact that configuration (48) is blocked by the GLBC, and that consequently facts motivating Ross's (1967) Sentential Subject Constraint follow from the GLBC without modification. He also points out that (48) is not restricted to subjects, thus explaining the anomaly of (i), which is outside the scope of Ross's constraint.

(i) *What do you believe that iron is *t* to be a fact well known to virtually everybody?

- g. *Who did you wonder *t* saw Kim? (structure (46))
- h. *Which did you buy the table *t* supported the book? (structure (46))
- i. *The fact, I put it down to *t* that Kim came. (structure (47))
- j. *The table, I put Kim on *t* which supported the book. (structure (47))
- k. *The table, that I put Kim on *t* surprised Kim. (structure (48))
- l. *The exam, whether Kim passes *t* matters to Sandy. (structure (48))

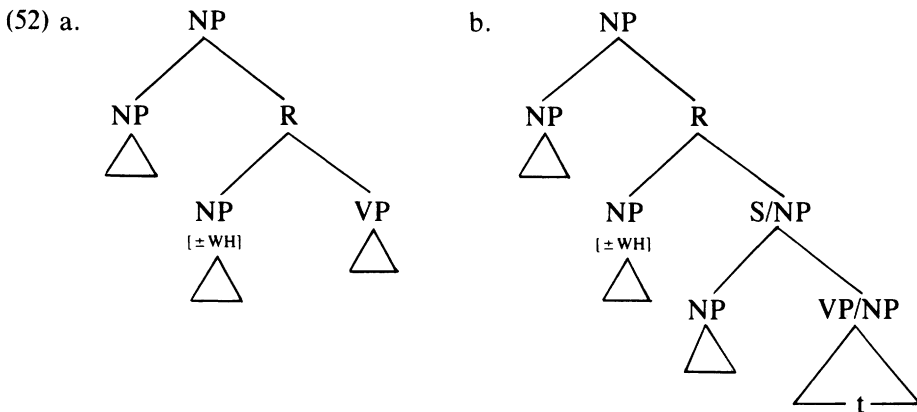
Notice that if *it*-clefts on NP are introduced by a rule like (50),

(50) [_{VP} V NP R]

then the GLBC will not block “extraction” of the NP introduced by the rule, since it is not on a left branch. Predictably, then, the following example (which I owe to Stan Peters) is acceptable:

(51) Who is it that Mary likes?

The proposals made here, which have so far been motivated on purely syntactic grounds, commit us to the claim that relatives and interrogatives with a dependency into the matrix subject argument have a rather different structure from all other relatives and interrogatives. This is illustrated in (52) and (53). Under our analysis, matrix subject relatives will have the structure shown in (52a), whereas other relatives will have the one shown in (52b):²²

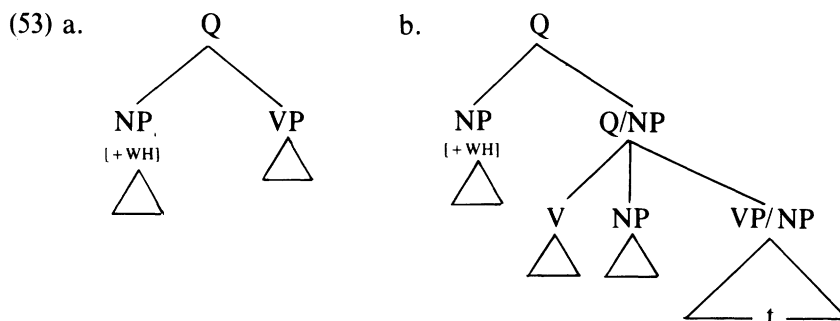


²² There is evidence to suggest that (52a) is the correct structure for matrix subject relatives in French also. I owe the following examples to Lisa Selkirk:

- (i) a. L'homme qui est parti . . .
the man who has left
- b. L'homme qu' est parti . . .
the man who has left

She notes that one would expect (ib) to be ungrammatical if a trace intervened between the relative pronoun and the copula. On the VP analysis, there is no trace there to block the contraction. And Elisabet Engdahl reports that in Swedish relatives the matrix subject can never be realized as a resumptive pronoun, although resumptive pronouns can occur in other relativized NP positions. Again, this would be predicted by a VP analysis.

Likewise, matrix subject questions will have the structure (53a), whereas other constituent questions will have structure (53b):



A derived category has to contain twice as much syntactic information as a basic category. This suggests that structures which involve derived categories will impose a heavier processing load than those that do not. If this is so, then our analysis predicts that matrix subject relatives and questions will be significantly easier to process than all other relatives and questions (excluding polar interrogatives). This prediction is borne out, both developmentally and for adult speakers, by a substantial body of recent psycholinguistic work.²³ Furthermore, that (52a) is a possible structure for matrix subject relatives makes it much less surprising that all languages allow the construction of such relatives (Keenan and Comrie (1977)), even those which have no mechanism inducing unbounded dependencies.

