In neither case can the derivation proceed any further, as the last two items \textit{(have and been)} cannot combine with their complements, nor can the complements combine with the S-node. The model permits the fronting of participial phrases as instances of topicalization, analogous to the fronting of Object NP (26c) or an Adverb (28c)\footnote{7}:

(33) (a) (She said she would be standing at the bus-stop, and)

\begin{align*}
\text{Standing at the bus-stop} & \quad \text{she was} \\
\text{Cing} & \quad \text{S/Cing} \\
\text{S} & \quad \text{B}
\end{align*}

(b) (I can't believe I've eaten the whole thing, but)

\begin{align*}
\text{Eaten the whole thing} & \quad \text{I have} \\
\text{Cen} & \quad \text{S/Cen} \\
\text{S} & \quad \text{B}
\end{align*}

3.2.3. \textit{Preposing from VP.} In Section 2.3 the category of the verb \textit{put} was assumed to be VP/PP/NP, as in:

(34) \begin{align*}
\text{Put} & \quad \text{the frog} \\
\text{on the table} & \quad \text{NP} \\
\text{VP/PP/NP} & \quad \text{PP} \\
\text{VP/PP} & \quad \text{F} \\
\text{VP} & \quad \text{F}
\end{align*}

The combination of \textit{put} with Tense by Affix Cancellation yields S/PP/NP/NP (22d). The model permits either the NP or the PP to be preposed:

(35) (a) On the table \text{he} \quad \text{put} \quad \text{the frog}

\begin{align*}
\text{PP} & \quad \text{NP} \\
\text{S/PP/NP/NP} & \quad \text{NP} \\
\text{S/PP/NP} & \quad \text{B} \\
\text{S/PP} & \quad \text{F} \\
\text{S/PP} & \quad \text{F} \\
\text{S} & \quad \text{B}
\end{align*}

4. \textbf{The Left Branch Condition and the Root Transformation Hypothesis}

4.1. \textit{Removing Overgeneralizations and Distinguishing Clause Functions}  

The account presented above is intended to suggest that the basic facts of English main clause constructions can be made to follow from just four very restricted rule schemata, together with an order-free categorial lexicon. However, there are a number of ways in which the model as it stands overgeneralizes, and which might seem to cast doubt upon the claim. Its acceptance of \textit{*her must he love, *love her must he}, and of sentences violating DO-support have already been mentioned (cf. Sections 3.1 and 3.2.1). A further class of overgeneralisations concerns sentences involving 'double topicalisations', such as the following, all of which are accepted by the model as it stands:

(b) This frog \text{he put} \quad \text{on the table}

\begin{align*}
\text{NP} & \quad \text{S/PP/NP} \\
\text{B} & \quad \text{PP} \\
\text{S/PP} & \quad \text{F} \\
\text{S} & \quad \text{B}
\end{align*}

The model also allows fronting from within a PP ('Preposition Stranding', a topic which will be further discussed below):

(36) This table \text{he put the frog} \quad \text{on}

\begin{align*}
\text{NP} & \quad \text{S/PP/NP} \\
\text{NP} & \quad \text{PP/NP} \\
\text{S/PP} & \quad \text{F} \\
\text{S/NP} & \quad \text{FP} \\
\text{S} & \quad \text{B}
\end{align*}

Finally, the model rules out any sentence where the PP occurs before the NP Object (except when the PP is topicalized to sentence-initial position)\footnote{12}:

(37) \text{*He put} \quad \text{on the table} \quad \text{the frog}

\begin{align*}
\text{S/PP/NP} & \quad \text{PP} \\
\text{NP} & \quad \text{F}
\end{align*}
(38) (a) *On the table a frog she put
(b) *In the garden apples he ate.
(c) *The pink one here I gave.

(The obvious extension to complex sentences with sentential complements that we discuss later would if nothing were done provide an opportunity for even more grotesque multiple topicalisations, like *the pink one her Fred Susan I said thought gave.) There are two sorts of question that are posed for any theory by such overgeneralisations. First, and most generally, does their exclusion involve any increase in the power of the system? (in particular, does it require negative constraints?) Secondly, how are the necessary restrictions to be explained?

As in the discussion of DO-support (cf. Note 13) we shall attempt no more in the present paper than to show that the phenomena can be handled without any extension to the power of the model, without claiming that the proposals offered here are correct or explanatory. One solution to the problem is to provide specific instances of the combination rules, which attach features identifying the derivational history of sentences as being +/- topicalised, +/- inverted, +/- Wh-fronted, and so on. The only effect of the elaboration is to slightly restrict the class of sentences that the model accepts, and to capture the different functional categories of clausal constructions within the grammar. It leaves the class from which the rules are drawn, and the power of the grammar, essentially unchanged.

