The question that syntax is the answer to...

When we leave the level of the word behind, and start considering the combinations of words that we call sentences, there is a question that arises which cannot be solved just by ‘eyeballing’ the data and deciding one way or the other. The question is, do sentences have a determinable structure, and if so, how do we determine it? When we first confront sentences, they appear to be nothing but shorter or longer chains of words. But is that all they are? How can we tell?

The key to discovering the structure of anything is to subject it to tests that reveal its internal ‘joint-points’, the places where its parts combine. Structure, in general, is just that: the way in which the parts of a thing are put together. In the case of sentences, one way we can get at their internal structure is to test various subportions of them—substrings of words—to see which ones hang together, so to speak, and which ones don’t. When we do that, what we see is that the number of substrings of a word which do behave as a unit is a lot smaller than the number which don’t (and the discrepancy increases very quickly as the number of words in the sentence increases).

The rearrangement test

Let’s start by identifying what subgroupings of words in a sentence can be rearranged. We can start with the fact that certain groups of words following what are commonly called auxiliaries can be rearranged so that they appear at the very front of sentences, rather than over on the right (thinking in terms of how they’re written) as they usually do. The auxiliaries are a small group of words in English which includes be, have, may and several others; I’ll have more to say about them later. But for the moment, consider the following pairs of sentences (where the ‘gap’ marker indicates the place where the material appearing in the b. sentences on the far left side of the sentence is normally found, corresponding to the a. sentences):

1. a. If Terry said she has eaten an apple, then she has eaten an apple.
   b. If Terry said she has eaten an apple, then eaten an apple she has __.

2. a. Terry said she may go to the movies, and indeed she may go to the movies.
   b. Terry said she may go to the movies, and indeed go to the movies she may __.

3. a. If Terry said that Robin has been giving Leslie a hard time, then Robin definitely has been giving Leslie a hard time.
   b. If Terry said that Robin has been giving Leslie a hard time, then giving Leslie a hard time Robin definitely has been __.

Compare these cases with those in (3), where various other substrings of the sentence in (3)a are rearranged—with awful results—in (3)b–d.
(4)  a. Terry says she is putting the book on the table, and putting the book on the table she is _

b. *Terry said she is putting the book on the table, and putting she is _the book on the table.

c. *Terry said she is putting the book on the table, and putting the book she is _on the table.

d. *Terry said she is putting the book on the table, and putting the book on she is _the table.

Typically, in a situation where you have a grouping of objects that seems to be independently manipulable, it’s a sign that that grouping actually forms some kind of whole, a subunit of a larger object. I’m going to assume that that’s the case here. The simple assumption is that if I can do this kind of fronting with a certain string of words, that substring is such a subunit of the sentence. If so, then all we have to do to get the difference between (1)–(3) on the one hand, and the bad cases in (4) on the other, is assume that whatever general pattern relates the normally-ordered data to the fronted examples only applies to such subunits. Easy, eh?

Well, maybe not so easy. In (5), we have some examples that seem to challenge the supposedly simple approach I suggested:

(5)  a. Terry said she would leave, and leave she did _

b. Terry said she would leave us, and leave us she did _

c. *Terry said she would leave us, and leave she did _us.

d. Terry said she would leave us the keys, and leave us the keys she did.

e. *Terry said she would leave us, and leave us she did _the keys.

You should have a least a general sense of why these examples mean trouble for my proposal. In (5)a, we have evidence that leave is a subunit of the right sort; after all, it does front in that example. But in (5)c, it’s clear that leave can’t just front on its own. And similarly with (5)b and (5)e. But there’s a wrinkle that we can’t ignore here: in the case where leave can be fronted, it’s not leaving anything behind, so to speak. Where it can’t be fronted, there’s an us that remains behind. Similarly, when leave us is fronted, there’s nothing else that follows the auxiliary. When it can’t front, as in (5)e, there seems to be something else associated left over. This turns out to be very important in understanding what’s going on. But for the moment, we need to adjust the overly simple formulation I gave earlier. The hypothesis should be modified to read, if a substring in a given sentence can be fronted, then that substring is a subunit of that particular sentence, but not necessarily of another sentence which contains the same substring.

