COORDINATE STRUCTURE AND UNBOUNDED DEPENDENCIES*

Gerald Gazdar, Ewan Klein,
Geoffrey K. Pullum, Ivan A. Sag

November 1982

Gerald Gazdar, University of Sussex

Ewan Klein, University of Newcastle

Geoffrey K. Pullum, University of California, Santa Cruz

Ivan A. Sag, Stanford University

* This paper depends on the content of Gazdar & Pullum (1982) ["GP82",henceforth] and is unlikely to be comprehensible to someone who hasn't read the latter. The acknowledgements noted in the preface to GP82 carry over to this paper, and will not be repeated here. Like GP82 this paper consists of draft material from a book we are preparing.

Cognitive Science Research Paper

Serial no: CSRP 006

The University of Sussex
Cognitive Studies Programme
School of Social Sciences
Falmer
Brighton BN1 9QN

X* A restriction ofi metarule application

Consider listed ID rules of the following form:

(1) K ~> CCi], ...

It is precisely rules liked), and no others, which determine the possibilities of subcategorization for lexical items. There is a clear sense in which these rules are the subcategorization frames for the language. Let us refer to this class of phrase structure rules as the "lexical ID rules". In terms of the feature system in GP82, a phrase structure rule is a lexical ID rule if and only if it requires a unique daughter to be C+LEXICALJ. We will refer to all phrase structure rules not meeting this condition as "nonlexical ID rules".

Earlier work in GPSG has assumed, implicitly or explicitly, that metarules are free to apply to any ordinary phrase structure rules in the grammar, provided their structural analysis is met. Most, and perhaps all, of the motivation for permitting this unrestricted application came either from facts about precedence possibilities, or from facts about the cooccurrence of features. In the present framework, the first class of facts can no longer provide that motivation since metarules must apply to ID rules and cannot, therefore, have any direct bearing on Linear precedence. Furthermore, given the very general approach to feature instantiation elaborated in the IULC paper, it does not seem appropriate to also express grammar-wide generalizations about feature-passing by means of metarules, and thus the second class of facts mentioned above no longer provides a motivation for an unrestricted domain of application.

Let us explore the implications of assuming the following restriction: metarules may only apply to lexical ID rules. <1> That is to say, we henceforth assume that metarules may only apply to rules which explicitly introduce a lexical category, i.e. a category having the general form O/-V, +/-N, [LEXICAL jjl]. This set of rules is, in fact, identical to the set of rules which introduce lexical heads. Under this revised conception of metarules then, their role is simply to express generalizations about possibilities of subcategorization.

?L^m Coordination schemata

All languages, as far as we know, make use of coordinate constructions. This apparently innocent empirical claim implicitly embodies an important metatheoretical assumption to the effect that there is a unitary notion of "coordinate construction", one that can be abstracted away from the evident differences that pertain between, say, English and Latin coordination. <2> <3>

The theory of coordination presented in this section, and continued in the next, is able to locate all the parochialism of coordinate constructions in two components of the grammar that stand already as the repository of such parochialism, namely the LP statements and the rules responsible for the realization of features. Everything else, and that includes the definition of coordinate structure, the coordination schemata, and the feature instantiation principle for conjuncts, is to be taken as universal. <4>

(2) Definition: a structure rooted in @ is coordinate
if and only if every daughter of @ is a conjunct.

(3) Definition: a category (constituent) @ is a conjunct
if and only if @ is [+CONJ].

The feature CONJ is to be introduced (optionally, like every other feature) as a potential coefficient of CAT'.

(4) [CAT' CAT FOOT CONJ]

The range of coefficients for CONJ itself, in English, is as shown in (5).

(5) [CONJ {and, both, but, e, either, neither, nor, or}]

And the [CONJ @] categories expand in the manner shown below.

(6) CAT'[[BAR $][CONJ @]] --> @, CAT'[[BAR $][- CONJ]
where @ is in {and, both, but, e, either, neither, nor, or}

Here e is the empty string, and the rationale for including it will become apparent in the discussion below. <5>

There are two possible coordinate constructions, in one iteration of a class of conjuncts is permitted, in the other iteration is not possible, and exactly two conjuncts appear. In (7a) and (8a), respectively, we exhibit the two ID schemata responsible for these two constructions, and in (7b) and (8b) we provide the English coefficients for the CONJ features required by them.

(7) a. CAT' --> CAT'[CONJ @0], CAT'[CONJ @1]+
b. @ is in {<and, e>, <e, and>, <neither, nor>,
<or, e>, <e, or>}

(8) a. CAT' --> CAT'[CONJ @0], CAT'[CONJ @1]
b. @ is in {<both, and>, <either, or>, <e, but>}

A number of points need to be made concerning these two schemata. <6> Firstly, the only variables in (7a) and (8a) are @0 and @1. Unlike all previous proposals concerning coordination schemata, "CAT'" is not a variable ranging over categories which enforces categorial identity across the conjuncts, by means of a uniform substitution principle implicit in the interpretation of such earlier schemata. Instead, CAT' just is a category, albeit a minimal one. In effect, and modulo CONJ, all (7a) and (8a) say is that the mother is a category, and so are the daughters. If feature instantiation were completely free, which, of course, it is not, then (8a) would in principle be consistent with, for example, a coordinate structure in which a preposition expanded as an NP conjoined with a VP. The only information that (7a) and (8a) contain concerns the distribution of CONJ and the possibility of iteration. Everything else has been factored out of them and is to be found in the conjunct realization principle discussed below.

A consequence of this is that the schema in (8a) collapses exactly three English coordination rules, namely those arrived by substituting in the three possible coefficient pairs for @. These three rules will hav
---- instantiations after feature instantiation, but they are no

schemata over those instantiations. We are stressing this point because it has important empirical consequences to which attention will be drawn subsequently. In particular, the approach we have adopted which abjures alpha-variables ranging over fully specified categories does not entail that every conjunct be categorially identical to each of its sisters, although near-identity usually follows as a consequence of the interaction of the various principles of feature instantiation. This point is pursued in detail below.

The schema in (7a) collapses infinitely many rules in virtue of the presence of "+" on CAT'CCONJ 311. This symbol denotes the positive closure of the set containing this category (see Hopcroft and Ullman (1979:28)), that is to say that one or more instances of the category may appear with no upper bound on how many. The use of regular expressions in phrase structure rule schemata allows the resulting grammar to contain infinitely many rules, but this has no undesirable formal consequences, in particular, the resulting infinite grammars have exactly the same weak generative capacity as CF-PSG's which are restricted to a finite set of rule (this is trivial to prove, and it follows in any case as a corollary of a more general theorem proved by Langendoen (1976)).

Two minor points are worth noting here. Firstly, (7a) is a schema over ID rules and the formalism for the latter separates the names of categories with commas, thus, very pedantically, what we really have here is the positive closure of a set whose only member is a string of the form

<comma>~<space>*CAT^{f}CCON<T<space>~<value for 31>^{A}]

Here "~" is the concatenation symbol, and angle brackets enclose descriptions of characters in the string.

Secondly, notice that there is an alternative interpretation available for expressions of the form shown in (7a), one in which it constitutes not a schema but a single rule (modulo 3) which admits infinitely many mother-plus-daughters *tree* fragments. This alternative interpretation, which is perfectly coherent, formally speaking, was adopted, uniquely as far as we know, in Gazdar (1982). The interpretation of (7a) that we adopt here is, by contrast, the conventional one in which it constitutes not a single rule, but rather a grammar for an infinite set of rules. The difference in interpretation makes an empirical difference given the theory of feature instantiation presented in GP82. That theory instantiates features on rules rather than tree fragments. A consequence of this is that maintaining the Gazdar (1982) interpretation of the positive closure notation would predict that all the iterated conjuncts, though not the single non-iterated conjunct (i.e. GAT'CCONJ 303 in (7a)) would have to be identical in respect of every coefficient for every feature (since the rule, under that interpretation, only contains a single expression making reference to the iterated conjuncts). The more familiar interpretation, and the one that we adopt here, does not make this prediction. The prediction is false. <7>

Notice, incidentally, that none of the rules collapsed by (7a), or (8a) either, for that matter, is a lexical ID rule, and thus no conflict arises between the infinite cardinality of this set of rules, and the finite closure condition on metarule application. The coordination rules, like "S —> NP, VP" and every other nonlexical ID rule, are outside the domain of metarules.

Apart from the putatively universal feature instantiation principle for conjuncts discussed in the next section, our analysis of coordination needs one further parochial component before we can explore what claims it makes concerning the structure of English. We need LP "statements to express the ordering constraints that hold across the various types of conjunct characterized by distinct coefficients for the feature CONJ. These LP statements can be collapsed into a single schema, as shown in (9).

(9) LCONJ 303 < tCONJ [illegible]
where 80 is in {both, e, either, neither>,
and 31 is in -Cand, but, nor, or>.

The three schemas shown in (7), (8), and (9) interact to make a very wide range of detailed predictions concerning possible and impossible coordinate constructions in English. We will illustrate these predictions by reference to examples involving coordinate VP's, and concern ourselves only with the predictions made in respect of iterability and choice of conjunction morpheme, since these are the issues that (7) - (9) address. The categorial identity, or lack of it, between mother and conjunct, and between conjunct and conjunct, is a topic that we leave to the next section. Here we will simply assume that constituents of the same category can conjoin to form a coordinate constituent of that category.

Since there are eight distinct coefficients for CONJ in English, it follows that there are 64 logically possible two-conjunct coordinate structures. However, only six of these 64 possibilities are, in fact, grammatical, and these six are exhibited in (10): <8>

(10) a. Kim made a speech and stuttered.
b. Kim made a speech or stuttered.
c. Kim neither made a speech nor stuttered.
d. Kim both made a speech and stuttered.
e. Kim either made a speech or stuttered.
f. Kim made a speech but stuttered.

Schemas (7) and (9) interact to provide the rules necessary for (10a), (10b), (10c), and no others. Schemas (8) and (9) interact to provide the rules necessary for (10d), (10e), (10f), and no others. Thus the schemas we have given induce all and only the six grammatical two-conjunct grammatical English coordinate structures.

We turn our attention now to the flat three-conjunct constructions. <9> Here there are 512 logical possibilities. Of these only five are grammatical. They are shown in (11).

(11) a. Sandy whimpered, shouted, and screamed.
b. Sandy whimpered, and shouted, and screamed.
c. Sandy whimpered, shouted, or screamed.
d. Sandy whimpered, or shouted, or screamed.
e. Sandy neither whimpered, nor shouted, nor screamed.

