Structure

The concept of structure is fundamental to the study of syntax. But it is a very general concept that can be applied to any complex thing, whether it’s a bicycle, a commercial company, or a carbon molecule. When we say of a thing that it is complex, we mean, not that it is complicated (though of course it may be), but that

(a) it is divisible into parts (called constituents),
(b) there are different kinds of parts (different categories of constituents),
(c) the constituents are arranged in a specifiable way,
(d) that each constituent has a certain specifiable function in the structure of the thing as a whole.

When anything can be analysed in this way, we say that it has structure. In considering structure it is important to note that, more often than not, the constituents of a complex thing are themselves complex. In other words, the parts themselves consist of parts, which may in turn consist of further parts. When this is so we may speak of a hierarchy of parts and of hierarchical structure.

It is obvious, for example, that a complex thing like a bicycle is not just a collection of randomly assembled bits and pieces. Suppose you gathered together all the components of a bicycle: metal tubes, hubs, spokes, chain, cable, and so on. Now try to imagine all the possible objects you could construct by fixing these components together. Some of these objects might be excellent bicycles, while others wouldn’t remotely resemble a bicycle (though they might make interesting sculptures). And, of course, there would be intermediate cases, things which we would probably want to say were bicycles, if only because they resembled bicycles more than anything else.

So, only some of the possible ways of fitting bicycle components together produce a bicycle. A bicycle consists not just of its components but, much more importantly, in the structure that results from fitting them together in a particular way.
When we turn to linguistic expressions, we find a similar state of affairs. Suppose you have a collection of words, say all the words in a dictionary. Can you imagine all the possible word-sequences you could construct by putting these words together? The possibilities are endless. Clearly not all the sequences would be acceptable expressions of English. And again, some would be odder than others. When a sequence of words fails to constitute a good expression in the language, I shall describe it as being ungrammatical (or ill-formed) and mark it with an asterisk (*). Here are some examples:

[1a] *the nevertheless procrastinate in foxtrot
[1b] *disappears none girls of the students
[1c] *Max will bought a frying pans.

More subtle examples of ungrammatical sentences were given in the Introduction.

Ultimately, a full syntactic description of any language consists in explaining why some strings of words of the language are well-formed expressions and why others are not. Just how this ultimate (and very ambitious) goal might be attempted is discussed in Chapter 11. It is enough to say here that it could not be achieved without recognising structure. Just as the concept of structure was required in distinguishing between the bicycles and the would-be bicycles, so the concept of structure is essential in distinguishing between the strings of words that are well-formed expressions and those that are not.

We can use diagrams to show how things are analysed into their constituent parts. For instance, [2] says that a bicycle can be analysed into two wheels, a frame, a chain, handlebars, among other things (the dots mean ‘and other things’):

![Tree Diagram]

Such diagrams are called tree diagrams (though the trees are upside-down).

I've mentioned that the constituents of a complex thing can themselves be complex. An example of this is a bicycle wheel. It is itself a constituent of the bicycle, but in turn consists of hub, spokes, rim, tyre, etc. Although it's true that spokes are constituents of bicycles, it's more important to note that they are constituents of bicycles only because they are constituents of the wheel which, in turn, is a constituent of the bicycle. The relation between spoke and bicycle is indirect, mediated by wheel. We might express this by saying that, although the spoke is a constituent of the bicycle, it is not an immediate constituent of it.

It is important to recognise the indirectness of the relationship between bicycle and spoke because, in giving a description of the structure of bicycles, we need to be able to say that wheels are parts of bicycles. But if we allowed that spokes were immediate constituents of bicycles rather than of wheels, this would
leave wheels out of the picture. It would imply that bicycles could have spokes independently of the fact that they have wheels, and that spokes were not a necessary part of the structure of wheels.

As mentioned, specifying the function of constituents is an important part of structural analysis. Notice that if we were to represent spokes as immediate constituents of bicycles, it would be impossible to specify correctly what the function of the spokes is. The spokes don’t have a function in respect of the bicycle directly, but only in respect of the wheels. In talking of the function of the spokes, then, we’re going to have to mention the wheels anyway.