For example, let us assume that partially complete sentences which have already found their subjects, like [He must$_1$$_{NP}$, are distinguished from those like [must$_1$$_{NP}$$_{VP}$ by a feature +/-subject – a feature that we take to be related to the phenomenon of subject-verb agreement. Let us further assume that sentences which are topicalised are distinguished from those which are not by a feature +/-topic. The lexical category for the tense morpheme will carry the features –subject and –topic, since it has yet to combine with either, and these values will be inherited by the tensed verb via the Affix Cancellation rule. We can then define an instance of the Backward Combination rule as follows:

\[
\text{X SS/X} \rightarrow \text{S S}
\]

\[
\text{[+subject] [topic]}
\]

(39) expresses the fact that topicalised sentences differ in functional category from non-topicalised ones. It also expresses the fact that sentences with more than one topic are ungrammatical, for the rule applies only to S nodes which have not previously been marked by the same rule as +topic. The rule therefore excludes all of the strings (38).

Assuming that Auxiliary verbs are distinguished as such in the lexicon, the DO-support phenomena can be captured in a similar way, by replacing the single Forward Combination schema by a number of instances, one of which only allows a tensed verb to find its subject by Forward Combination when it is marked as an Auxiliary. This apparatus could also be extended to include features identifying Yes/No and Wh-questions. Similar instances of the general rules could then be provided to rule out main clauses like *Where he saw a tree and *Over there did he see a tree, and complements like *apples I know eats Mary. Such a modification is equivalent to introducing (for example) a rule schema specifying constituent order for the relative clause into a PS grammar (cf. Gazdar, 1981a). Nothing in the account that follows hinges upon such details. We shall therefore pursue the matter no further, but simply assume that some such feature system could be devised to exclude these overgeneralisations of the basic grammar, without any extension to its intrinsic power, or any loss of the generalisations about the form of those specific rules that the basic theory captures.

4.2. The Left Branch Condition

Ungrammatical strings like *love he must her (26w) were ruled out by the grammar because the main verb is separated from its complement and because backwards combination is restricted to combining functions with arguments of the form Y rather than Y$. This same absence of a 'Backwards Partial Combination' rule in English also rules out the following sentences, which are of a kind which has forced the introduction of constraints like the A-over-A Principle (Chomsky, 1964), the Left Branch Condition (Ross, 1967) and the Relativised A-over-A principle (Bresnan, 1976).

(40) (a) *Whose did you peel grapes

\[
\text{NP/N} \quad \text{S/NP} \quad \text{N}
\]

(b) *How strong is she a woman

\[
\text{NP/NP} \quad \text{S/NP} \quad \text{NP}
\]

On the assumption that whose in whose grapes is NP/N, any derivation of (40a) blocks because whose is not adjacent to grapes, and because Backward Combination cannot pick up incomplete objects like NP/N. On the assumption that the category of How strong in How strong a woman is NP/NP, (40b) blocks in the same way.
4.3. The Root Transformation Hypothesis

A number of similar phenomena that Emonds (1976) adduces in support
of his two major constraints on Root Transformations are also accounted
for by the model as it stands. For example, the following ungrammatical string was ruled out by Emonds on the grounds that it
would involve two Root Transformations, VP Preposing and Topical-
ization, applying in a single clause.

(41) (a) *Apples eat I may

As shown in example (26s), such strings are ruled out because verbs
must combine with their complements by Forward Combination, and because [I may]_{snp} can only backward-combine with a complete VP.
The following instances where two Root Transformations have applied,
are ruled out for exactly the same reasons:

(41) (b) VP Preposing and Directional Adverb Preposing:

*Into the house run they did

(c) Topicalization and Adverb Preposing:

*The garden in we found the apples

(d) Topicalization and Directional Adverb Preposing:

*This house into they ran

(The sentences in (38) might also be considered as instances where two
Root Transformations have applied. They were ruled out by our stipula-
tion that sentences can have only one topic, expressed in the restricted
version of the Backward Combination rule (39).)

Emonds' second major constraint on Root Transformations was that
they could not apply in subordinate clauses. Without going into the
question of how relative clauses attach to their matrices, it is simplest to
assume that English subordinate clauses are constructed by the same
rules as main clauses. Assuming that relativisation is another instance of
topicalization, the following non-sentences which in part motivated
Emonds' constraint are ruled out for precisely the same reasons as their
main clause analogues (38, 41).

(42) (a) *This is the table on which the frog he put

(b) *This is the pink one which her I gave

(43) (a) *These are the apples which eat I may

(b) *This is the house into which run they did

(c) *This is the garden in we found the apples

(d) *This is the house which into they ran

However, the following sentence, which appears to allow Directional
Adverb Preposing in a subordinate clause, contradicts Emonds' hypo-
thesis, but is correctly predicted to be grammatical by the present
model:

(44) This is the rock from which springs the Ganges

Emonds' hypothesis also excludes cases where the Root Transformation
has applied in a subordinate clause that is not a relative, but is for
example a verbal complement:

(45) *He realised did John come for dinner.

It seems reasonable to suppose that the anomaly arises because of the
subcategorisation of the verb realise and the functional category of the
subordinate sentence, both of which can be expressed in the feature
system outlined in Section 4.1.

5. Extraction phenomena

5.1. Noun-phrase Constraints

In Section 2 it was argued on both linguistic and psychological grounds
that incomplete entities like *[He must love]_{snp} must have a place in the
grammar. However, no similar argument could be advanced for the
reality of entities like *[He must love the]_{snp}. English does not allow
preposings like

(46) *Woman he must love the.