3. Constraints on Coordinate Structure

A consequence of the analysis of unbounded phenomena that we have put forward is that subtrees that contain an externally controlled hole are of a different syntactic category from those that do not. An \bar{S} which has an NP missing somewhere will be of category \bar{S}/NP , not \bar{S} , and a VP which contains a missing PP will be of category VP/PP . Now according to the coordination schema in (9), only items of the *same* syntactic category can be conjoined.²⁴ It follows that, while we can conjoin a VP with a VP, for

²³ Questions (i) adults: Read, Kraak, and Boves (1980) on Dutch; (ii) children: Stewart and Sinclair (1975), Tyack and Ingram (1977). Relative clauses (i) adults: Fodor, Bever, and Garrett (1974, 364), Sheldon (1976; 1977), Wanner and Maratsos (1978), Holmes (1979), Mills (1979) on German, Frauenfelder, Segui, and Mehler (1979) on French; (ii) children: Brown (1971), Smith (1974), Ferreiro et al. (1976), and de Villiers et al. (1979). I am grateful to Anne Cutler and Paloma Garcia-Bellido for drawing my attention to some of this literature.

²⁴ As Tom Wasow and one of my referees have reminded me, there are apparent exceptions to this claim such as *slowly and with great care* and *longwinded and a bully*. Such expressions are permitted to appear in some slot in a sentence only when either conjunct could appear alone in that slot. Thus, we find (ia–c) but not (ii) or (iii):

- (i) a. He was longwinded.
- b. He was a bully.
- c. He was longwinded and a bully.
- (ii) *The [longwinded and a bully] man entered.
- (iii) *[Longwinded and a bully] entered.

Beyond this observation, I have nothing to say about such examples.

example, we cannot conjoin a VP with a VP/NP or a VP/PP any more than we can conjoin it with an \bar{S} . But we can conjoin an \bar{S} /NP with an \bar{S} /NP, or a VP/PP with a VP/PP. Thus, the effect of (9) in place of a rule of CR is to entirely eliminate the need for Ross's (1967) Coordinate Structure Constraint (CSC) or Williams's (1977; 1978) Across-the-Board (ATB) Convention. Given (9), we simply cannot generate strings like (54) and (55):

(54) *John is easy to please and to love Mary. (= $\overline{VP/NP}$ & \overline{VP})

(55) *The man who Mary loves and Sally hates George computed my tax.
(= S/NP & S)

Permissibility of coordination has traditionally been taken as evidence for sameness of syntactic category, but no one has previously drawn the relevant inference from Ross's CSC facts, assuming instead a quite different explanation for the impossibility of coordination in these cases.²⁵

The ATB exceptions to the CSC are predicted by (9):

(56) John is easy to please and to love. (= $\overline{VP/NP}$ & $\overline{VP/NP}$)

(57) The man who Mary loves and Sally hates computed my tax.
(= S/NP & S/NP)

But we will not get ATB dependencies when the holes are of different categories:

(58) a. The kennel which Mary made and Fido sleeps in has been stolen.
(= S/NP & S/NP)

b. The kennel in which Mary keeps drugs and Fido sleeps has been stolen.
(= S/PP & S/PP)

c. *The kennel (in) which Mary made and Fido sleeps has been stolen.
(= S/NP & S/PP)

The same is true in the following comparative examples from Williams (1977, 421):

(59) a. John saw more horses than Bill saw or Pete talked to. (= S/NP & S/NP)

b. John saw more horses than Bill saw cows or Pete talked to cats.
(= S/QP & S/QP)

c. *John saw more horses than Bill saw cows or Pete talked to.
(= S/QP & S/NP)

As Williams points out (1977, 421), a theory employing a rule of CR and the CSC "will be hard pressed to avoid generating [59c]." In fact, Williams's own (1978) rule of CR avoids generating (59c) only because of an otherwise unmotivated constraint which prevents CR from "eating into Ss" (1978, 38).

Williams (1978) draws attention to a curious fact about the coordination of relative clauses: a relative with a matrix subject NP dependency cannot be conjoined with any

²⁵ Schachter (1977) almost draws the relevant inference but ends up explaining some CSC violations syntactically, others semantically.

other kind of relative, whereas a relative with an embedded subject NP dependency behaves quite normally with respect to ATB coordination. The relevant data are as follows ((60c) = (16) from Williams (1978, 34), and I owe (60d) to Paul Hirschbühler):²⁶

- (60) a. I know a man who Bill saw and Mary liked.
 (= S/NP & S/NP)
- b. I know a man who saw Bill and liked Mary.
 (= VP & VP)
- c. *I know a man who Bill saw and liked Mary.
 (= { S/NP & VP
 VP/NP & VP })
- d. I know a man who Mary likes and hopes will win.
 (= VP/NP & VP/NP)

This acceptability pattern is exactly as predicted by the VP analysis of subjectless relatives proposed in the last section. Notice that just the same pattern obtains for indirect questions, as our Σ /NP-to-VP metarule would lead us to expect:

- (61) a. I wonder who Bill saw and Mary liked.
 (= S/NP & S/NP)
- b. I wonder who saw Bill and liked Mary.
 (= VP & VP)
- c. *I wonder who Bill saw and liked Mary.
 (= S/NP & VP)
- d. I wonder who Mary likes and hopes will win.
 (= VP/NP & VP/NP)

The grammaticality distribution in (60) and (61) cannot be explained by a theory which requires all the holes in an ATB construction to have the same case: such a theory incorrectly predicts that (60d) and (61d) will be ungrammatical.²⁷

It is worth pointing out that the treatment of the interaction of unbounded dependencies and coordination detailed in this article makes exactly the right predictions about what Grimshaw (1978) refers to as “complementizer conjoined *wh* phrases”. Consider her examples:

- (62) *John asked who and where Bill had seen.
 [= Grimshaw’s (4b)]
- (63) *John asked who and what bought.
 [= (6)]

²⁶ Paloma Garcia-Bellido informs me that the counterpart of (60c) is grammatical in Spanish. This would be predictable if, as is arguably the case, the GLBC does not apply in Spanish: (60c) could then have the structure S/NP & S/NP.

²⁷ This observation is due to Paul Hirschbühler and Andrew Radford.

- (64) Which book and which pencil did John buy? [= (8a)]
- (65) *Where and when did Bill put the book? [= (9a)]
- (66) On which table and under which flower pot did John put the keys? [= (11)]
- (67) a. To which city and to which conference did Bill go? [= (12)]
- b. To which city and which conference did Bill go? [= (15a)]
- c. Which city and which conference did Bill go to? [= (15b)]
- d. *Which city and which conference did Bill go to ____ to ____ ? [= (14a)]
- e. *Which city and to which conference did Bill go to? [= (14b)]
- f. *To which city and which conference did Bill go to? [= (14c)]

In every case Grimshaw's explanation for the (un)grammaticality of an example carries over to the present analysis. Examples (62) and (67e) are ruled out because they involve coordination of unlike categories; (63) because *who* and *what* cannot be both subject and object of *bought* and yet there are both subject and object holes in the sentence; (65) because *put* subcategorizes for a locative adverbial (hence, the *wh*-phrase must also be locative, but *where* and *when* is not); (67d) because one (coordinate) *wh*-NP cannot bear a dependency relation to two (noncoordinate) holes; and (67f) because a *wh*-PP cannot bear a dependency relation to an NP hole. Thus, no extra apparatus is needed to handle conjoined *wh*-phrases: the pattern of acceptability exhibited in (62)–(67) simply follows from the coordination and dependency schemata given as (9) and (15) taken together with uncontroversial assumptions about subcategorization.

However, the approach advocated here cannot, at present, explain the contrast between (68a) and (68b) noted by Williams (1978, 35), since both will be generated:

- (68) a. *John, who and whose friends you saw, is a fool. (= NP & NP)
- b. John, to who and to whose friends that letter was addressed, is a fool.
[= PP & PP]

Williams's own explanation for this contrast depends crucially on what he refers to as a "quite particular" asymmetric redefinition of *factor* (1978, 32) and an equally particular assumption about the structural description for the rule of *Wh* Movement (1978, 35). There are two problems with his account, one theoretical and the other empirical. The theoretical problem is this: if the rule of *Wh* Movement looks like (69), then Williams's "quite particular" redefinition of *factor* will not do any work:

| | | | | | | |
|------|-----|---|------|---|-----------|---|
| (69) | SD: | W | COMP | Y | \bar{X} | Z |
| | | | | | [+WH] | |
| | | 1 | 2 | 3 | 4 | 5 |
| | SC: | 1 | 4 | 3 | ϕ | 5 |

So, instead of the rather simple and elegant (69), Williams must state *Wh* Movement in a way which makes essential reference to tree structure in the structural description, contrary to the definition of a transformation as an operation on an unbracketed string of category symbols as in classical transformational grammar.

The empirical problem is that Williams's redefinition of *factor* predicts that the following English sentences will be ungrammatical:²⁸

- (70) a. I wonder when and how often she went that day.
 b. I wonder who and whose friends he handed over to the FBI.

Since both sentences are entirely grammatical, Williams would be forced to modify his definition of *factor* to accommodate them. But any such modification will inevitably allow (68a) to be generated; thus, in the end Williams's analysis is also unable to explain the contrast between (68a) and (68b). These examples are therefore irrelevant to choosing between his proposals and those made here.²⁹

4. Rightward Dependencies

As is well known, rightward displacement of constituents is subject to a constraint having to do with "heaviness": roughly speaking, the heavier the displaced constituent, the better the sentence sounds. This fact is hard to capture in the formalism of a generative grammar (whether phrase structure or transformational) and it seems reasonable, and probably not controversial, to suppose that it may be a fact that ought to be captured not in that formalism, but rather in one's model of language perception and/or production.³⁰

Another familiar fact about rightward displacements is their apparent clause-boundness. This was first noted by Ross (1967, 166), and the stipulation he made against unbounded rightward movement has come to be known as the "Right Roof Constraint" (RRC, hereafter). Subsequent work has shown that the RRC is neither universal nor absolute. Languages as diverse as Circassian, German, Hindi, and Navajo have been alleged to admit constructions which appear to violate it.³¹ And even in English, the facts are not exactly as the RRC predicts. Thus, Postal (1974, 92n) cites Witten (1972, IV-93) as the source for the examples in (71):

²⁸ I am indebted to Geoff Pullum for the examples in (70).

