Phrase Structure Rules

The view of trees that we concluded the last reading with—that tree representations can be decomposed into sets of ‘minimal’ trees, each of which consists solely of a single mother category and one or more daughter categories—has, as I, suggested, quite far-reaching effects. Think of it this way: there are an unbounded number of word strings that satisfy the requirements for sentencehood in any human language. And for each of this infinite set of sentences, there is a tree. But if each such tree can be broken down into a finite number of minimal trees, and if there are only a finite number of types of such trees, then we have a way to use finite means to achieve infinite ends. A finite number of types of trees means that we only need a finite number of statements to say what those types are; once we have them, we can combine them at will, without any restrictions on the size of the result. And since that set of statements will (if we’re right about them) give us all the trees we need to identify what is in the set of sentences of English (or, with a different set of appropriate rules, French, or Mandarin, or . . . ), we will have achieved probably the key part of the essential goal that syntax sets for itself: specifying exactly what it is that makes a string of words a sentence. The idea is this: if for any string of words, we can show, by the appropriate empirical tests, that it can be represented as a tree with certain branchings (justified by our probes for structure, such as displacement and replacement), and if this specific tree can be broken down into a set of minimal trees \( \{T_1, \ldots, T_n\} \), and if every minimal tree \( T_j \) in this set, there is a rule which \( T_j \) satisfies, then and only then do we have a structure which is legal for an English sentence.

There’s only one more thing we have to do, in that case, to ensure that the resulting string of words is a sentence—and we’ll get to that very soon.

VPs

Let’s begin with verb phrases. To see what rules we need, we need to see what minimum trees we need, i.e., what kind of trees of the form

\[
\text{VP} \quad \text{\ldots} \quad V \quad \text{\ldots} \quad \text{\ldots}
\]

I’m taking nothing for granted in this schematic graphic for VP, except for the assumption that since the head of a VP must be some kind of
verb-type object, there must be a V daughter of the VP somewhere in and among the other daughters. Even this is an oversimplification, as it happens (could a VP have a head daughter which was itself a VP?), but it’ll do for the moment. In order to flesh out this skeletal tree structure any further, we need to look at what kinds of VPs are out there. The following is a very substantial fraction of the kinds of VPs that you’ll find if you look (note that I’m treating infinitival phrases such as to drive her mother to the station as VPs, for reasons that will become clear a bit later on this quarter).

(1)  

| a. Robin \[\text{hesitated}\] |  
| b. Robin \[\text{ate a sandwich}\] |  
| c. Robin \[\text{gave the twins a message}\] |  
| d. Robin \[\text{gave a message to the twins}\] |  
| e. Robin \[\text{spoke to me of the future}\] |  
| f. Robin \[\text{seems unhappy with her results}\] |  
| g. Robin \[\text{knows that the twins are up to no good}\] |  
| h. Robin \[\text{told the committee that the twins were up to no good}\] |  
| i. Robin \[\text{bet ten dollars with the twins that the game would go into overtime}\] |  
| j. Robin \[\text{bet the twins ten dollars that the game would go into overtime}\] |  
| k. Robin \[\text{persuaded the twins to surrender quietly}\] |  
| l. Robin \[\text{seems to some of us unhappy with his results}\] |  
| m. Robin \[\text{to me unhappy with her results}\] |  
| n. Robin \[\text{offered ten dollars to the twins to drive her mother to the station}\] |  
| o. Robin \[\text{offered the twins ten dollars to drive her mother to the station}\] |  
| p. Robin \[\text{negotiated with the twins to drive her mother to the station}\]. |  

There may be some other possibilities, but this list is definitely enough to be getting on with. The items on the list are the raw data; but how should we translate them into rules?

The first thing to recognize is that every VP on the list is directly convertible into a minimal tree. Thus, instead of the VP in (1)g, we could write

```
VP
V  NP  S
```
One approach to the solution would be to simply write a separate rule for every possible form of VP displayed in the above list. We would then have a list that matched (1) in one-to-one fashion, along the lines of (2):

(2)  
   a. VP → V  
   b. VP → V NP  
   c. VP → V NP NP  
   ;  

But to proceed in that fashion would miss some critical generalizations—truths which the nature of a list such as (2) fails to express. In a sense, a list such as (2) wouldn’t be *false*—it would contradict no aspect of the observable world—but it would wrong; that is, it would be making claims that the world is a more complex and idiosyncratic place than it really is. Take, for example, the position of NPs in relation to other kinds of phrasal constituents: do NPs ever occur to the right of any other kind of category? Run your eye over the list in (1)—and what you’ll see is that any NP daughter in a VP always precedes every other kind of phrase. An NP may occur to the left of another NP, but that’s all. Similarly, compare the position of the V head in these examples with that of all other elements. Do we ever encounter an instance of a V anywhere but at the left edge of the string of daughters? Again, the answer is no. But these constituent order facts are nowhere expressed in a rule system which simply mimics the examples in (1).