Schema (8) is no help with any of these since it introduces exactly tv* conjuncts. However, schema (7) interacts with the LP schema in (9) t -:-- 'ic the rules needed for these five examples, but not for any of tt

507 other possibilities. Finally, consider the case of flat four-conjunct constructions. There are 4096 possibilities, but, again, only five are grammatical, and just those five are legitimated by the interaction of (7) and (9).

(12) a. Sandy moaned, whimpered, shouted, and screamed.
b. Sandy moaned, and whimpered, and shouted, and screamed.
c. Sandy moaned, whimpered, shouted, or screamed.
d. Sandy moaned, or whimpered, or shouted, or screamed.
e. Sandy neither moaned, nor whimpered, nor shouted, nor screamed.

There is some variation among speakers with respect to *either* and *neither*. More liberal varieties than ours permit the examples in (13):

(13) a. %Sandy either whimpered, or shouted, or screamed.
b. %Sandy either moaned, or whimpered, or shouted, or screamed.

And more conservative varieties probably do not permit (11e) and (12e), above. Such varieties are straightforwardly catered for in the present analysis by minor changes to the parochial components of (7) and (8). Thus, to allow the more liberal varieties, one simply moves <*either*, *or*> from the list in (8b) to the list in (7b). And to handle the more conservative varieties, one just shifts <*neither*, *nor*> in the opposite direction.

3. The Conjunct Realization Principle

The coordination schemata that we introduced above dealt only with the distribution of the coordination morphemes, and with the possibility of iteration. Nothing was said about the categorial status of conjuncts, or about the category of the mother given the categories of conjunct daughters. The conventional wisdom on this topic has it that conjuncts must all be of the same category, say @, and that the mother of these conjuncts will also be of category @. And the conventional wisdom is widely known to be wrong, or at best, seriously incomplete. However, in the absence of any other candidate analyses, it has yet to be replaced.

There are two classes of phenomena that show that the conventional wisdom is wrong. Firstly, the case, person, number, and gender of coordinate NP's does not behave in the manner it would lead you to expect. Thus, for example, a singular NP can conjoin with a plural NP to form a plural NP. In common with the rest of the field, we have no analysis of any of the NP facts, which we can only leave for further research. A good starting point for such research is Corbett (1983), who brings together and elucidates most of the cross-linguistic generalizations there are to be made concerning the agreement features of coordinate NP's.

The second class of phenomena concerns what one might call, nontechnically, predicates, as in the following examples.

(14) She walked slowly and with great care.
(15) He was longwinded and a bully.

Here we apparently have a coordination of AdvP with PP in (14) and of AP with NP in (15). Two questions arise immediately: (i) What category is the mother of the conjuncts in each case? And (ii) why is the coordination of AP with NP that appears to be possible in (15), not possible in (16) or (17)?

(16) *The [longwinded and a bully] man entered.
(17) *[Longwinded and a bully] entered.

It is questions like this that the theory of conjunct realization outlined here is intended to address.

Given the theory of features developed in GP82, and, in particular, the notion "extension of a category" defined there, it is very simple to formulate the principle of feature instantiation for conjuncts that we are proposing. Instead of saying that conjuncts in a coordinate structure must have an identical featural make-up, both to their sisters and to their mother, we maintain only that each conjunct has to be an extension of the mother. This can be expressed formally as follows. <10>

(18) Conjunct Realization Principle (CRP)
If \$i is a conjunct, then Ext(@i, @0).

Notice that this does not entail that any conjunct is identical to the mother, although it is consistent with any or every conjunct being identical to the mother. Nor does it entail than any conjunct is identical to any other, although, again, it is consistent with some or all of the conjuncts being featurally identical. It does not even entail that the various conjuncts are nondistinct from each other, although it does entail that every conjunct is nondistinct from the mother.

To see how this works in practice, consider the case of tensed VP's. The feature TENSE in English has two possible coefficients as shown in (19).

(19) [TENSE {PAST, NONPAST}]

Thus VP[TENSE PAST] is an extension of VP[TENSE], as is VP[TENSE NONPAST]. The CRP allows nine possibilities for expanding VP[TENSE] as two conjuncts. These are shown in (20).

(20) a. VP[TENSE] --> VP[TENSE], VP[and, TENSE]
b. VP[TENSE] --> VP[TENSE], VP[and, [TENSE PAST]]
c. VP[TENSE] --> VP[TENSE], VP[and, [TENSE NONPAST]]
d. VP[TENSE] --> VP[TENSE PAST], VP[and, TENSE]
e. VP[TENSE] --> VP[TENSE PAST], VP[and, [TENSE PAST]]
f. VP[TENSE] --> VP[TENSE PAST], VP[and, [TENSE NONPAST]]
g. VP[TENSE] --> VP[TENSE NONPAST], VP[and, TENSE]
h. VP[TENSE] --> VP[TENSE NONPAST], VP[and, [TENSE PAST]]
i. VP[TENSE] --> VP[TENSE NONPAST], VP[and, [TENSE NONPAST

Now the rule responsible for canonical tensed NP VP sentences only requires that the VP be tensed, it does not specify which tense is to appear (although, of course, the lexicon ultimately forces a choice). And since the present framework treats partly specified categories in exactly the same way as fully specified categories, namely just as categories, there is nothing to stop the partly specified VP[TENSE] appearing as a node label in a structural description. And, since that is true, then there is nothing to stop any of the nine rules listed above admitting some coordinate mother-daughters substructure of a tree. In particular, the instantiated rules (20e), (20f), (20h), and (20i), will be responsible for the examples (21e), (21f), (21h), and (21i), below, respectively.

(21) e. Kim liked cats and alienated dogs.
f. Kim liked cats and alienates dogs.
h. Kim likes cats and alienated dogs.
i. Kim likes cats and alienates dogs.

But none of the nine coordination rules that CRP provides us with will allow for strings like those in (22).

(22) a. *Kim alienated cats and liking dogs.
b. *Kim alienates cats and to like dogs.
c. *Kim alienated cats and liked by dogs.

The CRP makes reference only to conjuncts, not to coordinate structures. It therefore predicts that a conjunct will be an extension of its mother even when it appears in a noncoordinate constituent. At first sight, this appears to be merely an artefactual possibility, with no empirical consequences. However, this is not so. Consider the following rule of English grammar.

(23) VP --> V[99], (PP), VP[CONJ and]
where at least come and go are instances of V[99].

Given the definitions in the previous section, the complement VP here is a conjunct, even though the construction in which it appears is not a coordinate one. The CRP thus leads us to expect that it will be an extension of the mother. Thus, if the mother VP is tensed, past or nonpast, we will expect it to be the same. <11> And if the mother is a bare infinitive, then it should be a bare infinitive too. Surprisingly, these predictions turn out to be true.

(24) a. Sandy went and got a drink.
b. *Sandy goes and got a drink.
c. *Sandy went and gets a drink.
d. Sandy goes and gets a drink.
e. *Sandy goes and get a drink.
f. I want Sandy to go and get a drink.
g. Go and get a drink!

These facts do not follow from the HFC, since the complement VP is not the head of the construction. Nor do they follow from the CAP since that only deals with agreement features, not verb form features like choice of tense.

We will be returning to this interesting, but neglected, construction in section 10 below, when we discuss, in some detail, unbounded dependencies and their interaction with conjuncts and coordinate structures. We conclude this section by briefly noting a theorem that follows from FFP and CRP, one which has a crucial bearing on the interaction just mentioned.

(25) In a coordinate construction, the mother and all the daughters have identical foot feature coefficients.

We have seen that this is not true for other kinds of feature - why, then, should it be true for foot features?

The coordination schemata themselves make no mention of foot features. It follows therefore that any foot feature instantiated on a coordinate daughter or mother will be an increment resulting from feature instantiation. By FFP, the mother's foot feature will consist of the unification of the foot features of all its daughters. <12> By definition, a coordinate structure is one in which every daughter is a conjunct. Hence, by CRP, it follows that every daughter is an extension of the mother in a coordinate structure, and thus that every daughter's foot feature is an extension of the mother's foot feature. But the latter, as we have seen, is just the unification of the foot features of all the daughters, and hence an extension of the foot feature of each daughter. Since the mother's foot feature is an extension of the daughter's, and the daughter's is an extension of the mother's, they must be identical. And, since this holds for each daughter, every daughter's foot feature must be identical to the mother's, and hence every daughter's foot feature must be identical to every other daughter's foot feature.

4. The percolation of SLASH features

An "unbounded dependency" construction (UDC, hereafter) is one in which (i) a syntactic relation of some kind holds between two substructures in the construction, and (ii) the structural distance between these two substructures is not restricted in any way (e.g. by a requirement that both be substructures of the same simple clause). Historically, topicalization, relative clauses, constituent questions, free relatives, clefts, and various other constructions in English have been taken to involve a dependency of this kind. It is analytically useful to think of such constructions as having three parts, the top, the middle, and the bottom. The top is the substructure which introduces the dependency, the middle is the domain of structure that the dependency spans, and the bottom is the substructure in which the dependency ends, or is eliminated. We shall be claiming that the principles which govern the bottom and the middle are completely general in character, and do not serve to distinguish one UDC from another.

The theory of foot features elaborated in GP82 is crucial to our analysis of UDC's. Indeed, our analysis of the middle of UDC's consists of no more than the feature SLASH taken together with the Foot Feature Principle (FFP). Since this is at once the simplest and the most important part of our treatment of UDC's, we will begin with it. That is to say, we will begin in the middle.

GP82 introduced categories that can be conveniently notated as @/$, where this is an abbreviation for [CAT' @ [FOOT [SLASH $]]]. Conceptually, an @/$ is to be thought of as an @ that has the category $ as the coefficient of its SLASH feature. And the intuitive interpretation of @/$ is that of 'an @ with a $ missing'. Thus, canonically, a constituent of category @/$ will be a constituent o category @ which has (at least) a hole in it where one would expect t find a constituent of category $. So, for example, an S/NP is sentence which is missing a noun phrase.

In earlier GPSG work (e.g. Gazdar 1981), the rules responsible for passing slash category information through a tree were arrived at by a metarule known variously as the "derived rule schema", or the "slash introduction metarule". This metarule has no place in the present analysis. Everything that it did, and some other things, now follows as a special case of the application of the FFP in feature instantiation. Furthermore, given the restriction on metarule application to lexical ID rules, it is no longer be possible to formulate a "slash introduction metarule" with the necessary properties.