As it stands the Forward Partial Combination Rule will allow the formation
of such a construction by partially combining *[He must love]_{snp} and
*[the]_{snp}. What is needed is a way of restricting the Forward Partial rule to the
combination of 'verb-like' argument functions.

In fact, the device of restricting rules to apply only to certain sorts of
category has been implicit all along, in the proviso that Backwards
Combination apply only with functions yielding S, and in the restriction of
Affix Cancellation to affixed verbs. The required restriction on the Forward
Partial Combination rule differs only in that it refers to a whole class of
node-types, rather than to a single type. In order to impose this restriction
we adopt the following tenets of the X-bar hypothesis (Chomsky, 1970;
Jackendoff, 1977):

(47) (a) All the basic category symbols, both terminal and non-terminal,
are considered to be associated with a bundle of grammatical
features.
The major categories Nominal, Verbal, Adjectival and Prepositional are considered to be associated with specific combinations of the major features +/− N and +/− V.

All constituents are considered to be of the form
\[ X^* = (\text{Specifier}) \ X^{*−1} \text{ Complement} \]
where \( X \) is a major category, and \( X^{*−1} \) is termed the 'head' of the constituent \( X^* \).

Under this notation, for example, the noun phrase might be regarded as having the determiner as specifier, the noun as head, and a relative clause or prepositional phrase as an (optional) complement.

It will be noted that this notation draws a distinction between two varieties of function category which has not hitherto been made in the present theory, in which both heads and specifiers are merely functions of the form \( X^*/Y \). The difference between them is expressed in (47c) above, which makes it clear that while 'Head' functions, such as nouns and verbs, are already associated with major features, 'Specifier' functions are not. On the contrary, when specifiers combine with their head they yield an entity which bears the features of that head. We therefore make the following additional stipulation, which is already implicit in the \( X \)-bar system:

\[
\text{(48)} \quad \text{No combination, and in particular no partial combination, can apply where the argument term does not bear major features.}
\]

It follows that there can occur no partial combination of specifiers which have not yet found their heads, and therefore that specifiers (a class which, with Jackendoff (1977), we take to include determiners, complementisers and affixes) cannot 'strand', unlike heads such as prepositions and verbs. Thus, for example, \([\text{He must love}]_{\text{SNP}}\) and \([\text{the}]_{\text{NP/N}}\) cannot combine because the determiner is a specifier and bears no major features. A slightly more stringent restriction upon Forward Partial Combination will enable us to capture several further constraints on extraction.

5.2. The NP Constraint

Besides the general constraint that functions can only combine with arguments that bear major features, we can further specify which major features a particular rule allows, just as we restricted the functions to which the Backwards Combination rule could apply. The following restriction upon the Forward Partial Combination rule will rule out sentences that would violate the NP constraint of Horn (1974) and Bach and Horn (1976).

\[
(49) \quad \text{Forward Partial Combination}
\]
\[
X^*/Y \quad Y^*/Z \Rightarrow X^*/Z
\]

Consider, for example:

\[
(50) \quad *\text{This man I burned a book about.}
\]

The construction would be possible if \( I \text{ burned a book about} \) could be formed into an incomplete entity of type S/NP. There are two possible views of the internal structure of Complex NPs like a book about this man, but on neither will the revised form of Partial Combination allow the extraction. If the NP is regarded as \([\text{a}]_{\text{NP/N}} \quad \text{book about this man}]_{\text{NP}}\) then it is blocked because \( \text{I burned}]_{\text{SNP}} \) cannot combine with \([\text{a}]_{\text{NP/N}}\). If the NP is \([\text{a book}]_{\text{NP/PP}} \quad \text{about this man}]_{\text{NP}}\), then it is blocked because \( \text{I burned}]_{\text{SNP}} \) cannot partially combine with \([\text{a book}]_{\text{NP/PP}}\). Both combinations are excluded because NP naturally does not bear the feature +N. (Following Bach and Horn, 1976, we take it that in one reading of I wrote a book about this man, the PP is dominated by VP, not by NP. That is to say that in terms of the present theory, one lexical entry for write identifies it as VP/PP/NP. Hence the acceptability of This man I wrote a book about, analogous to (36).)

Extraction from other complex NPs is blocked for just the same reason as in (50): they are not dominated by a node bearing the feature +N. For example:

\[
(51) \quad (a) \quad *\text{This man I met a girl who knew}
\]
\[
\begin{array}{c|c|c|c|c|c}
\text{NP} & \text{S/NP} & \text{NP/N} & \text{N/S} & \text{S/NP} \\
\hline
\end{array}
\]

(b) *This man I met a girl who knew
\[
\begin{array}{c|c|c|c|c|c}
\text{NP} & \text{S/NP} & \text{NP/S} & \text{S/NP} \\
\hline
\end{array}
\]

5.3. Preposition Stranding

If Prepositional Phrases are included in the class of −N constituents and Prepositions are their heads, as is implied by both Chomsky (1970) and Jackendoff (1977), then the grammar also expresses the fact that extraction from within PP in post verbal position is allowed in English.

\[
(52) \quad \text{This table he put the frog on (cf. example 36).}
\]

It is worth pointing out that the model does not require the node PP to be characterised as extractable per se. It has already been shown how other
general features of the model for English rule out preposition stranding in PPs that have already been fronted (41c, 43c). The account appears to have a descriptive advantage over theories that label PP as a bounding node in English to explain the ill-formedness of (41c, 43c), but which must then invoke an ‘escape zone’ hypothesis to explain the acceptability of preposition stranding in post verbal position. (Cf. Koster, 1980).