The replacement test

But there’s a more serious problem with the proposal I made earlier about subunits—namely that, at the moment, all the notion ‘subunit’ amounts is another way of saying ‘can be rearranged’! In order to avoid circularity—it’s shown to be a subunit by the fact that you can rearrange it, and the reason you’re able to rearrange it is that it’s a subunit—we need something else, some independent reason to believe that what we’ve isolated actually corresponds to a real entity, a genuine part of the structure of the sentence which is giving us signs of its existence.

Fortunately, there is a second, confirming test of the existence of these subunits. Consider the data in (6):
(6)  a. If Terry said she would go to the movies, then she will go to the movies.
    b. If Terry said she would go to the movies, then she will DO so.

(7)  a. If Terry said she was putting the book on the table, then she was putting the book on the table.
    b. If Terry said she was putting the book on the table, then she was DOING so.

You can front the strings of words in boldface, but you can also replace them with some form of do so. What does do so mean? Clearly, it has no meaning of its own; it takes its meaning from the sequence words that it (in effect) replaces, a sequence which occurs earlier in the sentence. Notice also that the material which can’t be fronted—putting, putting the book, putting the book on—cannot be replaced either.

(8)  e. *If Terry said she was putting the book on the table, then she was DOING SO the book on the table.
    f. *If Terry said she was putting the book on the table, then she was DOING SO on the table.

It looks as if those blocks of words definable on the basis of rearrangeability are also referenced when it comes to identifying what is replaceable by an expression which doesn’t have its own meaning. In all cases, these word blocks take the form, summarized in (9)a of a verb (V) followed by a series of other words (W), where V W can be either fronted (as in (9)b) as a rearrangement of the sentence, or replaced (as in (9)c) by a form—commonly referred to as a pro-form—which stands in for it.

(9)  a. N aux V W (W a cover symbol for a series of words following the verb).
    b. V W N aux
    c. ...N aux do so.

We can apply these test freely; we get results such as those in

(10)  a. Robin has eaten breakfast.
    b. ... and eaten breakfast, Robin has
    c. Terry has eaten breakfast, and Robin has \{ eaten breakfast
         done so \} as well.

(11)  a. Robin will talk to Terry.
    b. ... and talk to Terry, Robin will
    c. ... and Robin will do so.

(12)  a. Terry should signal to Leslie to leave immediately.
    b. ... and signal to Leslie to leave immediately, Terry definitely SHOULD
    c. ... and Terry should DO so.

I’m going to put brackets around these word blocks, indicating that they pass the tests for syntactic unithood—or, using a more common phrasing in syntax, for constituency. A constituent is a grouping of words which displaces these unitary properties, which go together to CONSTITUTE a larger (sub)unit of a sentence. Thus we have
a. Terry has [eaten an apple].
b. Terry may [go to the movies].
c. Robin has been [giving Leslie a hard time].
d. Robin has [eaten breakfast].
e. Robin will [talk to Terry].
f. Terry should [signal to Leslie to leave immediately].

The term *phrase* is also applied to such blocks of words which behave as a unit, and I’m going to talk about these blocks as phrases of a certain kind. If you take certain verbs, none of the elements which follow them can be omitted:

c. *Robin put on the table.*
d. *Robin put.*

But in other cases, there are omissible elements:

(15) a. Terry withdrew his name from the candidate pool.
b. Terry withdrew from the candidate pool.
c. Terry withdrew his name.
d. Terry withdrew.

(16) a. Robin ate a steak.
b. Robin ate.

(17) a. Leslie signaled to Robin to leave immediately.
b. Leslie signaled to Robin.
c. Leslie signaled to leave immediately.
d. Leslie signaled.