²⁹ Edwin Williams has suggested to me that the difference in acceptability between (70b) and (68a) may result from the fact that relatives require the complement clause to characterize a unique individual whereas interrogatives do not. If conjoined indefinite NPs like *who* and *whose friends* are necessarily disjoint in reference, then this would explain why they do not show up as relative heads.

³⁰ Cf. Ross (1967, chapter 3), Fiengo (1977, 48-49).

³¹ Circassian: Colarusso (1976); German: Kohrt (1975); Hindi: Satyanarayana and Subbarao (1973); Navajo: Kaufman (1974), Perkins (1975).

- (71) a. I have wanted to know exactly what happened to Rosa Luxemburg for many years.
- b. I have wanted to know for many years exactly what happened to Rosa Luxemburg.

Example (71b) is clearly grammatical and equally clearly in violation of the RRC. Consider also the following examples, suggested to me by Janet Fodor:

- (72) a. I had hoped that it was true that Rosa Luxemburg had actually defected to Iceland for many years.
- b. I had hoped that it was true for many years that Rosa Luxemburg had actually defected to Iceland.
- (73) a. I have wanted to meet the man who spent so much money planning the assassination of Kennedy for many years.
- b. I have wanted to meet for many years the man who spent so much money planning the assassination of Kennedy.

Again, (72b) and (73b) are clear violations of the RRC. In fact, as Andrews (1975, 112) points out, all the starred sentences that Ross originally used to motivate the RRC were, in any case, excluded by the Sentential Subject Constraint. Andrews goes on to show that when the Sentential Subject Constraint does not interfere, the resulting RRC-violating strings are much more acceptable than those that Ross originally listed.³²

In the light of these observations, I want to suggest, following Grosu (1972), that the RRC is not part of the grammar at all, and instead that the facts it purports to explain are to be better explained in terms of performance considerations, in particular parsing strategies. Frazier (1979) and Frazier and Fodor (1978) have recently argued at length for a model of the natural language parser which attaches incoming material as low as it possibly can on the parse tree. This theory offers a natural explanation for the fact that the preferred reading of (74a) is (74b) rather than (74c):

- (74) a. The woman believed that the man was ill who was here.
- b. The woman believed that the man who was here was ill.
- c. The woman who was here believed that the man was ill.

It also explains why Perlmutter's example (75a) (from Grosu (1971, 423)) is interpreted as being synonymous with (75b) rather than (75c):

- (75) a. A woman hit a girl who was pregnant.

³² Andrews exhibits the following "rather massive RRC violations" (1975, 234):

- (i) a. People are said to do crazier things at higher speeds there by Dorothy than they are by other people.
- b. People are said to do such crazy things at such high speeds there by Dorothy that I am getting skeptical.

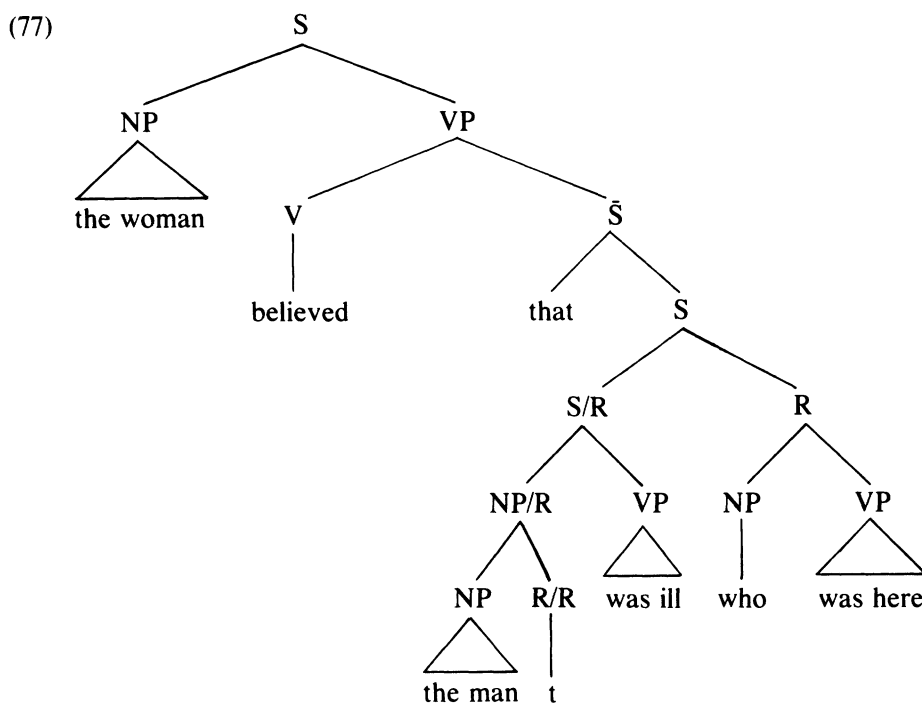
As Andrews remarks, these examples "strengthen the suspicion . . . that the RRC should be retired" (1975, 234).