The suggestions advanced above are broadly related to the idea of a ‘projection path’ (Koster, 1978). As has been noted by Fodor (1980), the Forward Partial rule determines how right-branching structures are built up, while the restrictions upon the types of categories to which it can apply in a given language does the work of Subjacency constraints and principles of Government (cf. Koster, 1980), in determining which nodes are on the projection path.

5.4. Extraction from Sentential Complements

On the assumptions (a) that one lexical category of the verb believe is VP/S, and (b) that S nodes, like all verbal nodes, carry the feature – N, as implied by the X-bar hypothesis, then it follows that extraction will be possible from both subject and object positions in simple sentential complements, as follows:

(53) (a) Who(m) do you think he loves?

```
   NP  S/S  S/NP
       S/NP  FP
       S    B
```

(b) Who do you think loves him?

```
   NP  S/S  S/NP/NS  NP
       S/NP/NS  FP
       S/NP    B
       S      F
```

On the assumptions that the category of that identifies it as a function which we might write as Cthat/S, which maps sentences onto some sort of complement Cthat, and that believe is also categorised as combining with complements of category Cthat, i.e. as VP/Cthat, then extraction from object positions in that complements can proceed more or less as in 53a):

(54) Who(m) do you think that he loves?

```
   NP  S/Cthat  Cthat/S  S/NP
       Cthat/NP  FP
       S/NP  FP
       S    B
```

(Do you think may combine with that he loves, just as with he loves. But, given the stipulation (48) and the form of the Partial Combination rule (49), it cannot combine with the specifier that until the latter has ‘inherited’ some features from its head he loves, and in particular the feature – N which all verbal categories including S bear.)

It remains to be explained why extraction from subject position in that complements is not allowed in English (Chomsky and Lasnik, 1977).

(55) *Who do you think that loves him

```
   NP  S/Cthat  Cthat/S  S/NP/NS  NP
       S/NP/NS  FP
       S/NP    B
       S      F
```

In Section 4, certain features were introduced to express the different functional meanings of the various clause constructions. One of these was a feature +/− subject, which distinguished S-nodes that had already found their subjects from those that had not. A simple solution to the above problem is to further stipulate that the word that is categorised as a function over sentences which bear the feature + subject – that is as Cthat/S/+subject. Given this categorisation, that cannot combine with [loves]S/NP/NP, because the latter is still − subject. The derivation can therefore proceed no further.

The proposal outlined above is once again intended mainly to show that the present system can be made to obey the constraints upon ‘movement’ that have been identified by transformationalists, without the inclusion of any negatively stated rules. Whether or not this particular account will stand the test of time is another question, and depends of course on whether further reasons can be produced to justify the mysterious feature +/− subject, and the categorisation of the morpheme that with respect to it. It is, however, suggestive that in some heavily
inflected languages, like Italian and Spanish, where the finite verb inflection is very similar to a personal pronoun, the equivalent of (55) is grammatical (Perlmutter, 1968; Chomsky and Lasnik, 1977, p. 450f). In such languages one might suspect that the finite verb inflection already bears the feature + subject.

6. Conclusions

There is no doubt that the model presented here is incomplete. Many important categories, particularly negation and the adverbials, have been entirely ignored, and the treatment of Tense and the affixes is certainly inadequate. It also remains to be seen how the many constructions that have been ignored here are to be accommodated within the framework that has been outlined. However, the fact that a standard categorial lexicon, plus the four rule schemata, seems to come close to exhaustively specifying the main clause constructions of English, and also seems to explain a number of major constraints on transformations, encourages us to compare the theory with certain alternatives, and to examine its broader implications.

6.1. General Linguistic Properties

In this model, the lexicon has the familiar role of specifying both the categories of words and the entities that they may co-occur with, as in other generative frameworks. However, because it is merely a projection of semantic dependencies, it is free with respect to the left-right order of functions and arguments, and to that extent also constitutes the equivalent of an order-free base. If any order is allowed by the dependencies and the constraint that none of the combination rules operates on other than adjacent entities, then the dependencies will also potentially allow the mirror image order. For any given language, it is the specific form of the combination rules that further determines which of the orders allowed by the lexicon are grammatical in that language.

The model may be contrasted in this respect with other theories in which semantic dependencies are specified not only in the lexicon, but also in syntactic base rules and/or the semantic rules (which Brame (1976, 1978) and Heny (1979) have repeatedly pointed out all look suspiciously alike). Similarly, the present model is to be contrasted with others which divide the task of specifying constituent order between different components of the grammar, such as an orthodox PS base and either a transformational component or a number of metarules which are similarly extrinsic to the grammar and the corresponding semantics and to the processor. While the reallocation of duties may be a 'notational variant' of these other theories, it seems to be one that is on the side of parsimony, and one which effects considerable savings in the number of different varieties of rules the system includes. The clear separation between surface ordering and semantic dependency may also offer the theorist a freer hand to approach linguistic universals. As relational grammarians and others have frequently pointed out, these are notably resistant to analysis with an ordered base (cf. Cole and Sadow, 1977; Venneman, 1973).