Of course, the fact that we can say *Robin ate* implies that Robin ate something or other, and the fact that we can say *Leslie signaled* implies that Leslie signaled to someone or other. But the important point, so far as the form of the sentence is concerned, is that we don’t actually have to include a sequence of words identifying what it was that Robin ate or who it was that Leslie signaled to (unlike the case of *put*, where we have to say explicitly what Robin put where). But notice something further: if we leave out the verb here, we have a nonsentence:

(18) a. *Terry his name from the candidate pool.*
b. *Robin a steak.*
c. *Leslie to Robin to leave immediately.*

There is a generalization here: in all such blocks of words—the ones corresponding to V W in (9)—if any element is omissible, it is not the verb. It follows that the verb will in many cases be the sole indispensable part of the sequence—the phrase, as I’m referring to it. In view of this property of such phrases, I will refer to them as *verb phrases*, abbreviated VP, and will use VP as the label for the brackets I put around such word strings to indicate their phrasehood. Thus we have

(19) a. Terry has [VP eaten an apple].
b. Terry may \([_{VP} \text{go to the movies}]\).
c. Robin has been \([_{VP} \text{giving Leslie a hard time}]\).
d. Robin has \([_{VP} \text{eaten breakfast}]\).
e. Robin will \([_{VP} \text{talk to Terry}]\).
f. Terry should \([_{VP} \text{signal to Leslie to leave immediately}]\).

Now what about these auxiliaries—what are they? As I indicated in class, there are four properties which distinguish auxiliaries from all other words in the languages, in particular the so-called NICE properties (where when two curly braces appear in the same sentence, you should match up elements in them in the obvious way):

\[
(20)\]

\begin{enumerate}
\item \textbf{N(egation)}: Terry \[
\begin{align*}
\text{is} & \quad \text{has} \\
\text{can} & \\
\text{not} & \quad \text{speaking} \\
\text{spoken} & \quad \text{speak}
\end{align*}
\]
\item \textbf{I(inversion)}: \[
\begin{align*}
\text{is} & \quad \text{has} \\
\text{can} & \\
\text{speaking} & \quad \text{spoken} \\
\text{spoken} & \quad \text{speak}
\end{align*}
\] Terry \[?
\]
\item \textbf{C(ontraction)}: Terry \[
\begin{align*}
\text{is} & \quad \text{has} \\
\text{can} & \\
\text{not} & \quad \text{speaking} \\
\text{spoken} & \quad \text{speak}
\end{align*}
\]
\item \textbf{E(llipsis)}: Terry \[
\begin{align*}
\text{is} & \quad \text{has} \\
\text{can} & \\
\text{not} & \quad \text{speaking} \\
\text{spoken} & \quad \text{speak}
\end{align*}
\], but Robin \[
\begin{align*}
\text{is} & \quad \text{has} \\
\text{can} & \\
\text{not} & \quad \text{speaking} \\
\text{spoken} & \quad \text{speak}
\end{align*}
\]
\end{enumerate}

But what exactly are auxiliaries? In the case of \textit{have} and \textit{be}, that’s pretty clear: they’re verbs. After all, they display tense:

\[
(21)\]

\begin{enumerate}
\item Robin \[
\begin{align*}
\text{is} & \quad \text{was} \\
\text{can} & \\
\text{speaking} & \quad \text{spoken}
\end{align*}
\]
\item Terry \[
\begin{align*}
\text{has} & \quad \text{had} \\
\text{spoken} & \quad \text{spoken}
\end{align*}
\]
\end{enumerate}

There is a tense distinction in both cases, exactly the same as in \textit{Robin is/was a spy, Terry has/had a major problem}. Only verbs in English display tense morphology. It follows that these forms, at any rate, are verbs. But what about, e.g., \textit{can, should} and so on? We’ll get to those in due course.

But meanwhile, we have some important ideas on the table: sentences have subparts which we can establish by empirical tests (rearrangement and replacement) which converge on the same results—a very important point—and these subparts may have special, ‘privileged’ elements which are, for example, obligatory if any element within them is obligatory. As we’ll soon see, VPs are from a unique case in any of these respect. But for the time being, notice that we can break the seemingly homogeneous string of words \textit{Robin had put this book on that table} is, we now know, to be represented as

\[
(22) \quad [s \text{ Robin had } [_{VP} \text{put this book on that table}]])
\]

where I’m bracketing the whole sentence as \([s , to indicate that it a syntactic unit—the biggest syntactic unit, in fact, the one whose internal structure syntax is concerned with. But we’ve only just gotten started...}