The best way to grasp the effect of the FFP in respect of slash categories is to inspect an example of its application. Consider the set of ID rules shown in (26): <13>

```
(26)    i.      S   -->  NP,  VP
                NP  -->  Det,  N'
                S'  -->  that,  S
                VP  -->  VP,  PP
                AP  -->  A'

        ii.     N'  -->  N[1]
                N'  -->  N[2],  PP
                PP  -->  P[3],  NP
                VP  -->  V[4]
                VP  -->  V[5],  NP
                VP  -->  V[6],  PP
                VP  -->  V[7],  AP
                VP  -->  V[8],  VP
                VP  -->  V[9],  S
                VP  -->  V[10],  S'
                VP  -->  V[11],  NP,  PP
                VP  -->  V[12,there],  NP, PP
                VP  -->  V[13],  NP,  VP
                VP  -->  V[14],  NP,  S
                VP  -->  V[15],  NP,  S'
                VP  -->  V[16],  PP,  S'
                A'  -->  A[17],  PP
```

Suppose that this set of rules is instantiated by FFP with respect to an NP coefficient for SLASH. The grammar will retain all the rules shown in (26), and will now include, in addition, all the rules shown in (27). <14>

```
(27)    i.      S/NP   -->  NP/NP,  VP
                S/NP   -->  NP,  VP/NP
                S/NP   -->  NP/NP,  VP/NP
                NP/NP  -->  Det/NP,  N'
                NP/NP  -->  Det,  N'/NP
                NP/NP  -->  Det/NP,  N'/NP
                S'/NP  -->  that,  S/NP
                VP/NP  -->  VP/NP,  PP
                VP/NP  -->  VP,  PP/NP
                VP/NP  -->  VP/NP,  PP/NP
                AP/NP  -->  A'/NP
```

ii. N'/NP --> N[2], PP/NP
PP/NP --> P[3], NP/NP
VP/NP --> V[5], NP/NP
VP/NP --> V[6], PP/NP
VP/NP --> V[7], AP/NP
VP/NP --> V[8], VP/NP
VP/NP --> V[9], S/NP
VP/NP --> V[10], S'/NP
VP/NP --> V[11], NP/NP, PP
VP/NP --> V[11], NP, PP/NP
VP/NP --> V[11], NP/NP, PP/NP
VP/NP --> V[12,there], NP/NP, PP
VP/NP --> V[12,there], NP, PP/NP
VP/NP --> V[12,there], NP/NP, PP/NP
VP/NP --> V[13], NP/NP, VP
VP/NP --> V[13], NP, VP/NP
VP/NP --> V[13], NP/NP, VP/NP
VP/NP --> V[14], NP/NP, S
VP/NP --> V[14], NP, S/NP
VP/NP --> V[14], NP/NP, S/NP
VP/NP --> V[15], NP/NP, S'
VP/NP --> V[15], NP, S'/NP
VP/NP --> V[15], NP/NP, S'/NP
VP/NP --> V[16], PP/NP, S'
VP/NP --> V[16], PP, S'/NP
VP/NP --> V[16], PP/NP, S'/NP
A'/NP --> A[17], PP/NP

This is a rather large set of rules, but there are vastly more which the FFP will not legitimate. Some representative examples of improper instantiations of the "S --> NP, VP" rule are shown in (28).

(28) S --> NP/NP, VP
S --> NP, VP/NP
S/NP --> NP, VP
S/NP --> NP/PP, VP
S/NP --> NP, VP/PP
S/NP --> NP/PP, VP/NP

In the case of the simple rules we have been considering, the generalizations that the FFP imposes are these: (i) at least one, but possibly more than one, daughter has its slash feature instantiated in a manner identical to the instantiation on the mother, and (ii) no daughter gets its slash feature instantiated in a manner distinct from the instantiation on the mother. We will consider the application of FFP in more complex instances, in particular its application to coordinate rules, and to rules whose uninstantiated counterparts already include a slashed category, later on.

5. Slash termination, metarule 1

We will turn our attention now to the bottom of UDC's, the part of the structure in which the chain of slashed categories (i.e. the "projection path" in the sense of Fodor (1982)) comes to an end and we reach an incomplete constituent or gap. In the earliest GPSG work on UDC's (i.e. Gazdar 1981, written in 1979), this part of the structure was handled by two mechanisms: (i) a phrase structure rule schema expanding categories of the form @/@ as the empty string, and (ii) a metarule that allowed an S/NP to be replaced by VP. For all its elegance and generality, it turned out that (i) entailed false empirical predictions about coordinate structures (e.g. _Which book did you read and a review of_?). <15> And (ii) suffered from a lack of both elegance and generality. In the present work, we shall continue to maintain that two distinct mechanisms are involved in slash termination, but both will be realized as metarules. Our replacement for (i) will be a variant of the metarule analysis originally proposed in Sag (1981), and our replacement for (ii) will have much of the effect of (ii) following as a special case. Initially, however, we will only introduce and discuss the former. <16>

(29) Slash termination metarule 1 (STM1)

```
@  -->  W,  $[-CASE]

            |
            V

@/$  -->  W
```

This says that any rule which introduces a caseless category has a counterpart in which that daughter is missing, but where the mother has the category of the daughter assigned to be the value of its SLASH feature. Notice that, given the way our grammar is organized, STM1, being a metarule, can only apply to rules that specifically introduce a lexical category, i.e. the lexical ID rules, and must apply prior to feature instantiation. To show the effect it has, we exhibit in (30) all the rules that result from the application of STM1 to the rules in (26ii) when the value of $ is AP, NP, or PP. <17>

```
(30)    N'/PP   -->  N[2]
        PP/NP   -->  P[3]
        VP/NP   -->  V[5]
        VP/PP   -->  V[6]
        VP/AP   -->  V[7]
        VP/NP   -->  V[11],  PP
        VP/PP   -->  V[11],  NP
        VP/NP   -->  V[12,there],  PP
        VP/PP   -->  V[12,there],  NP
        VP/NP   -->  V[13],  VP
        VP/NP   -->  V[14],  S
        VP/NP   -->  V[15],  S'
        VP/PP   -->  V[16],  S'
        A'/PP   -->  A[17]
```

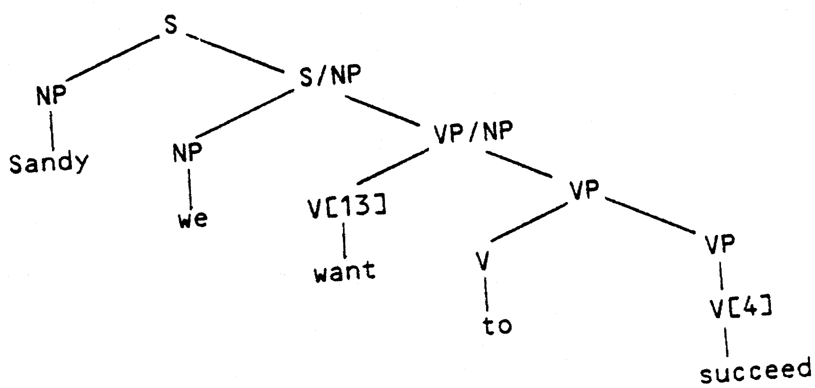
6. Two topicalization constructions

We are, at last, now in a position to discuss the top of UDC's and, in the light of our discussion so far, provide some concrete illustrative examples. The most basic and general rule responsible for introducing UDC's in English is that shown in (31).

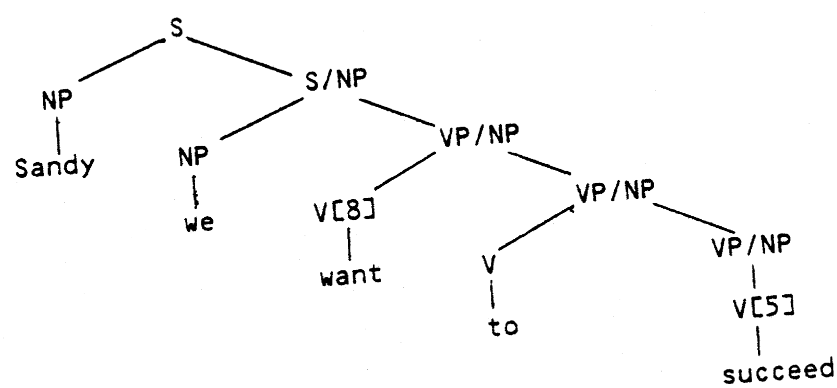
(31) S --> @ S/@

This rule says that a sentence can consist of an @ followed by a sentence with an @-type hole in it, and it is responsible for, among other things, the topicalization construction. The trees shown in (32) - (36), below, illustrate this.

(32)

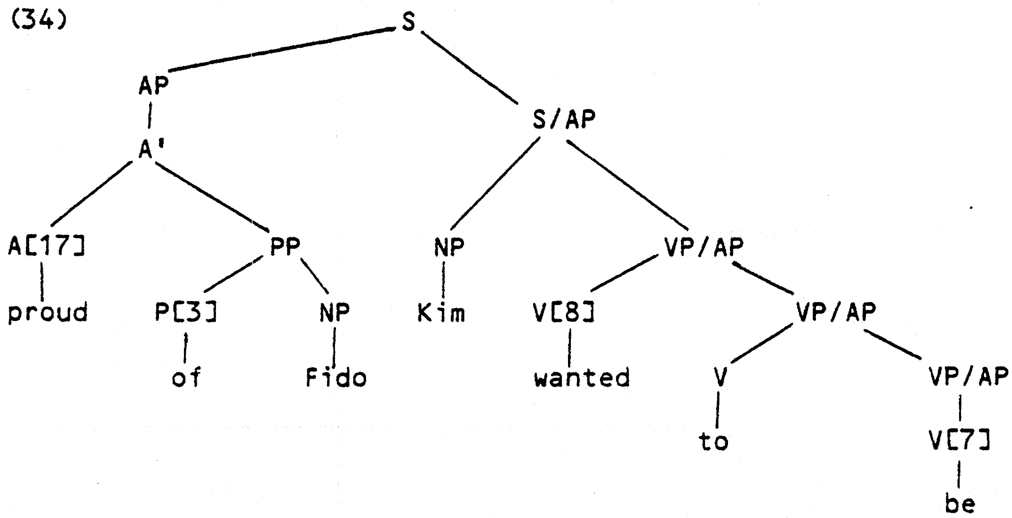


(33)