6.2. Constraints on 'Movement'

A grammar ought to need no more than the addition of a mechanism for ambiguity resolution in order to turn it into a theory of performance. Given this requirement, negative rules to constrain movement have a very problematic status indeed, as was pointed out earlier. Their absence from the present theory is therefore a desirable feature.

There are also advantages from a linguistic point of view in handling the constraint phenomena as the present theory does. All extraction constraints are reduced to a single underlying mechanism. Extraction is possible if and only if a structure can be 'penetrated' by Partial Combination. Thus the Left Branch Condition and many constraints upon Root constructions hold in English because it includes no 'Backwards Partial Combination' rule. Similarly, the noun-phrase constraints apply because the Forward Partial Combination rule cannot penetrate noun-phrases, so that there is no way that a 'bridge' (Shir, 1977) can be built between the moved item and the extraction site. The absence of negative rules may also have implications for language learning; but this issue is too controversial to pursue here (cf. Levelt, 1979).

The device of partial combination plays a somewhat similar role in the present theory to the metarule which introduces 'derived PS rules' in the Base-generative theory of Gazdar (1981a), in that it is the means by which certain entities such as S/X, a sentence with a 'gap', 'hole', 'trace' or 'empty node', are given grammatical status. There is some similarity between our freedom to exclude from English grammar a rule of 'Backward Partial Combination', (parallel to (17) but with the order of the 'consuming' and 'consumed' functions reversed), and Gazdar's (1981a) freedom to exclude certain derived rules under the 'Generalised Left Branch Condition', although the proposals differ in their further implications. A further similarity is that the restrictions which were imposed upon Forward Partial combination in order to capture the Noun-phrase Constraint are identical to the parallel restrictions on
induction of derived rules in the Gazdar framework. However, (Gazdar, 1981a; Fodor, 1981) the rule is not used to introduce new rules into the base grammar defining constituents with holes or gaps. The categorial grammar already does that. All that partial combination does is to compose ‘hole categories’ (that is, functions), and assemble the corresponding translation as described in the Appendix. This is more than a notational variant. Because the categorial grammar allows more than one gap in a category (for example, the verb put has the category VP/PP/NP) the existence in English of constructions that include two gaps, like (2a), This wood is too rough for these nails to be easy for us to hammer into, and of some of the double extractions that Maling and Zaenen (1981) and Engdahl (1980) have noted in Scandinavian languages, are unproblematic. The context free nature of the categorial component will also ensure that such gaps must nest, rather than cross. Moreover, the apparatus of variable binding that goes with the Lambda Calculus will ensure the correct association of functions and arguments, without recourse to any further ‘Storage Mechanisms’ or devices like Distinguished Variables.

6.3. Psychological Properties
Two features of the model mean that it can be translated easily into a psychologically believable process. First, because there is a one-to-one correspondence between rules of grammar and rules of semantics, it does not require that the processor build autonomous syntactic representations. however temporary, for subsequent semantic interpretation. According to this model, syntax is something that a speaker or hearer does in getting from strings to meanings, or vice versa, not something that is built. Of course, many other grammars offer the possibility of non-autonomy in processing. However, because it includes the Partial Combination mechanism as a rule of grammar, the present one maps directly onto a processor which can build semantic interpretations on an almost word-by-word basis, even for incomplete fragments of the right-branched structures which abound in English. The psychological advantages are considerable. Apart from the fact that it allows incomplete sentences to receive an interpretation, it seems likely that it is some such mechanism that underlies semantic and contextual facilitation effects upon ambiguity resolution (Tyler and Marslen-Wilson, 1977) and upon word recognition (Marslen Wilson and Tyler, 1980). For example, a listener possessing a semantic interpretation of ‘I’m going to drop . . .’ or ‘He might have . . .’ will be able to apply the functions in question to putative subsequent words and/or constituents, and thereby more rapidly identify them.

We would suggest, finally, that this sort of mechanism should be given more weight in discussions of ambiguity resolution. Recent accounts have placed greater emphasis upon structural mechanisms (Frazier and Fodor, 1978; Frazier, 1980), lookahead (Marcus, 1977) and rule-ordering (Wanner, 1980) as ways of resolving ambiguities, and upon ‘garden path’ sentences as evidence. However, as Crain (1980) has pointed out, all garden path effects seem to be eliminable by a suitable choice of semantic or pragmatic context. Since all accounts admit the need for some kind of semantic interaction, it seems likely that systems like the present one which allow the immediate assembly of semantic interpretations will play an important part in psychological theories of ambiguity resolution. (Indeed, it remains unclear how much more is needed than such a system.) If so, the present account will in some sense have come full circle from our initial exclusion of local ambiguity questions. However, such considerations will remain logically secondary to the present attempt to explain the apparent vagaries of a significant fragment of English grammar.