Let’s reason things out a bit more tightly. A more insightful rule system would start from the premise that the invariably left-most position of the verb, and the leftmost position following the verb of any NP(s) that happened to be present, should be built into the rule system; in other words, that to the extent that the rule system encodes our knowledge of what can be a possible VP, it incorporates these two insights, and therefore must express the latter by virtue of the very form of the rule. But we can go further. We have V NP NP, V NP PP and V PP PP as possibilities, but we do not have V VP S, V S VP, V VP VP or V S S. So these restrictions must be accounted for as well. The trick seems to be that the rule system must have, built into it, the possibility that there can be a series of NPs or PPs, as well as NP + PP, but that there can be at most one instance of S mentioned in the rule, one instance of VP, and the choice among the two must be mutually exclusive. Setting things up that way will rule out VP VP, S S, S VP and VP S following the V, because we only get to choose one instance of S per rule and one instance of VP, and if we pick one, we can’t pick the other. If this seems a bit abstract and difficult to encode, think of it as something which naturally falls out from a pull-down menu in a software application, where you can pick a certain option only from one menu—it’s just not available on any of the other pull-down menus—and
no matter which one you pick, you can’t pick any of the others on that same menu. In syntax, the notation for this sort of thing is

\[
\begin{cases}
X \\
Y \\
Z
\end{cases}
\]

indicating that you have a choice of X or Y or Z, and if you don’t have another such trio of bracketed possibilities somewhere to the right or left of the first one, then that’s the only time you get to pick one of these three.

Putting these desiderata together, we wind up with a single rule that looks very much like (3):

\[
(3) \quad VP \rightarrow V(NP) \left( \left\{ \begin{array}{c} NP \\
PP (PP) \end{array} \right\} \left( \left\{ \begin{array}{c} AP \\
S \\
VP \end{array} \right\} \right) \right)
\]

Here’s how to read this rule:

We’re told that a VP can take the form of a lexical (i.e., word-level) head V, with an optional NP following. Whether or not we take that option, there are two possible continuations. We can choose the options corresponding to

\[
\left( \left\{ \begin{array}{c} NP \\
PP (PP) \end{array} \right\} \right) \left( \left\{ \begin{array}{c} AP \\
S \\
VP \end{array} \right\} \right)
\]

where we can choose an NP or a PP, or a sequence PP PP, or none of these possibilities, followed by either an S or an infinitive VP, or neither. The choice NP NP S corresponds to the material following the verb in *bet Robin ten dollars that the game would go into overtime*; the sequence NP PP PP is needed for examples such as *bet ten dollars with Robin on the game*, and so on. Or we can forget about these possibilities and pick the options corresponding to ((PP) AP), which include *seems rather unhappy*, *seems to me rather unhappy*, and *consider Robin very unpleasant*. The possibilities associated with AP are somewhat different from those involving NPs and PPs, as already mentioned, and the rule in (3) incorporates these differences.

(3) accounts for all of the cases displayed in (1), and several others as well, and does so in a way which ensures that the V head occurs first within any VP with a lexical head, that no non-NP constituent ever precedes an NP, that no sequences of S and/or VP ever occur in any order, and
that an AP cannot be followed by either S or VP within a VP with a V head—among a number of other correct observations. It replaces the list of rules abbreviated in (2) with a single (though admittedly rather bulky) rule which expresses—simultaneously, so to speak—all these generalizations. If we add to this VP rule a few others accounting for NPs and the other major phrasal categories, we will have a compact set of principles that determine the form of all the subtypes of minimal trees that English allows, and therefore all of the members of the infinite set of English sentences.