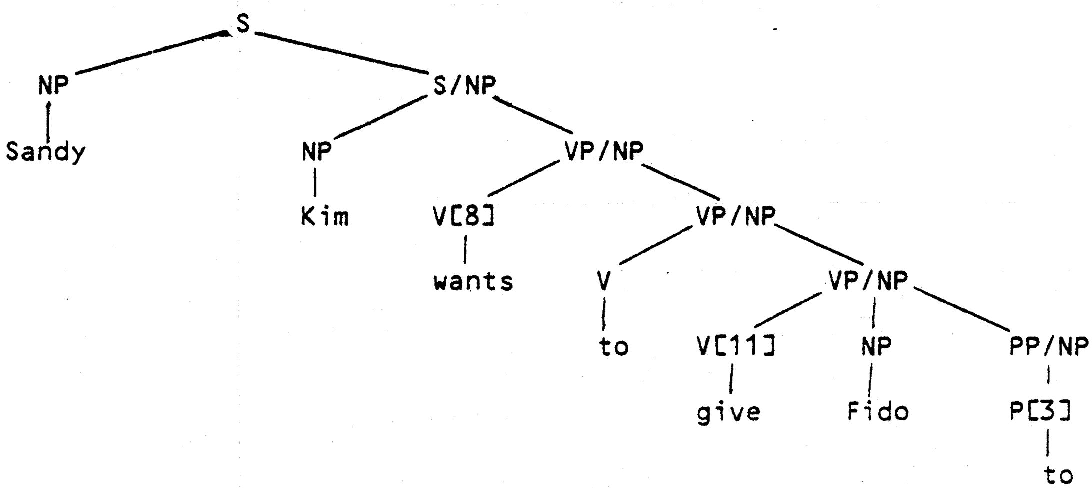


In (32), we have the structure associated with the "we want Sandy to succeed" reading of <u>Sandy we want to succeed</u>, whereas in (33), we have the structure associated with the "we want to succeed Sandy" reading.

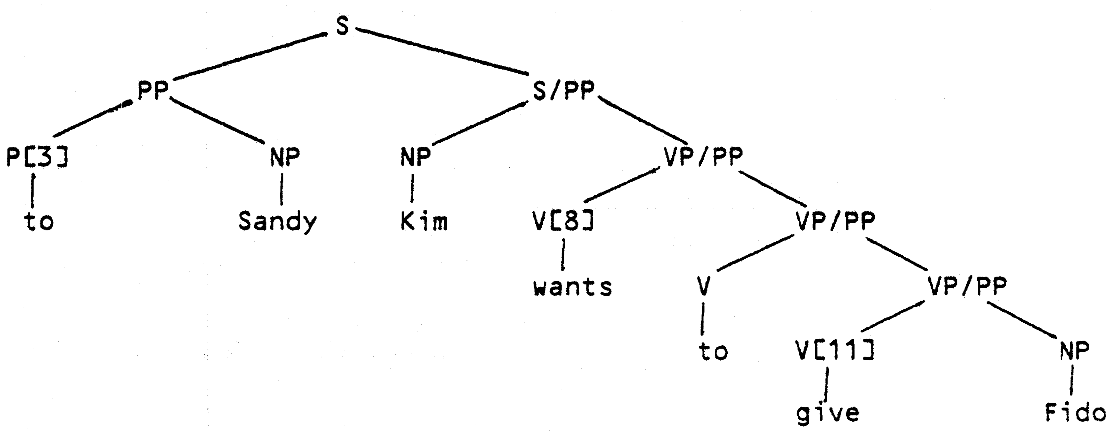
(34)



(35)



(36)



Examples (35) and (36) show that the rules given allow both for NP's to be topicalized so as to strand prepositions, or for the whole PP to be topicalized. Thus the analysis gets the pied-piping facts exactly right, and will not permit examples such as (37) (from Iwakura's (1980) critique of Chomsky (1977)).

(37) *To John, Julie gave a copy of her book to.

The rules permit a topicalized PP to bear a dependency into an N' within a subject NP, and thus we generate (38) (cf. Chomsky 1977:112).

(38) Of Kim's book, a review was published.

And the C-CASE3 requirement on STM1 captures, inter alia, the contrasts noted in (39) (the examples are adapted from Grosu (1974)),

(39) a, A book of Kim's, I read.
b. *Kim's, I read a book of.
c. Kim, I saw a picture of.

Topicalization is permitted in embedded clauses, thus we get the examples in (40) (cited by Iwakura 1980:59).

(40) a. Harry said that Max, Joan would never be willing to marry.
b. The inspector explained that each part he had examined very carefully.

If (31) is the most canonical UDC rule, then the next rule we introduce must be a good candidate for the least. We introduce it here, not because of it has any intrinsic importance, but because it interacts in interesting and unexpected ways with the WH feature and the second slash termination metarule as we shall show below.

(41) S —> *P?* VPttherel/PP

This odd looking rule is responsible for the construction illustrated by the examples in (42) ((42b) is due to Postal (1977:146)). <18>

(42) a. In the garden is a fountain.
b. Near the fountain sat a large purple gorilla.
c. Behind the trees appeared to stand a large building of some kind.

Example (42c) is important because it shows that the construction involves an unbounded dependency, and not merely a clause-bounded phenomenon such as "subject-auxiliary inversion". There is good evidence, too, that the construction crucially involves the existential Ctherel feature: Postal (1977) contains a whole series of arguments which, under the assumptions made here, have this as a consequence; Gazdar and Pullum (1980) provide an argument to this effect based on subject-verb agreement; and, classically, tag questions point to the same conclusion. <19>

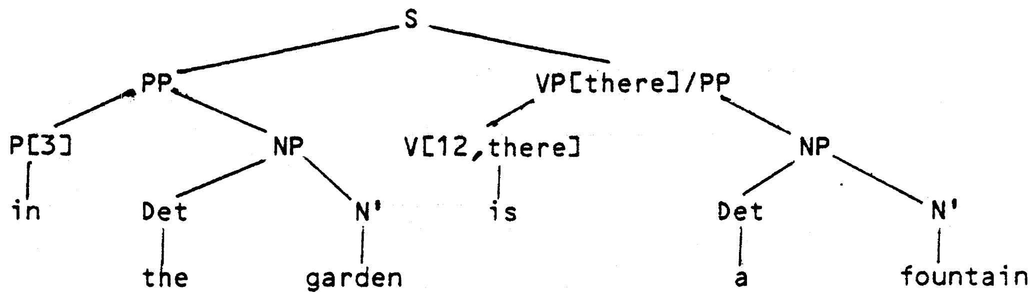
(43) Near the fountain sat a large purple gorilla, didn't there?

Notice that our analysis of ordinary topicalization already provides for the examples in (44). Nothing special has to be said about them.

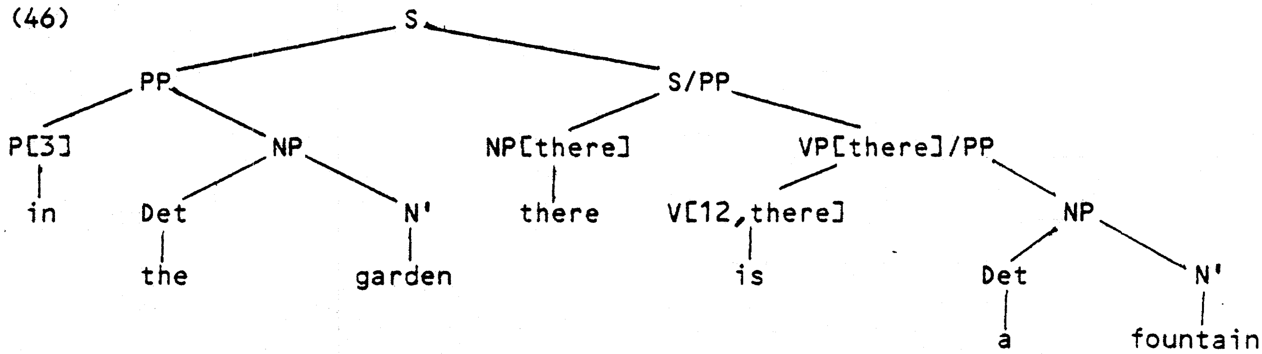
(44) a. In the garden, there is a fountain.
b. Near the fountain, there sat a large purple gorilla.
c. Behind the trees, there appeared to stand a large building of some kind.

The trees in (45) and (46) are the structures assigned to examples (42a) and (44a), respectively.

(45)



(46)



7. *WH constructions*

Apart from SLASH, there is another foot feature associated with UDC's, and that is the feature WH. This feature, according to GP82, has the internal structure shown in (47).

(47) [WH AGR WHMOR]

We will ignore the agreement feature AGR, for the moment. The feature WHMOR encodes the morphological type of the *wh* expression involved. It has at least three, and perhaps four, distinct coefficients:

(48) [WHMOR {Q, R, FR, EX}]

The two that we shall be concerned with here are Q, for interrogatives, and R, for relatives. The features FR and EX are intended for free relatives and exclamatives (*what a nice day*!), respectively. FR, for example, will allow the -*ever* forms peculiar to free relatives, whilst the distinction between Q and R will allow the lexicon to exclude the possibility of *what* being used as a relative pronoun in standard English. We will use [+Q] and [+R] as abbreviations for [WH AGR [WHMOR Q]] and [WH AGR [WHMOR R]], respectively. And we will assume the existence of FCR's which forbid the appearance of the WH feature on VP or A'. The impossibility of either questioning or relativizing on VP constituents in English provides support for the VP FCR. Note the contrast between the examples in (49).

(49) a. and going to that party, you certainly are!
b. *Going to what party are you?

And the A' FCR is motivated by the following contrast, pointed out to us by Stan Peters.

(50) a. How fond of him, are you?
b. *Fond of whom, are you?

The proposed FCR will not forbid the appearance of WH on AP, thus it will permit (50a), but it forbid its instantiation on the daughter A', thus preventing generation of (50b). The effect of the FCR is thus to restrict the appearance of wh words in AP's to the determiner position.

We will also assume an FCR making WH and SLASH mutually exclusive, an FCR which can be motivated by the need to prevent the generation of examples like (51). <20>

(51) *Which books did you wonder whose reviews of had annoyed me?

The following ID rules will now serve to introduce relative clauses and embedded questions, respectively.

(52) N' --> N', S[+R]
(53) VP --> V[18], S[+Q]

These rules naturally lead one to ask what the rules are which allow S[+R] and S[+Q] to be expanded. And the answer to this question, or at least part of the answer, is that these rules cannot help but include the rules that we already have which expand S. We have employed three such rules in this paper already, and they are repeated in (54) below, for convenience.