- b. A woman hit a pregnant girl.
- c. A pregnant woman hit a girl.

If the boundedness of rightward dependencies is a by-product of parser operation, then it need not be built into the syntactic rules which permit such dependencies. Accordingly, we may propose the following very general schema for rightward displacement,³³

$$(76) \langle 9, [\alpha \alpha/\beta \beta], \lambda h_{\beta}[(\alpha/\beta)'](\beta') \rangle$$

where α ranges over clausal categories and β can be any phrasal or clausal category. This schema will then give us extraposed relatives and will, for example, assign the following surface structure to (74a):



Extrapolation of relative clauses *cannot* in general be the result of a movement or deletion rule, as the following examples (based on Perlmutter and Ross (1970)) clearly demonstrate:

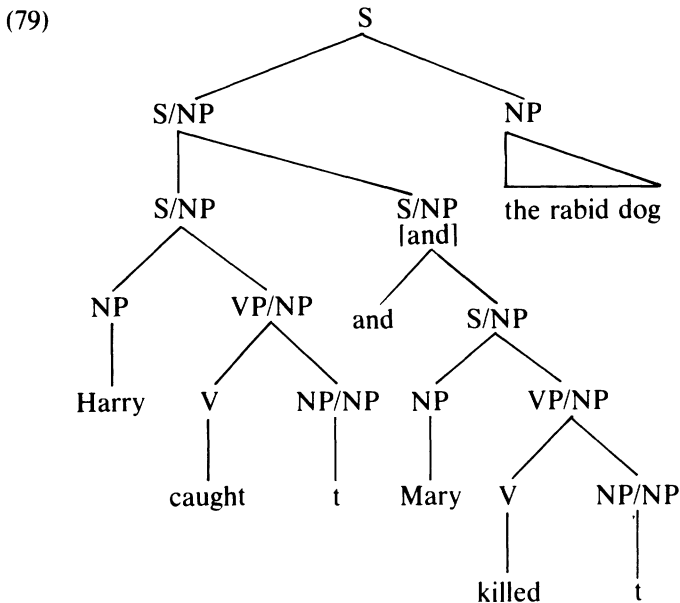
- (78) a. A man just came in and a woman went out who were similar in all kinds of ways.

³³ As Emmon Bach has pointed out to me, the semantic rule in (76) is too simplistic to handle the examples in (78) and (82), below, correctly. However, in the absence of any serious proposals for the semantics of symmetric predicates, it is not obvious what to replace it with.

- b. A man just came in and a woman went out who hate each other like poison and always have.

These sentences show that extraposed relatives must be generated in situ, exactly as (76) proposes.

Extraposed relatives are not all that (76) will account for. Most interestingly, it interacts with the coordination schema given in (9) to achieve all the effects of the Right Node Raising (RNR) transformation without the need for additional machinery. This is a remarkable result for, as Jackendoff has pointed out, there are “no remotely coherent formalizations of RNR” (1977, 193). But we have an entirely coherent formalization that makes RNR otiose. The following tree makes it clear how this comes about:



RNR is not bound by the RRC. This fact is puzzling when the former is a rightward movement rule and the latter a constraint on rightward movement rules. However, far from being puzzling, it is actually predicted by the present analysis. The Frazier–Fodor parser requires incoming material to be attached in the clause being parsed *when possible*. But the effect of the RNR construction is to make it impossible to attach the material following the hole in the clause which precedes the hole. For example, consider tree (79) above. When the parser has attached *caught* to the first VP, it is looking for the NP that *catch* subcategorizes for: but the incoming material (*and*, *and Mary*, *and Mary killed*, . . .) cannot be this NP. So, of necessity, the parser closes the VP and begins the analysis of a new clause.

RNR is often used as a test for constituenthood. Schema (76) predicts that only constituents can appear in the rightmost β position. Thus, the following strings cannot be generated:³⁴

- (80) a. *I find it easy to believe—but Joan finds it hard to believe—Tom to be dishonest.
- b. *John offered, and Harry gave, Sally a Cadillac.
- c. *John told, and Harry showed, Seymour that Sally was a virgin.

But all the sentences in (81) will be:³⁵

- (81) a. Jack may be *t* and Tony certainly is *t* a werewolf.
- b. Tom said he would *t* and Bill actually did *t* eat a raw eggplant.
- c. Tony used to be *t* and George still is *t* very suspicious.
- d. Harry has claimed *t* but I do not believe *t* that Melvin is a Communist.
- e. I like *t* but Tom doesn't like *t* to visit new places.
- f. I can tell you when *t*, but I can't tell you why *t*, he left me.
- g. I've been wondering whether *t*, but wouldn't positively want to state that *t*, your theory is correct.