NPs

NPs offer less variety than VPs, but there are still a reasonable number of possibilities that we need to account for:

(4) a. [np [s books]]
   b. [np [det the] [s books]]
   c. [np [s books] [pp of poetry]]
   d. [np [det the] [s books] [pp of poetry]]
   e. [np [slender] [s volumes] [pp of poetry]]
   f. [np [det those] [slender] [s volumes] [pp of poetry]]
   g. [np [det your] [s belief] [vp that Robin is guilty]]
   h. [np [det your] [groundless] [s belief] [vp that Robin is guilty]]
   i. [np [det your] [unrealistic] [s hope] [vp to compete in the Olympics]]
   j. [np [det several] [s signals] [vp to leave the building immediately]]
   k. [np [det my] [s disagreements] [pp with Robin] [pp about politics]]

These too can be collapsed into a single rule, which we state as (5):

(5) NP → (Det) (A) N (PP) (PP)

In formulating this rule, bear in mind that when two forms of the NP differ only in terms of an optional element, any rule you write for the situation in which the optional element is missing can handle the case in which it’s present if you write that element into the rule using option braces. So, for example, if you have two possibilities of the form

\[ X \rightarrow A \ C \ D \]
\[ X \rightarrow A \ B \ C \ D \]

then you can collapse the two together as

\[ X \rightarrow A \ (B) \ C \ D \]

In the same way, two rules
X → A B Y
X → A B Z
can be reduced to a single rule
X → A B \{ Y \}
If you encounter a situation where no two parts of the rule can be collapsed, as with
X → A B C
X → D E F
you can always write the rule
X → \{ A B C \}
\{ D E F \}
But in such cases, we shouldn’t kid ourselves into thinking that we’ve done any more than state the content of two separate rules by means of a single arrow. Such a rule is irredeemably disjunctive: it’s sort of like saying that in 99% of American families who own a pet, that pet is a ‘blickit’, where a blickit can take the form of a dog in one family and a cat in another. What has been gained here over the more straightforward statement that in 99% of American families who own a pet, that pet is either a dog or a cat?
Note further however that if the situation corresponds more to a pair of rules of the form
X → Y Z A B C
X → Y Z D E F
then the situation is significantly different. It is presumably no coincidence that Y and Z are present in both rules. The fact that the rules then diverge is less important than the fact that we have both Y and Z in that order as expressions of both rules, suggesting that they correspond to a single rule with partially branching possibilities.
But having come this far, we need to ask if there might not be more to the structure of NPs than we’re seeing. Compare the sentences in (6):

(6) a. Robin and Terry were arguing with each other in a restaurant.
   b. Terry and Robin were arguing with each other in a restaurant.

Constructions in which and, or and a few other words which correspond to English ‘logical’ vocabulary are said to exhibit the coordination phenomenon. The question of what is actually being coordinated isn’t necessarily always obvious, however. One excellent test is to look at the string
containing the coordination and then switch the positions of the strings you think are being coordinated, as in (6). For example, it’s conceivable that the coordinated strings in (6) are *Robin on the one hand and Terry were; but if that were so, we’d expect *Terry were and Robin arguing with each other in a restaurant to be good—which it definitely isn’t. We can therefore be confident that in (7)a, it’s not The King of France that’s being coordinated with Queen of Spain:

(7) a. The King of France and Queen of Spain were arguing with each other in a restaurant.
   b. *Queen of France and the King of Spain were arguing with each other in a restaurant.
   c. The Queen of Spain and King of France were arguing with each other in a restaurant.

What does work is switching the positions of King of France on the one hand and Queen of Spain on the other. Since, in general, strings that can be coordinated correspond to constituents, we have some plausible evidence that Queen of Spain and King of France are constituents.

But the evidence goes from fairly good to excellent when we consider the replacement pattern in (8)

(8) a. I like **this** King of England better than ****that** one. (*one = ‘King of England’)
   b. Some students of chemistry are smarter than other ones. (*ones = ‘students of chemistry’)
   c. Leslie found a new book of poetry after Terry bought a used one. (*one = ‘book of poetry’)

One here seems to have no meaning of its own; but it does pick up a meaning from the syntactic context appears in. It also has a plural form:

(9) Robin likes these books of poetry better than those other ones.

On this basis, it seems fair to conclude One(s) is a proform for a subportion of an NP, which in turn means that that subportion is itself a constituent. But how are we to label it in phrase structure trees such as (10)?