(54) a. S --> NP, VP
b. S --> PP, VP[there]/PP
c. S --> α, S/α

The features [+R] and [+Q] are WH features, and thus foot features. So they will be instantiated by the FFP in exactly the same way as SLASH features, provided, of course, that the resulting instantiations are consistent with the prevailing FCR's. This means that the grammar will automatically contain the instantiations of (54) shown in (55) and (56).

(55) a. S[+R] --> NP[+R], VP
b. S[+R] --> PP[+R], VP[there]/PP
c. i. S[+R] --> NP[+R], S/NP
ii. S[+R] --> PP[+R], S/PP
iii. S[+R] --> AP[+R], S/AP

(56) a. S[+Q] --> NP[+Q], VP
b. S[+Q] --> PP[+Q], VP[there]/PP
c. i. S[+Q] --> NP[+Q], S/NP
ii. S[+Q] --> PP[+Q], S/PP
iii. S[+Q] --> AP[+Q], S/AP

These rule instantiations, none of which have had to be specially listed in the grammar, interact with rules (52) and (53) so as to generate all the examples in (57) and (58).

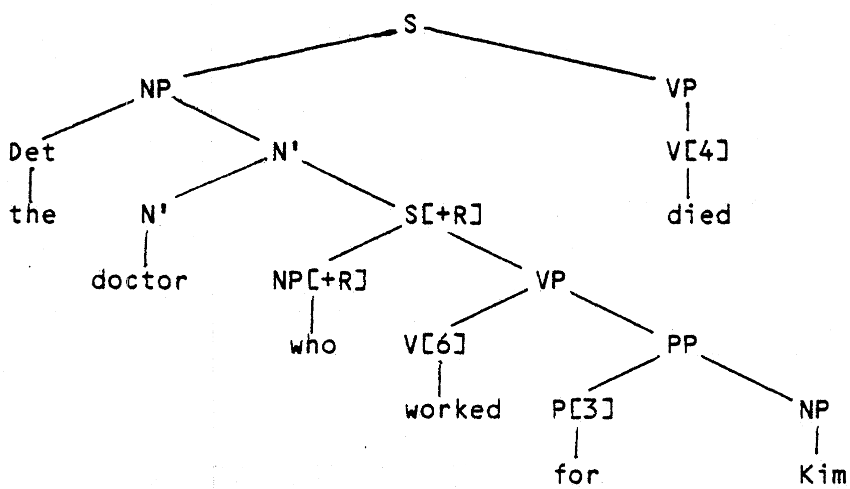
(57) a. The doctor who worked for Kim died.
b. The villages in whose cafes are to be found the best wines are the least hospitable.
c. i. The doctor who Kim worked for died.
ii. The doctor for whom Kim worked died.

Note that there is no exemplar for (55c.iii) in English, although some other languages permit adjectival relatives of this kind. Their absence in English is explained partly by the A' FCR discussed above, and partly by the fact that the English lexicon contains no [+R] counterpart to the interrogative adjectival determiner <u>how</u>.

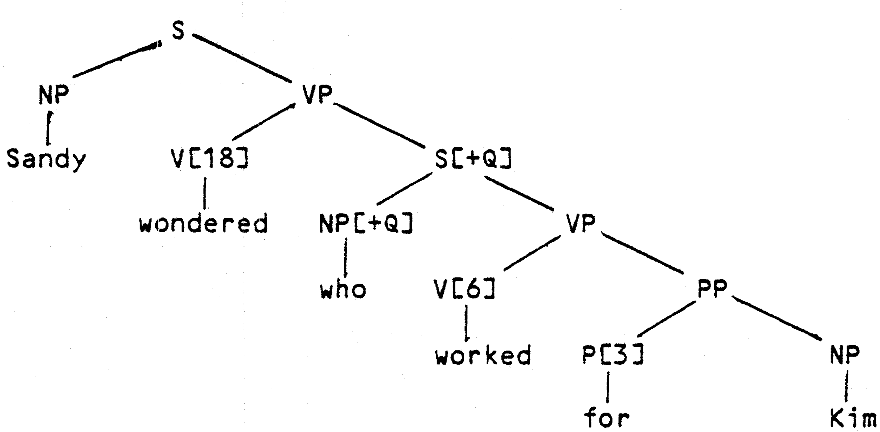
(58)
- a. Sandy wondered who worked for Kim.
- b. Sandy wondered in which villages were to be found the best wines.
- c.
  - i. Sandy wondered which doctor Kim worked for.
  - ii. Sandy wondered for which doctor Kim worked.
  - iii. Sandy wondered how expensive the wine was.

It is a surprising consequence of the present proposals that matrix subject relatives and constituent questions like those in (57a) and (58a) must have, and can only have, the simple NP VP structures shown in (59) and (60).

(59)



(60)



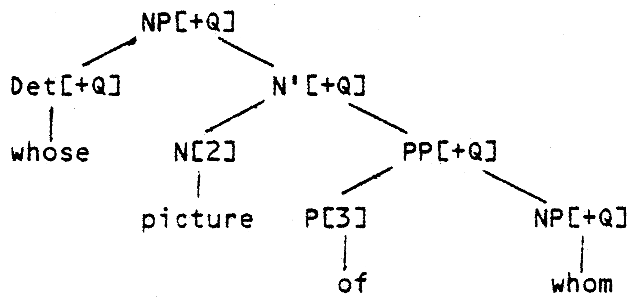
They must have this structure because of the existence of the "S --> NP, VP" rule and the principle that governs the way foot features work. And they can only have this structure, and not, say, one in which the <u>wh</u> NP is followed by a sentence with a missing subject, because no other rules are available. In particular, STM1, being a metarule, can only apply to lexical ID rules and cannot, therefore, apply to the "S --> NP, VP" rule so as to eliminate the subject in favour of a slash on the mother. The present framework thus offers no counterpart to the vacuous movement analysis of such sentences standardly offered in transformational accounts. <21>

The FFP will instantiate WH features on NP and PP rules, as well as the S rules shown in (54). Thus among the instantiated extensions of the rules given in (26) will be the following.

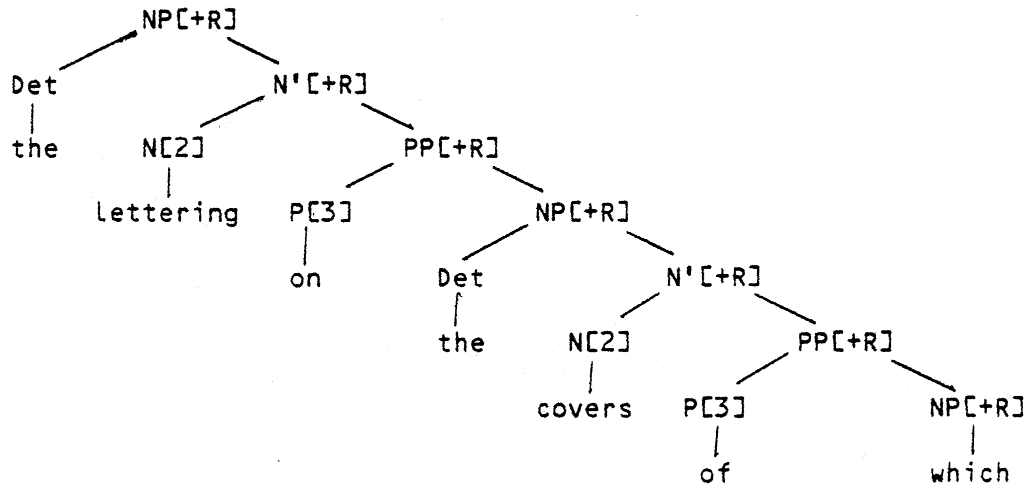
```
(61)    N'[+R]   -->  N[2]   PP[+R]
        N'[+Q]   -->  N[2]   PP[+Q]
        NP[+R]   -->  Det   N'[+R]
        NP[+Q]   -->  Det[+Q]   N'[+Q]
        PP[+R]   -->  P[3]   NP[+R]
        PP[+Q]   -->  P[3]   NP[+Q]
```

And these rules will induce structures such as those shown in (62) and (63). <22>

(62)



(63)



Let us briefly consider *that* relatives. These have a slightly different distribution from *who/which* relatives (see Brame (1981:277, n1) for relevant data), and thus it is reasonable to suppose that this will be reflected in the WH feature in some way, say in a coefficient for R. If we make this supposition, and if we assume, in addition, the existence of an FCR requiring a category bearing this coefficient to be an NP, then we immediately obtain an account for the contrast noted in (65).

(64) a. The doctor who Kim worked for died.
     b. The doctor that Kim worked for died.

(65) a. The doctor for whom Kim worked died.
     b. *The doctor for that Kim worked died.

As far as we can see, nothing else has to be said about *that* relatives.

Something does have to be said, however, about *that* -less relatives

Weisler (1980) shows that these have rather different properties from relative clauses containing overt relative pronouns like <u>who</u> or <u>that</u>. In particular, they do not stack and must therefore precede any other relative clause in the NP, they cannot appear in extraposed positions, they cannot modify free relatives, and they do not coordinate felicitously with other relatives. We propose that they are simply expressions of category S/NP, and are introduced directly into (lexical) N' rules by the metarule shown below.

(66) N' --> W

       |
       V

      N' --> W, S/NP

Weisler proposes that they be introduced by a rule of the form "NP -> N' S", but this entails introducing determiners (which he does not discuss) as daughters of N', rather than NP, which is an intolerable consequence. Our variant of his analysis avoids this consequence, but retains all the virtues of his proposal. Note that finite closure will ensure that <u>that</u> -less relatives do not iterate.

<u>8. Slash termination, metarule 2</u>

The reader will recall that we have, as yet, only provided half of our account of the bottom of UDC's, namely STM1. Although, as we have seen, a considerable amount follows from STM1, it does not, and cannot, provide an exhaustive account of slash termination. For example, given only what we have presented so far, the grammar will not contain a rule that will permit the following sentence to be generated.

(67) The doctor who we believe worked for Kim is dead.

The reason for this is simple: STM1 is a metarule, metarules apply only to lexical ID rules, the subject of an English declarative is not introduced by a lexical ID rule, therefore it cannot be eliminated by STM1.

(68) Slash termination metarule 2 (STM2)
      ====================================

      @ --> W, $

       |
       V

      @/I --> W, J   where "$ --> I, J" is a nonlexical ID rule.