Appendix
1. The grammar

1.1 A Fragment of Lexicon

<table>
<thead>
<tr>
<th>the:</th>
<th>NP/N</th>
<th>ing:</th>
<th>Cing/VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>that:</td>
<td>Cthat/S[+ subject]</td>
<td>Tense:</td>
<td>S/VP/NP</td>
</tr>
<tr>
<td>eat:</td>
<td>VP/NP</td>
<td>put:</td>
<td>VP/PP/NP</td>
</tr>
<tr>
<td>be:</td>
<td>VP/Cing</td>
<td>think:</td>
<td>VP/Cthat</td>
</tr>
<tr>
<td>on:</td>
<td>PP/NP</td>
<td>flog:</td>
<td>N</td>
</tr>
</tbody>
</table>

1.2. The Combination Rules

In what follows, X, Y, Z etc. stand for atomic category symbols such as NP, VP and so on, and $S, S$ etc. stand for members of the set containing the empty string and all strings composed of alternating slashes and atomic category symbols, and beginning with a slash and ending with a category, such as /NP, /PP/NP, /NP/NP/NP and so on. (Thus XS, X$S$, Y$S$/Z etc. stand for either atomic categories or functions like S/NP, VP/PP/NP and so on. Similarly X$/Y$, Y$/Z$ etc. are functions over the categories Y, Z etc.
The combination rules are written:

Forward Combination:
\[ X$ / Y \ Y \Rightarrow X$ \]
Backward Combination:
\[ Y \ S$/Y \Rightarrow S \]
Forward Partial Combination:
\[ X$/Y \ Y$/Z \Rightarrow X$/S$/Z \]
Affix Cancellation:
\[ Y$ \cdot X$/Y$ \Rightarrow X$/S$ \]

2. THE PROCESSOR

When provided with a single Push-Down store, or Stack, and the following trivial control mechanism, the above rules specify a Non-deterministic Push-Down Automaton:

Until the string is empty and no rule matches the topmost items on the stack, either: (a) apply a rule to the topmost items and replace them with the result, or: (b) put the next word of the string on top of the stack.

3. THE SEMANTICS OF THE COMBINATION RULES

Items of category \( X$/Y \) are functions in every sense of the word, including the sense used in Montague semantics and other systems related to the Lambda Calculus. The semantics of Forward and Backward Combination is simply the application of a function to its argument.

The semantics of Forward Partial Combination is marginally less straightforward. Intuitively, though, it is fairly obvious that if we have a definition of a function that we might call I-WILL-SNP (the function over VP interpretations which is produced for the partial sentence I will by Backward combination), and if we also have a definition of the function GIVE_{VP\rightarrow\text{SNP}} (a function from NP intentions onto a function from NP intentions onto VP interpretations), then we know everything that we need to know in order to define a new function I-WILL-GIVE_{SNP\rightarrow\text{VP}}. The new function is simply the composition of the old ones, which can be written

\[
\text{LAMBDA } x[\text{LAMBDA } y[\text{I-WILL-SNP (GIVE}_{\text{SNP}\rightarrow\text{VP}} (x) (y))]]
\]

- where \( x \) and \( y \) are variables which the category of the function identifies as being of the type of NP intentions.

Such a function is a complete semantic interpretation of the partial sentence I will give: it can combine directly with an argument without further modification, and in particular without any structural insertion of further material. Moreover, all such functions can be produced automatically by a single rule which maps pairs of component functions onto their composition.

Acknowledgements

We would like to thank William Marslen-Wilson, Stephen Isard, Janet Fodor, and Phil Johnson-Laird for continual moral and conceptual support. They, and many of our colleagues in Sussex, Warwick, Edinburgh, Amherst, Austin, and Nijmegen gave us valuable comments on two earlier versions which were circulated under slightly different titles (Ades and Steedman, 1979; Steedman and Ades, 1981).

The alphabetically first author was supported by a Fellowship from the Royal Society European Science Exchange Programme, held at the Max Planck-Gesellschaft Projektruppe fur Psycholinguistik, Nijmegen, Netherlands. The research was also supported by a grant from the Social Science Research Council and a Visiting Fellowship from the Sloan Foundation, held at the University of Texas, to the second author. Certain portions of the paper were presented to the autumn meeting of the Linguistic Association of Great Britain, Sussex, September 1979, and the AISB conference on Artificial Intelligence, Amsterdam, March 1980 (Steedman and Ades, 1980).

Reprint requests should be addressed to Steedman, at the Department of Psychology, University of Warwick, Coventry CV4 7AL, England.

Max Planck Institute for Psycholinguistics, Nijmegen
University of Warwick, Coventry

Notes

1. The tendency to demand nested dependencies is overwhelming. There may be very limited occasions on which apparently crossing dependencies are allowed (see accounts of Dutch infinitival complements by Huybregts, 1976 and of Norwegian relative clauses by Engdahl, 1980). However, these languages still predominantly nest dependencies. There are no languages which only allow crossing dependencies, and there are not even any which predominantly cross, allowing just a little nesting.

2. It is often argued that the rules are restricted in this way because Context Free languages can provably all be parsed using resources of time and space proportional to the cube of
the length of sentence or less. But some strictly Context Sensitive languages can be processed with similar economy, so their non appearance remains unexplained.