(10) NP
    Det ?
    this King of England

NP
    Det ?
    some students of chemistry

The standard notation for the category of this type of constituent is $\bar{N}$. $\bar{N}$ doesn’t include the determiner of the NP—*but it does include everything
else. \( \tilde{N} \) is thus a phrase, and headed by a noun—but it’s not an NP, because we reserve that symbol for the maximal-sized phrase headed by a nominal, something which in principle CANNOT appear with a determiner. Adding a determiner to \( \tilde{N} \) in effect ‘closes off’ the phrase, ensuring that there is no way to expand it ‘upwards’ to form a bigger unit. We thus have to think of the NP in terms of the following kind of structural skeleton:

\[
\begin{array}{c}
\text{NP} \\
\text{Det} \\
\ldots \\
\tilde{N} \\
\ldots \\
\end{array}
\]

But even this may not be enough. Consider the following example(s):

(12) I bought this new book of poetry, and robin bought that (used) one.

When \textit{used} does not appear, \textit{that} \textit{one} clearly refers to ‘that new book of poetry’. In other words, \textit{one} replaces \textit{new book of poetry}, so that the adjective is inside the NP. But when \textit{used} appears, \textit{one} is replacing only \textit{book of poetry}. In other words, the adjective is OUTSIDE the \( N \). How can the adjective both be outside and inside the \( \tilde{N} \)?

But this is something we’ve seen before—within VPs, where there was an adjunct (e.g., an adverb) sister to a smaller VP. If we compare the tree for the VP in \textit{Robin quickly ate a sandwich} with the \( \tilde{N} \) \textit{used book of poetry}, we see a remarkable parallelism:

\[
\begin{array}{c}
\text{VP} \\
\text{Adv} \\
\text{quickly} \\
\end{array}
\begin{array}{c}
\text{VP} \\
\text{ate a sandwich} \\
\end{array}
\begin{array}{c}
\text{\( \tilde{N} \)} \\
\text{A} \\
\text{used} \\
\text{book of poetry} \\
\end{array}
\]

The evidence base for the analysis in both cases is proform replacement—so for the \( \tilde{N} \), we have, e.g.,

(14) \[
\begin{array}{c}
[\text{NP Det new book of poetry}] \\
\text{one (= \( \tilde{N} \))} \\
\end{array}
\]

(15) \[
\begin{array}{c}
[\text{NP Det [\text{new book of poetry}]]} \\
\end{array}
\]
Given this analysis, we can now make a parallel distinction between nominal 
adjuncts—elements appearing under NP which are not sisters to the actual 
noun heading the phrase—and complements, which are sisters to the head. 
The contrast emerges in data such as (16):

(16) a. I like the student of Scottish origin better than the one from 
    Ostrogothia. (one = ‘student’)
   b. I like THIS student of Scottish origin better than THAT one. (one 
      = ‘student of Scottish origin’)
   c. I like THIS student of chemistry better than THAT one. (one = 
      ‘student of chemistry’)
   d. *I like the student of chemistry better than the one of physics.

One replacement has no trouble with student in student of Scottish origin/student 
from Ostrogothia, but fails in the case of student of physics. This is no differ- 
ent from the contrast between Robin hasn’t on oil rigs/done so for a 
long time on the one hand and Robin was waiting/*doing so for Leslie (as 
in *I thought Robin was waiting for Terry, but it turned out she was doing 
so for Leslie. We therefore apparently have the contrast

(17) \[
\begin{array}{cc}
\text{NP} & \text{NP} \\
\text{Det} & \text{Det} \\
\text{the} & \text{this} \\
\text{N} & \text{N} \\
\text{of Scottish origin} & \text{of chemistry} \\
\text{student} & \\
\end{array}
\]

In the case of student of chemistry, an \( \text{\tilde{N}} \) branches directly to an N and a 
PP, and one cannot replace student, because the latter isn’t an instance of 
an \( \text{\tilde{N}} \). But in the case of student of Scottish origin, where replacement is 
possible, the correct analysis must be that of Scottish origin is a sister of 
an \( \text{\tilde{N}} \)—that it’s in an ADJUNCT position. Hence student is a manifestation 
of \( \text{\tilde{N}} \) in this case and one replacement is possible.

These considerations lead us to revise our rule for NP along the following 
lines:

(18) \[
\begin{align*}
\text{NP} & \rightarrow (\text{Det}) \text{\( \tilde{N} \)} \\
\text{N} & \rightarrow \text{\( \tilde{N} \)} \text{PP} \\
\text{N} & \rightarrow \text{N} \text{PP}
\end{align*}
\]