This says that any rule which introduces a category $ for which the nonlexical rules provide an expansion I, J, has a counterpart in which $ is replaced by J, and the mother has I assigned to be the value of its SLASH feature. Crucially, STM2, like STM1, is a metarule and thus can only apply to rules that specifically introduce a lexical category, i.e. the lexical ID rules. To show the effect it has, we exhibit in (69) half the rules that result from the application of STM2 to the rules in

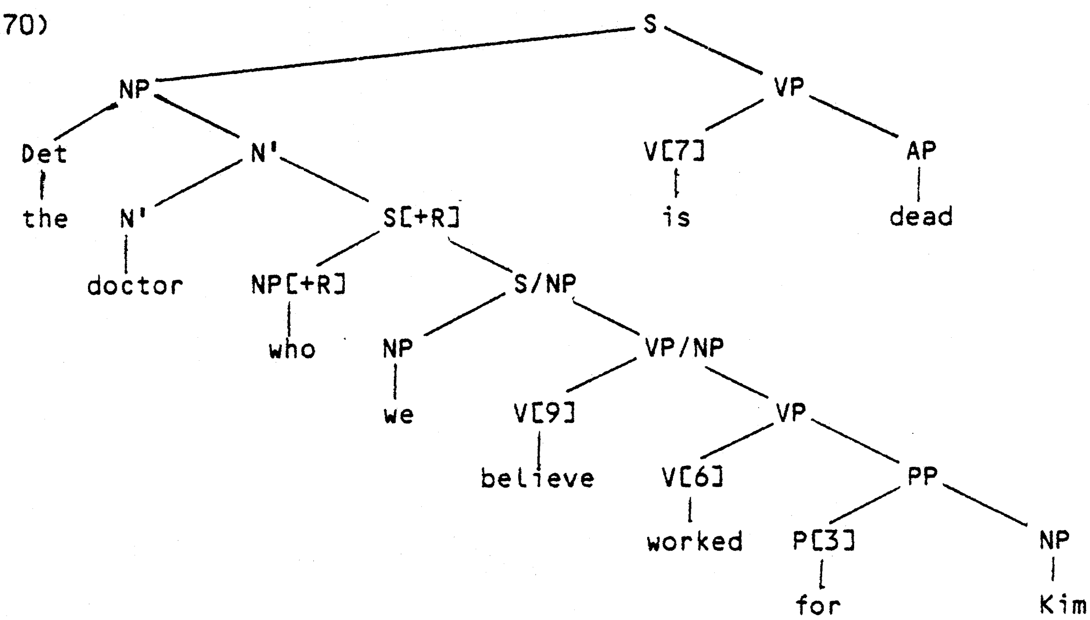
(26ii) when the value of $ is S, and we consider only the expansions of S shown in (54a) and (54b). <23>

(69) VP/NP --> V[9], VP
VP/NP --> V[14], NP, VP
VP/PP --> V[9], VP[there]/PP
VP/PP --> V[14], NP, VP[there]/PP

Note that the VP's introduced will all be tensed since the S's that they replace were tensed, although the informal category notation in (26) and (69) does not show this.

Our grammar now has the rule it needs to generate (67), and will assign it the structure shown in (70).

(70)



And we are now also able to generate examples such as (71).

(71) Kim wondered in which villages Sandy had said were to be found the best wines.

Notice that STM2 will not apply to rules which expand S', since these are not lexical ID rules. Furthermore, it will never be able to replace a daughter S' with a tensed VP, since S' has no expansions involving tensed VP's. These properties of STM2 make a range of interesting predictions. For example, the so-called "complementizer-trace" facts follow immediately. Thus our grammar makes exactly the right claims about the following examples.

(72) a. *Kim wondered in which villages Sandy had said that were to be found the best wines.
b. Kim wondered in which villages Sandy had said that there were to be found the best wines.

(73) a. The man that chased Fido returned.
b. *The man chased Fido returned.
c. The man that Fido chased returned.
d. The man Fido chased returned.

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

(75) a. *The man who I wondered whether chased Fido returned.
b. *The man who I wondered if chased Fido returned.
c. *The man who I wondered chased Fido returned.
d. *The man who I was keen for to chase Fido returned.
e. *The man who I was keen to chase Fido returned.

Some verbs subcategori2e only for S^1, and will not permit a simple tensed S. STM2 cannot interact with the lexical ID rules responsible for introducing such verbs so as to produce rules containing tensed VP's in the result. Thus we predict the grammaticality distribution shown in (76) in the case of a VC1Q3 like <u>regret</u>.

(76) a. Who do you regret that you saw?
b. *Who do you regret you saw?
c. *Who do you regret that saw you?
d. *Who do you regret saw you?

<u>9. Parasitic gaps</u>

Inspect the set of rules derived from those in (26) via instantiation by FFP. This set of rules is shown in (27), above. Consider, in particular, the subset of (27) shown in (77), below.

(77) i. S/NP --> NP/NP, VP/NP
HP/HP ~> Det/NP, N'/NP
VP/NP ~> VP/NP, PP/NP

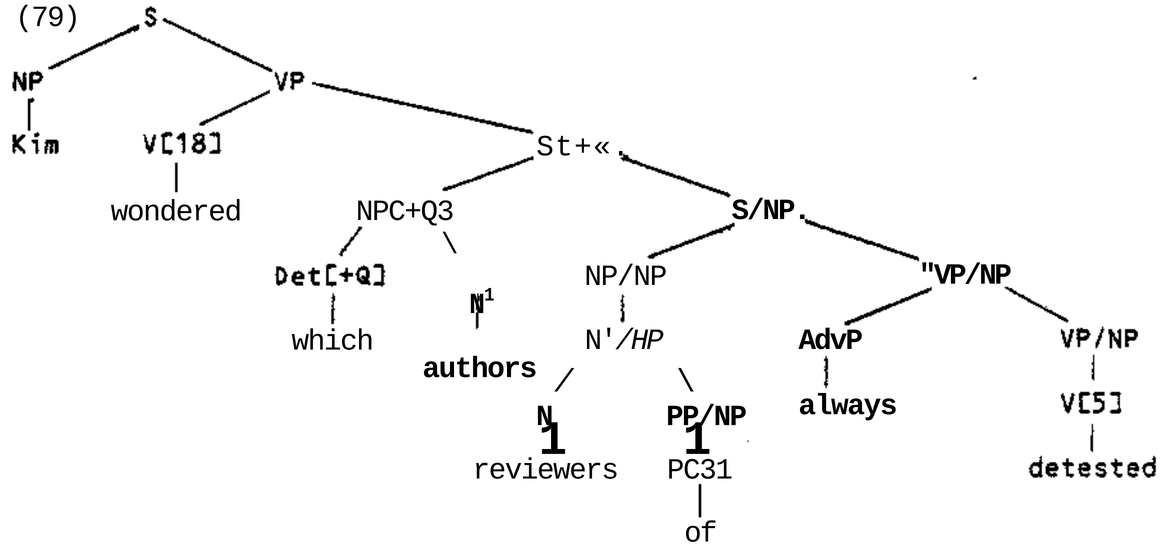
ii. VP/NP —> VC113, NP/NP, PP/NP
VP/NP —> VC12,thereD, NP/NP, PP/NP
VP/NP —> VC133, NP/NP, VP/NP
VP/NP —> VCUD, NP/NP, S/NP
VP/NP —> VC15], NP/NP, SVNP
VP/NP —> VC163, PP/NP, S'/NP

These rules have two things in common: (i) we have made no reference to any of them, nor have we invoked any of them in providing example sentences or trees, and (ii) they all introduce two daughters with a slash feature identical to that of the mother. But they exist, and they have consequences for what the grammar will generate. For example, they predict that the strings in (78) should all be grammatical. <24>

(78) a. Kim wondered which author reviewers of always detested.
b. Kim wondered which models Sandy had sent pictures of to,
c. Kim wondered which dolls house there was a replica of in.
d. Kim wondered which authors the editor wanted reviewers of to please.
e. Kim wondered which authors the editor had told reviewers of that they should pan.

And, somewhat surprisingly, they *are* all grammatical, although (78c) is pretty indigestible, for reasons, we assume, that have more to do with its Escherlike semantics than with its syntax, Thus (78a), for example, will be assigned the structure shown in (79).

(79)



However, as Engdaht (1982) shows, the appearance of these "parasitic gaps" is highly constrained. For example, they only show up in UDC's, and not in virtue of passive, "Equi", or "Raising" constructions, a fact which, as Engdahl points out, supports a distinction between local and nonlocal dependencies, such as the C+/-SLASH3 distinction which is fundamental to the present framework. The grammaticality of the examples shown above follows from the operation of the FFP on SLASH features required by the UDC's involved. Passive, "Equi", and "Raising" constructions are not UDC's, nor can they be properly analysed as such, and thus they cannot, in themselves, legitimate the appearance of parasitic gaps. <25>

Engdahl also provides data bearing on an altogether more subtle constraint illustrated in the examples below, which are taken from her paper.

(80) a. *Who did you say _ was bothered by John's talking to J?
b. *Which slave did Cleopatra give _ to _?

These examples clearly involve UDC's, and yet parasitic gaps are not possible. The theory of UDC's outlined in GP82 and this paper is in the enviable position of not needing to say anything about such examples or, indeed, the examples in (78). The latter will be generated, those in (80) will not. Since this fact is not immediately obvious, we will pursue the topic a little.

Consider the subject gap in (80a). This can only be induced by the rule shown in (81a).

(81) a. VP/NP --> VC91, VP
b. VP/NP ~> VC91, *VP/HP*
c. VP —> VC91, VP

Inspection of (68) will confirm that (81a) is indeed in the grammar thanks to STM2. However, (81a) will not allow us to generate (80a because the VP it introduces is not slashed, and the VP in (80a) must b ^Lashed since it has a gap in it which is not legitimated by any UD internal to the VP. To get (80a), we would need the rule shown in (81b), but this rule is not in the grammar, and cannot be* It cannot arise through instantiation of (81a), given the definition of FFP. It could arise through instantiation of (81c), but the latter is not, and must not be in the grammar. If it was, then we would be claiming that strings like (82) were grammatical standing on their own.

(82) *You said was bothered by Sandy.

Consider the direct object gap in (80b). This can only be induced by the rule shown in (83a) or that in (83b).

(83) a. VP/NP -> VI113, *PP*
b. VP/NP -> VC113, PP/NP
c. VP --> VC11D, *pp*

Inspection of (30) will confirm that (83a) is indeed in the grammar, thanks to STM1. But (83a) will not allow us to generate (80b) because the *PP* it introduces is not slashed, and the *PP* in (80b) must be slashed since it has a gap in it which is not legitimated by any UDC internal to that" PP. To get (80b), we would need the rule shown in (83b), however this rule is not in the grammar, and cannot be. It cannot arise through instantiation of (83a), given the definition of FFP. It could arise through instantiation of (83c), but the latter is not in the grammar. The structurally similar, and lexically intersecting, rule shown in (84) will be in the grammar, of course.