RNR cannot be a movement or deletion rule, as the following examples clearly demonstrate:³⁶

- (82) a. John hummed *t*, and Mary sang *t*, the same tune.
- b. John hummed *t*, and Mary sang *t*, at equal volumes.
- c. John gave Mary *t*, and Joan presented *t* to Fred, books which looked remarkably similar.
- d. The Red Sox beat *t*, and the Giants were beaten by *t*, different teams.

These examples require no extra syntactic apparatus or special devices. On the present approach, they are as straightforward as the examples in (81). But they pose an insuperable problem for conventional analyses.

5. Conclusion

This article has made five main proposals:

- I. A single, very general, schema for constituent coordination.
- II. A single, very general, schema allowing unbounded dependencies.
- III. A generalization of Ross's Left Branch Constraint.

³⁴ Example (80a) is from Postal (1974, 128), and (80b,c) are due to Hankamer (1971, 76, cited by Abbott (1976, 639)). The problematic case for the present analysis, as with every known alternative analysis of RNR, is Abbott's (1976, 639) example:

(i) Smith loaned, and his widow later donated, a valuable collection of manuscripts to the library.

³⁵ Examples (81a–e) are from Postal (1974, 126–128) and (81f–g) are from Bresnan (1974, 618).

³⁶ Example (82a) is from Jackendoff (1977, 192) and (82c) is due to Abbott (1976, 642).

- IV. A tensed VP analysis of apparent subject dependencies.
- V. A single, very general, schema allowing rightward dependencies.

Each proposal has been independently motivated. However, a wide variety of facts follow as theorems from the interaction of the various proposals. Thus, the distribution of complementizer *that* follows from II, III, and IV; the “complementizer conjoined *wh*-phrase” facts and the coordinate structure constraint facts follow from I and II; the across-the-board facts follow from I, II, III, and IV; and the “Right Node Raising” facts follow from I, II, and V. These facts are explained in a grammatical framework that does not employ any rules of NP Movement, *Wh* Movement, or Coordination Reduction. Indeed, the explanations given are incompatible with the existence of such rules. Thus, the assumption made at the beginning to the effect that there are no transformations turns out to be not just *tenable*, but in fact *necessary*, if we are to achieve an explanatory, rather than merely stipulative, account of the data discussed in this article.

References

- Abbott, B. (1976) “Right Node Raising as a Test for Constituents,” *Linguistic Inquiry* 7, 639–642.
- Andrews, A. D. (1975) *Studies in the Syntax of Relative and Comparative Clauses*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Bach, E. (1976) “An Extension of Classical Transformational Grammar,” mimeo, University of Massachusetts at Amherst.
- Bach, E. and R. Cooper (1978) “The NP-S Analysis of Relative Clauses and Compositional Semantics,” *Linguistics and Philosophy* 2, 145–150.
- Baker, C. L. (1978) *Introduction to Generative-Transformational Syntax*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Bever, T. G., J. M. Carroll, and R. Hurtig (1976) “Analogy or Ungrammatical Sequences That Are Utterable and Comprehensible Are the Origins of New Grammars in Language Acquisition and Linguistic Evolution,” in T. G. Bever, J. J. Katz, and D. T. Langendoen, eds., *An Integrated Theory of Linguistic Ability*, Thomas Crowell, New York, 149–182.
- Bresnan, J. W. (1974) “The Position of Certain Clause-Particles in Phrase Structure,” *Linguistic Inquiry* 5, 614–619.
- Bresnan, J. W. (1976) “Evidence for a Theory of Unbounded Transformations,” *Linguistic Analysis* 2, 353–393.
- Bresnan, J. W. (1977) “Variables in the Theory of Transformations,” in P. W. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax*, Academic Press, New York, 157–196.
- Brown, H. D. (1971) “Children’s Comprehension of Relativized English Sentences,” *Child Development* 42, 1923–1936.
- Colarusso, J. (1976) “An Instance of Unbounded Rightward Movement: Wh-Movement in Circassian,” unpublished paper, University of Vienna.
- Cooper, R. (1975) *Montague’s Semantic Theory and Transformational Syntax*, Doctoral dissertation, University of Massachusetts, Amherst, Massachusetts.
- Cooper, R. (forthcoming) *Formal Semantics and Transformational Grammar*, Reidel, Dordrecht.
- Dougherty, R. C. (1970) “A Grammar of Coordinate Conjoined Structures I,” *Language* 46, 850–898.