3 Structure preserving transformations have been studied extensively within the Base Generation framework. For instance Peters (personal communication), Gazdar (1981a, 1981b), Brame (1978) and Bresnan (1978), have all proposed various means of capturing the active-passive relation without transformations. We shall assume that such a approach is feasible and compatible with the model developed in Section 3.

4 Following Emonds (1976) we isolate and exclude from further consideration those Root Transformations that induce comma intonation and/or involve more than one clause. This class includes Left and Right Dislocations, Parenthetical Formation, and Tag Questions.

5 The poetic or stylistic constructions of English also seem to obey the NDC, as for example:

(i) [In Xandu] did Kubla Khan 0

[a stately pleasure dome] decree 0

See Note 1 for some remarks concerning exceptions to the NDC.

6 There seems to be widespread belief that Woods (1973) has proved the use of such incremental interpretation and evaluation to guide syntactic processing to be less efficient than the non-interactive alternative. In fact, the experiments that Woods performed only show this to be true of one particular processor, the LUNAR project ATN, which had a particularly cumbersome interface with its semantic module. They have no implications for other processors, still less for the psychological one.

7 The rule schema (17) is stated in its most general form, and almost certainly induces a grammar which is no longer strictly CF (Peters, personal communication) as the categorial grammar taken with the simple combination schemata is, but which is rather some limited generalisation of CF grammar. We are not sure at this point whether we need the fully general form, nor is it clear exactly what generalisation of CF grammar, if any, is involved. As far as the examples used in the present paper go, the more restricted form

(i) X\(YZ \rightarrow X\)Z

would do – which does not induce greater than CF power. The necessity for the slightly more general

(ii) X\(YZ \rightarrow X\)XZ

for English rests in part on the acceptability of sentences like

(iii) Who did you give the book that you bought for Mary?

<table>
<thead>
<tr>
<th>NP</th>
<th>S/VP</th>
<th>VP/NP/NP</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>S/NP/FP</td>
<td>B</td>
<td>S/NP</td>
<td>F</td>
</tr>
<tr>
<td>S</td>
<td>F</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- upon which matter opinions differ. However, work in progress on certain problems in other Germanic languages leads us to believe that they require at least the more general (ii), and we have therefore chosen not to prejudge the issue. It is worth noting in this connection that the rule is very closely related to the process of inducing 'derived PS rules' in the theory of Gazdar (1981a, 1981b). We return to this comparison in the concluding section.

8 Kimball’s motives for advancing such forms of semantic representation were like our own. He wanted to explain how semantic relationships could be elaborated before the end of right-branching structures was reached. This idea, which he attributed to Wise and Shapiro and called the ‘ongoing function hypothesis’, was to play a crucial role in permitting semantic guidance of parsing decisions at local ambiguity points (see Section 6.1). Speaking of the string Tom wanted to ask Susan..., and of the state of the parser after the first two words had been accepted, Kimball wrote (1975, 174): “I conjecture that at this point we have in the semantics a function which we might call (Tom want), that requires an argument.”

9 As for those cases where the subject object or object appears to take on a different functional role, see Note 3.

10 It will be noted that there is a second sequence in which the combination rules can apply to accept this sentence. Instead of partially combining [he must]\(SF\) and [love]\(NF\), we could input another word [her]\(SF\), combine it with the verb by Forward Combination to yield a complete VP, which could then be combined with [he must]\(SF\) by a further application of Forward Combination, thus:

(i) He must love her

\(SF \rightarrow VP \rightarrow VF \rightarrow S\)

The question of which analysis a processor pursues is of course a question quite external to the grammar. However the semantic interpretation of the ‘partial’ sentence [he must love]\(NF\) produced during the alternative analysis when applied to the object [her]\(SF\) results in exactly the same interpretation produced by (i) (see Appendix). There is therefore no ambiguity of sense of the sentence involved. As far as the examples used in the present paper go, the more restricted form

(i) X\(YZ \rightarrow X\)Z

would do – which does not induce greater than CF power. The necessity for the slightly more general

(ii) X\(YZ \rightarrow X\)XZ

for English rests in part on the acceptability of sentences like

(iii) Who did you give the book that you bought for Mary?

<table>
<thead>
<tr>
<th>NP</th>
<th>S/VP</th>
<th>VP/NP/NP</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>S/NP/FP</td>
<td>B</td>
<td>S/NP</td>
<td>F</td>
</tr>
<tr>
<td>S</td>
<td>F</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- upon which matter opinions differ. However, work in progress on certain problems in other Germanic languages leads us to believe that they require at least the more general (ii), and we have therefore chosen not to prejudge the issue. It is worth noting in this connection that the rule is very closely related to the process of inducing ‘derived PS rules’ in the theory of Gazdar (1981a, 1981b). We return to this comparison in the concluding section.

9 Kimball’s motives for advancing such forms of semantic representation were like our own. He wanted to explain how semantic relationships could be elaborated before the end of right-branching structures was reached. This idea, which he attributed to Wise and Shapiro and called the ‘ongoing function hypothesis’, was to play a crucial role in permitting semantic guidance of parsing decisions at local ambiguity points (see Section 6.1). Speaking of the string Tom wanted to ask Susan..., and of the state of the parser after the first two words had been accepted, Kimball wrote (1975, 174): “I conjecture that at this point we have in the semantics a function which we might call (Tom want), that requires an argument.”