(84) VP/NP -> VC63, PP/NP

And this rule will be responsible for examples such as (85).

(85) Which charities did you give to _?

But this only involves a single gap and is thus not pertinent to the matter in hand.

(86) Which caesar did Brutus imply _ was no good while ostensibly praising _?

Engdahl notes that the example in (86), due originally to Alan Prince, is grammatical, even though, at first sight, it appears to be similar to the ungrammatical (80a). But there is an important structural difference between them, and this difference means that the present analysis correctly predicts the grammaticality of the example. Crucially, the adverbial phrase modifies the VP <u>imply was no good</u>, not the VP <u>was no good</u>. The former is slashed, but the latter is not.

(87) a. VP ~> VP, AdvP
b. VP/NP -> VP/NP, AdvP/NP

The parasitic gap thus gets into (86) in virtue of the rule in (87b), which is itself an instantiated extension under FFP of the rule in (87a) whose presence in the grammar we take to be uncontroversiaL

There is one class of example where the FFP would lead us to expect grammaticality, but where all instances of the type in question appear to be pretty clearly unacceptable, even to those with a high tolerance for parasitic gaps. The construction is that in which a verb subcategorizes for two PP's. In this construction the FFP gives rise to rules which will permit both prepositions to be stranded.

(88) a. *Who did you complain about _ to _?
b. *Who do you seldom talk to _ about _?

Interestingly, Engdahl notes that the second example is acceptable in Swedish. We have no syntactic explanation to offer for the unacceptability of the English examples, and we suspect that Engdahl's hypothesis that the correct explanation is tied into language-particular conditions on bound anaphora is correct.

10. Coordinate structures

At the end of section 3, above, we showed how the principle shown in (89) followed as a theorem from FFP and CRP, given only the self-evident definition of coordinate construction provided in that section.

(89) In a coordinate construction, the mother and all the daughters have identical foot feature coefficients.

Since WH is a foot feature, this principle immediately provides us with an explanation for the contrast noted in (90).

(90) a. Which child and which android did you see?
b. *Which child and my android did you see?

And, since SLASH is also a foot feature, this principle also accounts for the contrasts in (91) and (92)

(91) a. The doctor who Kim worked for and Sandy relied on died.
b. *The doctor who Kim worked for and Sandy relied on Lee died.

(92) a. The doctor for whom Kim worked and Sandy ran errands died.
b. *The doctor for whom Kim worked and Sandy liked died.

Example (91a) is a coordination of S/NP with S/NP which is permissible under (89), whereas (91b) is an attempt to coordinate an S/NP with an S, and this is not consistent with (89). Example (92a) involves the legitimate coordination of S/PP[for] with S/PP[for], whereas (92b) is an illegitimate attempt to coordinate an S/PP[for] with an S/NP. Thus the facts which motivated Ross's (1967) Coordinate Structure Constraint (CSC) and Williams's (1977) Across-the-Board (ATB) Convention can be seen to follow directly as a special case of (89), which itself follows from the FFP and CRP.

More subtle facts also follow from the theory of UDC's outlined here (example (93c) is from Williams (1978:34), and d is due to Paul Hirschbuhler).

(93) a. I know a man who Bill saw and Mary liked.
b. I know a man who saw Bill and liked Mary,
c. *I know a man who Bill saw and liked Mary.
d. I know a man who Mary likes and hopes will win.

Examples a, b, and d, involve the coordination of two S/NP's, two VP's, and two VP/NP's, respectively. All three are consistent with (89). But, given STM1 and STM2, (93c) can only be an attempt to coordinate an S/NP with a VP, or to coordinate a VP/NP with a VP, and neither possibility is sanctioned by (89).

STM1 and STM2 are metarules, and thus, like every other metarule, restricted in their application to just the lexical ID rules. None of the coordination schemata are lexical ID rules, and so none of the following examples can be generated (cf. Gazdar, Pullum, Sag, and Wasow 1982).

(94) a. *I wonder who you saw C_ and _p?
b. *I wonder who you saw [_ and Ca picture of ^33?
c. *I wonder who you saw CCa picture of _3 and _3?

A coordinate structure is one in which every daughter is a conjunct. But, as we observed in section 3, there are idiomatic noncoordinate structures in which conjuncts appear. In such cases, the CRP and FFP will still interact to ensure that the conjunct and its mother carry the same foot features. Thus we predict the grammaticality pattern shown in (95).

(95) a. Which of the wines did you go and buy?
b. Which of the wines did you go to the liquor store and buy?
c. *Which of the liquor stores did you go to and buy wine?

Footnotes

1. This restriction was, as far as we know, first suggested by Henry Thompson for reasons motivated by parsing considerations. Subsequently, and independently, Dan Flickinger and Susan Stucky discussed the possibility of restricting slash termination metarules in this way.

2. Transformational grammar was never able to capture this unitary notion of coordinate construction for reasons that were fundamental to the nature of the theory. This is easily illustrated by reference to examples:

(i) Kim sang and Sandy danced.
(ii) Kim and Sandy met.
(iii) Kim sang and was accompanied by Sandy.

Examples (i) and (ii) would have been, and could only have been directly base generated, but example (iii) had to be derived in a completely different way, via a transformation of conjunction reduction (CR), in any grammar including that handled passive transformationally. Analogous triads of example can be constructed for almost every transformation ever proposed, so the problem is not crucially linked to passives. Conjunction reduction, though rarely formulated, needed to be formulated on a language-particular basis so as to produce structures that were isomorphic to those that would have been produced if everything had been base generated in the first place. Thus obvious and gross generalizations, such as the fact that coordinate VP's participate in the same structural configurations as coordinate NP's and coordinate S's, could only be stipulated by brute force, and were left entirely unexplained.

See Gazdar, Pullum, Sag, and Wasow (1982) for a detailed critique of one recent transformational theory of coordination, and George (1980) for the reductio ad absurdum of another.

In the absence of a formulation, the invocation of CR amounts to problem-naming, not problem-solving. However, the names of familiar would-be transformations are clearly still taken to have some magical explanatory quality, even in 1982: "these subjects are then deleted by CR (however formulated) and the VP's conjoined; the result is the surface forms", an approach which, we are assured, "accounts straightforwardly for the facts" (Rognavaldsson 1982:560). Given the history of the field, there is a sad irony in the fact that expressions that were once the names of rules are now no more than elements in a taxonomy of construction-types.

3. See Langendoen & Postal (1982, chapter 4, section 1) for an important recent discussion of coordinate constructions. Our definition of conjunct (see (3), below) is essentially identical to theirs.

4. Throughout the paper "@" is to be read as "alpha", and "$" as "beta". The integers 0 and 1 attached to @ and $ below are to be interpreted as subscripts.

5. The need to stipulate the bar coefficient identity in (6) is unfortunate, but it doesn't follow from anything else (e.g. the Head Feature Convention (HFC), the Control Agreement Principle (CAP), or the Foot Feature Principle (FFP), or the Conjunct Realization Principle (CRP) - for which see (18), below).

6. The kind of semantics appropriate to coordination schemata of this general kind has been discussed in a number of recent works, Keenan and Faltz (1978), Gazdar (1980), Cooper (1979), Partee and Rooth (1982), Rooth and Partee (1982), and Bergmann (in press).]

7. Paul Postal has drawn our attention to the fact that this second point, though a relatively minor one in itself, has an important moral. An uninterpreted formalism does not make any claims about anything. Since the work of Montague, many linguists have realized that an uninterpreted semantic formalism makes no claims about meaning (although recent discussion under the rubric of 'logical form' reveals, rather too clearly, that not all members of the community are yet privy to the insight). And it is just as true that an uninterpreted syntactic formalism makes no claims about grammatical structure.

8. <u>But</u> cannot be used to coordinate [+N, -V] categories (*<u>Kim but Sandy stuttered</u>). There are other facts of this kind, for example, <u>both</u> cannot be used to coordinate full sentences that lack complementizers (*<u>Both Kim sang and Sandy danced</u>). These idiosyncrasies are not relevant to our present concerns, and can be handled quite straightforwardly with feature cooccurrence restrictions.

9. The proviso "flat" is necessary since, obviously, there are three- conjunct constructions that involve two two-conjunct constructions, one embedded within the other. These have different semantic and intonational properties from the flat construction. We ignore them here, although (7) - (9) allow them to be generated, of course, as we would want.

10. This principle only makes sense when construed as an additional clause to the definition of <u>instantiated extension</u> given as (91) in GP82. The theory of coordination which it expresses is that of Sag, Wasow and Weisler (1982).

11. The mother VP has to be either past or nonpast in this construction, in contrast to the case of coordinate VP's, because its lexical head can only be one or the other, and the HFC requires the tense to be identical to that of its lexical head.

12. This notion of unification, which is crucial to the definition of the FFP in GP82, originates with Kay (1979), a paper which, regrettably, GP82 failed to cite.

13. For ease of subsequent exposition, we have divided the grammar in (26) into the lexical ID rules in (i), and the nonlexical ID rules in (ii). This distinction is irrelevant to feature instantiation, of course, but becomes relevant when we discuss certain metarules, below. Note that we are not taking Det to be a lexical category. For convenience, we show the rules in (26) in a partly instantiated form - the various heads, lexical and nonlexical, are specified. And

we use S' to abbreviate S[+that], and S to abbreviate S[-that].

14. In exhibiting (27), we have deliberately suppressed all rules that assign a coefficient to SLASH on a lexical category. These rules, if allowed to exist, are irrelevant to the grammar since slashed lexical categories necessarily have no expansions (i.e., there will not be any rules expanding a V/NP, for example). If conceptual hygiene dictates that these useless, but innocuous, rules be eliminated, then an FCR of the form [+SLASH] -> [-LEXICAL] is all that is necessary.

15. This was pointed out by a number of people independently: Maling and Zaenen (1982:279), Peters (p.c.), Sag (1982b:332), and Williams (1981:650). For some discussion of the latter, see Gazdar, Pullum, Sag, and Wasow (1982).

16. Notice that this formulation of STM1 does not introduce a phonologically null item. However, it can easily be modified to do so, and there may be phonological and semantic grounds for making such a modification. We leave the issue open here.