- Earley, J. (1970) "An Efficient Context-free Parsing Algorithm," *Communications of the ACM* 13, 94–102.
- Engdahl, E. (1980) *The Syntax and Semantics of Questions in Swedish*, Doctoral dissertation, University of Massachusetts at Amherst.
- Ferreiro, E., C. Othenin-Girard, H. Chipman, and H. Sinclair (1976) "How Do Children Handle Relative Clauses?" *Archives de Psychologie* 45, 229–266.
- Fiengo, R. (1977) "On Trace Theory," *Linguistic Inquiry* 8, 35–61.
- Fodor, J. A., T. G. Bever, and M. Garrett (1974) *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*, McGraw-Hill, New York.
- Frauenfelder, U., J. Segui, and J. Mehler (1979) "Monitoring around the Relative Clause," unpublished paper.
- Frazier, L. (1979) "On Comprehending Sentences: Syntactic Parsing Strategies," mimeo, Indiana University Linguistics Club, Bloomington, Indiana.
- Frazier, L. and J. D. Fodor (1978) "The Sausage Machine: A New Two-Stage Parsing Model," *Cognition* 6, 291–325.
- Gazdar, G. J. M. (1980a) "A Cross-categorical Semantics for Coordination," *Linguistics and Philosophy* 3, 407–409.
- Gazdar, G. J. M. (1980b) "A Phrase Structure Syntax for Comparative Clauses," in T. Hoekstra, H. v. d. Hulst, and M. Moortgat, eds., *Lexical Grammar*, Foris Publications, Dordrecht, 165–179.
- Gazdar, G. J. M. (forthcoming) "Phrase Structure Grammar," to appear in P. Jacobson and G. K. Pullum, eds., *The Nature of Syntactic Representation*, Reidel, Boston.
- Gazdar, G. J. M., G. K. Pullum, and I. Sag (1980) "A Phrase Structure Grammar of the English Auxiliary System," in I. Sag, ed., *Stanford Working Papers in Grammatical Theory*, Volume I, Stanford Cognitive Science Group, 1–124.
- Grimshaw, J. (1978) "On the Syntax of Conjoined *wh* Words in English," in R. M. Saenz, ed., *University of Massachusetts Occasional Papers in Linguistics* Volume 3, Graduate Linguistic Student Association, University of Massachusetts at Amherst, 1–10.
- Grosu, A. (1971) "On Perceptual and Grammatical Constraints," in D. Adams et al., eds., *Papers from the Seventh Regional Meeting of the Chicago Linguistic Society*, University of Chicago, Chicago, Illinois, 416–427.
- Grosu, A. (1972) "The Strategic Content of Island Constraints," *Ohio State University Working Papers in Linguistics* 13.
- Hankamer, J. (1971) *Constraints on Deletion in Syntax*, Doctoral dissertation, Yale University, New Haven, Connecticut.
- Harman, G. H. (1963) "Generative Grammars without Transformation Rules: A Defense of Phrase Structure," *Language* 39, 597–616.
- Hellan, L. (1977) "X-syntax, Categorical Syntax and Logical Form," in T. Fretheim and L. Hellan, eds., *Papers from the Trondheim Syntax Symposium*. Trondheim, 83–135.
- Holmes, V. M. (1979) "Some Hypotheses about Syntactic Processing in Sentence Comprehension," in W. E. Cooper and E. Walker, eds., *Sentence Processing Studies Presented to Merrill Garrett*, Erlbaum, Hillsdale, New Jersey.
- Horn, G. M. (1974) *The NP Constraint*, Doctoral dissertation, University of Massachusetts at Amherst. Also distributed by the Indiana University Linguistics Club, Bloomington, Indiana.
- Jackendoff, R. (1977) *\bar{X} Syntax: A Study of Phrase Structure*, Linguistic Inquiry Monograph 2, MIT Press, Cambridge, Massachusetts.
- Karttunen, L. (1977) "Syntax and Semantics of Questions," *Linguistics and Philosophy* 1, 3–44.