10 As for those cases where the subject object or object appears to take on a different functional role, see Note 3.

11 It will be noted that there is a second sequence in which the combination rules can apply to accept this sentence. Instead of partially combining [he must]\(SF\) and [love]\(NF\), we could input another word [her]\(SF\), combine it with the verb by Forward Combination to yield a complete VP, which could then be combined with [he must]\(SF\) by a further application of Forward Combination, thus:

(i) He must love her

\(SF \rightarrow VP \rightarrow VF \rightarrow S\)

The question of which analysis a processor pursues is of course a question quite external to the grammar. However the semantic interpretation of the ‘partial’ sentence [he must love]\(NF\) produced during the alternative analysis when applied to the object [her]\(SF\) results in exactly the same interpretation produced by (i) (see Appendix). There is therefore no ambiguity of sense of this sentence.

12 We owe this suggestion to Emonds (personal communication). It is possible that so is in fact a Pro-VP in the sentence So must I, in which case (26f) is grammatical as predicted by the model, provided that the VP takes this form.

13 Again, we ignore the question of parenthetical constructions.

14 It is possible to accommodate a Do-support analysis within the present framework. For example, we could assume, along with transformational treatments, that main verbs are distinguished in the lexicon by some feature, and that tensed main verbs are decomposable into two elements, a tensed modal-like element and a main verb stem. The sentences (29) would then be ruled ungrammatical for exactly the same reason as:

(i) *Must love he her

- that is, the Tense bearing item is not adjacent to the subject. One way of incorporating this analysis would be to assume that when a tensed main verb VPS/S/VP/NP is encountered, two items VP/VP/S/VP/NP (a modal plus Tense) and VPS (the main verb stem) are delivered, and are processed just as if they had occurred in the sentence.

However, we regard this solution to the overacceptances (29) as no more than makeshift, and a mere demonstration that a technical solution is possible within the constraints of the present model. The above solution would also necessitate a different analysis of Directional Adverb Preposing. Another solution is implicit in the feature system proposed in Section 4.1.

15 The model will also accept the following, which seem unacceptable to us:

(i) *Been eating beans he has

(ii) *Have stood on the corner he will

16 Why such strings are grammatical when the object NP is ‘heavy’ we do not know.
The range of data covered depends upon the feature composition of the different parts of speech. For example, on the reasonable assumption that adjectives are [...] + N [...] (Chomsky, 1970; Jackendoff, 1977), the Forward Partial Combination rule as stated rules out:

(i) *Angry we made them very

However, since so much doubt surrounds the distribution of features, and even the identity of the features, we regard the present indexing of the Forward Partial Combination rule as little more than illustrative.

This does not commit us to predicting preposition stranding in all languages which extract from within VP. Within other languages, the Partial Combination rule(s) may be indexed by a feature that distinguishes VP and PP, such as +V.

For example, the Generalised Left Branch Constraint has the effect of prohibiting Subject extraction, thus forcing the introduction of a number of metarules which are not required under the present proposal.

In fact, the Forward partial combination rule in the general form in which it is stated in (17) will permit a few cross dependencies under certain very restricted circumstances which do not arise in English. However, the basic point still stands: the Nested Dependency Constraint stems from the CF grammar of which this theory is some limited generalisation.

Such devices (cf. Cooper, 1978; Gazdar, 1981a) are usually considered to be in the domain of semantics rather than syntax. However, to the extent that they augment the corresponding automaton with extra storage devices, they can equally well be regarded as potentially powerful extensions to the syntax itself.

REFERENCES


ON THE PRAGMATICS OF MOOD* 

The basic assumption of the formal semantic program is that to give the meaning of a sentence is to specify its truth conditions. It has long been recognized that a major obstacle to the application of this program to natural languages is the fact that many of the meaningful sentences of these languages are not in the declarative mood and (at least some of them) lack truth conditions. Moreover, the manner in which speakers interpret such sentences does not appear to be significantly different from that in which they interpret declarative sentences. Various attempts have been made to accommodate non-declarative moods within a Tarskian framework, but none of them, I think, is fully convincing. In this paper, I will consider several of these proposals and indicate their shortcomings. I will then suggest an alternative approach for dealing with sentence mood, which avoids these difficulties.

First, however, I would like to recommend two conditions of adequacy which any satisfactory theory of mood ought to satisfy.¹

(i) A theory of mood must give a unified and systematic account of the way in which the mood of a sentence interacts with the truth conditional component of its meaning in order to yield a complete interpretation of the sentence.

(ii) The theory must exhibit what is particular in the way in which each mood determines the interpretation of the sentences in which it is present.

(i) requires that an adequate theory of mood provide a principled analysis of mood in general in relation to the semantic representations assigned to the constituents of a sentence. Such an analysis will entail a strategy for dealing with each particular mood. A theory which does not fulfill (i) will treat individual moods in an ad hoc manner, and fail to capture the general character of the influence of mood on the interpretation of sentences. (ii) simply specifies that the theory must also take account of the factors which distinguish different moods from each other.

The proposals for analysing mood which have been put forward to date by linguists and philosophers tend to exemplify two basic ap-