17. As formulated, STM1 will also give rise to rules such as "VP/V[5] --> NP", i.e. rules which have a lexical category as the coefficient of the mother's slash. In the absence of any rules introducing such categories, these rules are innocuous and can, and will, be ignored. As in the converse case, mentioned above, they can be legislated out of existence, either by an FCR, or by stipulating [-LEXICAL] in STM1. Actually, the restriction is almost certainly tighter than this suggests, since mid- bar categories like N' and A' also never seem to appear as coefficients of slash. But capturing the relevant generalization entails taking a stand on the problematic issue of the X-bar status of VP, S, and S' (see Gazdar, Klein, and Pullum (1983:2-5) for some discussion of this issue).

18. The PP in (41) needs to be featurally restricted in some way, presumably to locative, directional, and, as Pat Griffiths pointed out to us, temporal phrases. The precise nature of the restriction is not relevant to our present concerns. The construction has received intermittent attention in the literature, in particular, see Green (1977) and the references therein.

19. Example (43) is due to Postal ???

20. If this restriction is a fact about languages, as opposed to simply a fact about English, then it would be more elegantly expressed by modifying the syntax of the feature FOOT so as to prevent any possibility of both the WH and SLASH coefficients appearing at once.

Example (51) would have been blocked by the generalized left branch condition (GLBC) of Gazdar (1981). But the GLBC cannot be reconstructed in the version of GPSG assumed here for two reasons. It was a constraint on the slash introduction metarule, and that no longer exists. And it made reference to leftmost position, which is impossible given the ID format for rules. But the GLBC was, in any case, both too weak and too strong. The aspect of it which blocked (51) would also have blocked (78a), below, and the latter is clearly grammatical. Furthermore, its statement in positional terms failed to generalize to forbidden subject dependencies in VSO sentences, as pointed out for Welsh by Stephen Harlow (p.c), and for Polish by Borsley (in press).

A residual problem for the current analysis is the absence of an explanation for the impossibility of preposition-stranding in PP's which are left branches, as illustrated in the following examples:

(i) *Which fountain did you say that near sat a large purple gorilla?
(ii) *Who did you say that to , Kim had given the books ?

Contrast these with (iii), which appears to be grammatical/ given the appropriate stressing:

(iii) Which book did you say that reviews of , Sandy was collecting ?

21. For some relevant discussion of the NP-VP analysis, see Gazdar (1981:171-172, n22, n23), Dowty (1982:115), Chung and McCloskey (1982), and Fodor (in press).

22. As it stands, our syntax provides us with no explanation for the acceptability contrast illustrated in (i) and (ii), below.

(i) Whose picture of whom amused Sandy?
(ii) *The doctor whose picture of whom amused Sandy has died.

23. The other rules involve VP's appearing as the SLASH feature of the output's mother. These rules are unlikely to be able to play any role in the grammar of English in view of the virtual absence of rules introducing VP's into UDC's. Only topicalization permits their appearance since they cannot carry the WH features required in almost all other UDC's, and even here there are heavy restrictions - for example, the topicalized VP may not be tensed. See Gazdar, Pullum, and Sag (1982) for some precise proposals regarding these restrictions.

The effect of STM2 with respect to the S expansion rule (54c), is to define a whole set of rules all of which are already defined by STM1, which is why we do not exhibit them in (69). The fact that some rule is legitimated by more than one metarule has no consequences whatsoever.

24. The existence of examples of this kind was noted by Ross (1967), but they were almost entirely neglected from then until Engdahl's important (1982) paper which thoroughly explored both the facts and their theoretical implications. Most of the examples below are adapted from ones to be found in her paper. An earlier GPSG analysis of her data is to be found in Sag (1982a). We assume here the correctness of Sag's arguments against Engdahl's claim that there are grammatical differences with respect to island constraints between the two single-gap-containing subconstructions in a (noncoordinate) double gap structure. We also assume that the hierarchy of acceptability for such constructions, as outlined by Engdahl, is not to be explained in the syntax itself. Accordingly, we do not concern ourselves with ''degrees of acceptability[1]'' in what follows.

25. Although, of course, such constructions may be embedded in UDC's, and thus have parasitic gaps within them in virtue of the UDC in which they appear, as in the following example.

(i) Which mafiosi do you expect relatives of
to eliminate?

Bergmann, Merrie (in press) Cross-categorial semantics for conjoined common nouns. To appear in <u>Linguistics and Philosophy</u>,

Borsley, Robert (1982) A note on the Generalized Left Branch Condition, To appear in <u>Linguistic Inquiry</u> 14.1.

Brame, Michael (1981) Trace theory with filters vs, lexically based syntax without. <u>Linguistic Inquiry</u> 12, 275-293,

Chomsky, Noam (1977) On wh-movement. In P.W. Culicover, T« Wasow, and A, Akmajian (eds.) <u>Formal Syntax</u>, New York: Academic Press, 71-132.

Chung, Sandra, and James McCloskey (1982) On the interpretation of certain island facts in GPSG. Mimeo, UCSD.

Cooper, Robin (1979) Model theory for a fragment of English syntax. Unpublished manuscript. University of Wisconsin at Madison.

Corbett, Greville (1983) Resolution rules: agreement in person, number, and gender. In G. Gazdar, E.H, Klein, and G.K. Pullum (eds.) Order, Concord and <u>Constituency</u>, Foris Publications, Dordrecht, 1

Dowty, David (1982) Grammatical relations and Montague Grammar. In P. Jacobson & G.K, Pullum (eds,) <u>The Nature of Syntactic Representation</u>. D. Reidel, Dordrecht, 79-130,

Engdahl, Elisabet (1982) Parasitic gaps. To appear in <u>Linguistics and Philosophy</u>,

Fodor, Janet (1982) <u>Parsing, constraints, and the freedom o£ expression</u>. Montgomery, Vermont: Bradford Books, to appear.

Fodor, Janet (in press) Phrase structure parsing and the island constraints. To appear in <u>Linguistics and Philosophy</u>.

Gazdar, Gerald (1980) A cross-categorial semantics for coordination. <u>Linguistics and Philosophy</u> 3, 407-409.

Gazdar, Gerald (1981) Unbounded dependencies and coordinate structure. <u>Linguistic Inquiry</u> 12, 155-184.

Gazdar, Gerald (1982) Phrase structure grammar. In P. Jacobson & G.K. Pullum (eds,) <u>The Nature of Syntactic Representation</u>, 0. Reidel, Dordrecht, 131-186.

Gazdar, Gerald, Ewan Klein, *and* Geoffrey Fulium (1983) Chapter 1: Introduction. In G. Gazdar, E,H. Klein, and G.K, Pullum (eds,) Order, <u>Concord</u> and <u>Constituency</u>. Foris Publications, Dordrecht, [illegible]

Gazdar, Gerald, and Geoffrey Pullum (1980) <u>There</u> is there, <u>York Papers in Linguistics</u> 8, 199-200.

Gazdar, Gerald, and Geoffrey Pullum (1982) Generalized phrase structure grammar: a theoretical synopsis, Mimeo, Indiana University Linguistics Club, August 1982.

Gazdar, Gerald, Geoffrey Pullum, and Ivan Sag (1982) Auxiliaries and related phenomena in a restrictive theory of grammar. Language 58, 591-638.

Gazdar, Gerald, Geoffrey Pullum, Ivan Sag, and Tom Wasow (1982) Coordination and transformational grammar. Linguistic Inquiry 13, 663-676.

George, Leland (1980) Analogical generalizations of natural language syntax. PhD dissertation, MIT.

Green, Georgia (1977) Do inversions in English change grammatical relations. Studies in the Linguistic Sciences 7.1, 157-181.

Grosu, Alexander (1974) On the nature of the Left Branch Condition. Linguistic Inquiry 5, 308-319.

Hopcroft, John, and Jeffrey Ullman (1979) Introduction to automata theory, languages, and computation. Reading, Ma.: Addison-Wesley.

Iwakura, Kunihiro (1980) On wh-movement and constraints on rules. Linguistic Analysis 6, 53-95.

Kay, Martin (1979) Functional grammar. In C. Chiarello et al. (eds.) Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society. Berkeley: BLS, 142-158.

Keenan, Edward, and Leonard Faltz (1978) Logical types for natural language. Los Angeles: UCLA Occasional Papers in Linguistics 3.

Langendoen, Terence (1976) On the weak generative capacity of infinite grammars. CUNYForum 1, 13-24.

Langendoen, Terence, and Paul Postal (1982) The vastness of natural languages. Forthcoming.

Maling, Joan, and Annie Zaenen (1982) A phrase structure account of Scandinavian extraction phenomena. In P. Jacobson & G.K. Pullum (eds.) The Nature of Syntactic Representation. D. Reidel, Dordrecht, 229-282.

Partee, Barbara, and Mats Rooth (1982) Generalized conjunction and type ambiguity. To appear in A. von Stechow et al. (eds.) Meaning, Use, and Interpretation. Berlin: de Gruyter.

Postal, Paul (1977) About a "nonargument" for raising. Linguistic Inquiry 8, 141-154.

Rognvaldsson, Eirikur (1982) We need (some kind of a) rule of conjunction reduction. Linguistic Inquiry 13, 557-561.

Rooth, Mats, and Barbara Partee (1982) Conjunction, type ambiguity, and wide scope "or". In D. Flickinger, M. Macken, and N. Wiegand (eds.) Proceedings from the First Annual West Coast Conference on Formal Linguistics. Stanford: Stanford Linguistics Department, 353-362.

Ross, John Robert (1977) Constraints on variables in syntax. PhD thesis, MIT.

Sag, Ivan (1982a) On parasitic gaps. In D. Flickinger, M. Macken, and N. Wiegand (eds.) *Proceedings from the First Annual West Coast Conference O£ Formal Linguistics*. Stanford: Stanford Linguistics Department", 35-46.

Sag, Ivan (1982b) Coordination, extraction, and generalized phrase structure. *Linguistic Inquiry 13*, *329-336*.

Sag, Ivan, Thomas Wasow, and Steven Weisler (1982) Coordination and how to distinguish categories. Manuscript, Stanford University.

Weisler, Steven (1980) The syntax of *that*- less relatives. *Linguistic Inquiry* 11, 624-631.

Williams, Edwin (1977) Across-the-board application of rules. *Linguistic Inquiry* 8, 419-423.

Williams, Edwin (1978) Across-the-board rule application. *Linguistic Inquiry* 9, 31-43.

Williams, Edwin (1981) Transformationless grammar. *Linguistic Inquiry* 12, 645-653.