- Kaufman, E. S. (1974) "Navajo Spatial Enclitics: A Case for Unbounded Rightward Movement," *Linguistic Inquiry* 5, 507-533.
- Keenan, E. L. and B. Comrie (1977) "Noun Phrase Accessibility and Universal Grammar," *Linguistic Inquiry* 8, 63-99.
- Keenan, E. L. and L. Faltz (1978) "Logical Types for Natural Language," *UCLA Occasional Papers in Linguistics* 3.
- Klein, E. H. (1980) "Determiners and the Category Q," unpublished paper, University of Sussex.
- Kohrt, M. (1975) "A Note on Bounding," *Linguistic Inquiry* 6, 167-171.
- Langendoen, D. T. (1976) "On the Weak Generative Capacity of Infinite Grammars," *CUNY Forum* 1, 13-24.
- Levy, L. S. (forthcoming) "Automata on Trees: A Tutorial Survey," to appear in *Egyptian Computer Journal*.
- Maling, J. and A. Zaenen (forthcoming) "A Base-generated Account of 'Extraction Phenomena' in Scandinavian Languages," to appear in P. Jacobson and G. K. Pullum, eds., *The Nature of Syntactic Representation*.
- Marslen-Wilson, W. D. (1973) *Speech Shadowing and Speech Perception*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
- McCawley, J. D. (1968) "Concerning the Base Component of a Transformational Grammar," *Foundations of Language* 4, 243-269. Reprinted in *Grammar and Meaning*, Academic Press, New York, 35-58.
- McCloskey, J. (1978) "Questions and Relative Clauses in Modern Irish," *Texas Linguistic Forum* 12.
- McCloskey, J. (1979) *Transformational Syntax and Model Theoretic Semantics*, Reidel, Dordrecht.
- Mills, A. E. (1979) "Surface Structure Order and Comprehension of Ambiguous Clauses," paper presented to the Spring meeting of the Linguistics Association of Great Britain, Hull.
- Perkins, E. (1975) "Extraposition of Relative Clauses in Navajo," *The Navajo Language Review* 2.2 [cited by Andrews (1975, 114)].
- Perlmutter, D. M. and J. R. Ross (1970) "A Nonsource for Comparatives," *Linguistic Inquiry* 1, 127-128.
- Postal, P. M. (1974) *On Raising*, MIT Press, Cambridge, Massachusetts.
- Postal, P. M. and G. K. Pullum (1978) "Traces and the Description of English Complementizer Contraction," *Linguistic Inquiry* 9, 1-29.
- Read, C., A. Kraak, and L. Boves (1980) "The Interpretation of Ambiguous *Who*-Questions in Dutch: The Effect of Intonation," in W. Zonnefeld, ed., *Linguistics in the Netherlands, 1977-1979*, Foris Publications, Dordrecht.
- Roeper, T. and M. E. A. Siegel (1978) "A Lexical Transformation for Verbal Compounds," *Linguistic Inquiry* 9, 199-260.
- Ross, J. R. (1967) *Constraints on Variables in Syntax*, Doctoral dissertation, MIT, Cambridge, Massachusetts. Also distributed by the Indiana University Linguistics Club, Bloomington, Indiana.
- Satyanarayana, P. and K. V. Subbarao (1973) "Are Rightward Movement Rules Upward Bounded?" *Studies in the Linguistic Sciences* 3, 182-192.
- Schachter, P. (1976) "A Nontransformational Account of Gerundive Nominals in English," *Linguistic Inquiry* 7, 205-241.
- Schachter, P. (1977) "Constraints on Coordination," *Language* 53, 86-103.
- Sheldon, A. (1976) "Speakers' Intuitions about the Complexity of Relative Clauses in Japanese and English," in S. Mufwene, C. Walker, and S. Steever, eds., *Papers from the Twelfth*

- Regional Meeting of the Chicago Linguistic Society*, University of Chicago, Chicago, Illinois, 558–567.
- Sheldon, A. (1977) "On Strategies for Processing Relative Clauses: A Comparison of Children and Adults," *Journal of Psycholinguistic Research* 6, 305–318.
- Shir, N. E. (1977) "On the Nature of Island Constraints," mimeo, Indiana University Linguistics Club, Bloomington, Indiana.
- Smith, M. (1974) "Relative Clause Formation between 29–36 Months: A Preliminary Report," *Papers and Reports on Child Language Development* 8, 104–110.
- Stewart, J. and H. Sinclair (1975) "Comprehension of Questions by Children between 5 and 9," *International Journal of Psycholinguistics* 3, 17–26.
- Stump, G. T. (1978) "Interpretive Gapping in Montague Grammar," in D. Farkas, W. M. Jacobsen, and K. W. Todrys, eds., *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*, University of Chicago, Chicago, Illinois, 472–481.
- Thatcher, J. W. (1973) "Tree Automata: An Informal Survey," in A. V. Aho, ed., *Currents in the Theory of Computing*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Tyack, D. and D. Ingram (1977) "Children's Production and Comprehension of Questions," *Journal of Child Language* 4, 211–224.
- de Villiers, J. G., H. B. Tager Flusberg, K. Hakuta, and M. Cohen (1979) "Children's Comprehension of Relative Clauses," *Journal of Psycholinguistic Research* 8, 499–518.
- Vergnaud, J. R. (1973) "Formal Properties of Lexical Derivations," *Quarterly Progress Report of the Research Laboratory in Electronics* No. 108, MIT, 280–287.
- Wanner, E. and M. Maratsos (1978) "An ATN Approach to Comprehension," in M. Halle, J. Bresnan, and G. A. Miller, eds., *Linguistic Theory and Psychological Reality*, MIT Press, Cambridge, Massachusetts.
- Wijngaarden, A. van (1969) "Report on the Algorithmic Language ALGOL 68," *Numerische Mathematik* 14, 79–218.
- Williams, E. S. (1977) "Across-the-Board Application of Rules," *Linguistic Inquiry* 8, 419–423.
- Williams, E. S. (1978) "Across-the-Board Rule Application," *Linguistic Inquiry* 9, 31–43.
- Witten, E. (1972) "Centrality," in *Report No. NSF-28 to The National Science Foundation*, The Computation Laboratory of Harvard University, Cambridge, Massachusetts.
- Younger, D. H. (1967) "Recognition and Parsing of Context-free Languages in Time n^3 ," *Information and Control* 10, 189–208.

The University of Sussex
School of Social Sciences
Arts Building
Falmer
Brighton BN1 9QN
England