Which of the following tree diagrams best represents the structural relationship between bicycle and spoke just discussed?

Although each tree diagram is incomplete, the one that more accurately reflects the structural relationship between bicycle and spoke is [3b], since it says that spokes are constituents of wheels, which are, in turn, constituents of bicycle. It correctly describes the relation between bicycle, wheel, and spoke as being a hierarchical relation. [3a], on the other hand, says that spokes are immediate constituents of bicycles, independently of the fact that wheels are constituents of bicycles.

This book is concerned with syntactic structure – that is, with (a) analysing linguistic expressions into their constituents, (b) identifying the categories of those constituents, and (c) determining their functions. But what kind of expressions should we begin with? I’ll take the sentence as the starting point for analysis. I’ll assume (and in fact already have assumed) that you have an intuitive idea of what counts as a sentence of English.

The first question to be asked is, ‘What do sentences consist of?’ The answer might seem blindingly obvious: ‘Sentences consist of words.’ In the rest of this chapter (and, for that matter, the rest of the book), I’ll try to convince you that this apparently natural answer is not the most appropriate one. In fact, the discussion of hierarchical structure and the importance of recognising that sentences have such structure forces us very quickly to abandon the idea that sentences consist, in any simple way, of words.

This can be shown by asking whether the relationship between a sentence and its words is direct or whether it is indirect, mediated by parts of intermediate complexity. This amounts to asking: ‘Are words the immediate constituents of the sentences that contain them?’ It is only if the words
contained in a sentence are its immediate constituents that we can allow that sentences actually consist of words. As an aid to thinking about this question – and to gain practice in getting such diagrams to say what you want them to say – draw a tree diagram, starting with ‘Sentence’ at the top, which says of sentence [4] that its words are its immediate constituents, that it consists directly just of the words it contains. Having done that, ask yourself whether the diagram you have drawn gives an accurate representation of the structure of the sentence as you feel it to be.


The diagram that says of sentence [4] that its words are its immediate constituents looks like this:

```
    Sentence
      /    \
    Old   Sunbathed beside a stream
```

Do you feel that the diagram is wrong and/or unhelpful as a description of sentence [4]? How much does it tell us? Well, it tells us what words appear in the sentence. And in what order they appear. But nothing more. As well as being uninformative, the diagram is actually wrong as a description of the structure of the sentence. In essence, it says of sentence [4] that it has no structure – or no more structure than a sequence of numbers (1–2–3–4–5) or an ordered string of beads. This is surely wrong.

In not allowing that the sentence has constituents that mediate between it and its words, the diagram doesn’t allow that certain of the words seem to belong with others, that the words seem to work in groups. It says that the words have no relationship to each other except the relationship of being in a certain order in the same sentence. And, although the diagram tells us in what order the words occur, in failing to assign any but the simplest possible structure to the sentence, it fails to give any explanation of why they occur in that order to form a sentence, and why the orders in [6] and [7], for example, don’t form sentences of English.

[6] *Stream old Sam sunbathed beside a
[7] *Sunbathed old beside stream a Sam

We need to say that sentence [4] is more highly structured than [5] says it is. As we saw in the discussion of bicycles, the position of a spoke in the structure of a bicycle is determined by its being a constituent of the wheel, which itself has a certain position within the bicycle. If you reposition the spokes from out of their structural position in the wheel, you land up with an unworkable bicycle. A very similar thing has happened in [6] and [7]. The position of words in a sentence is determined by the fact that the words are not immediate constituents of the sentence, but belong with other words to form groups – **phrases** – which
have their own position in the structure of the sentence. It is these phrases (and further phrases made up of these phrases) that function as immediate constituents of the sentence. In short, **while sentences certainly contain words, they don’t consist of words. They consist of phrases.**

In addition, we need to be able to say what kinds (or categories) of words can combine to form structural groups. What’s wrong with [6] and [7] is that words have been displaced from positions in which they are capable of forming phrases with the words next to them to positions where they are not, given the kinds of words they are. But the diagram gives no information of this sort. Such information is needed to account for the ungrammaticality of [6] and [7], but it is also needed if we want to explain why replacing *stream* with *road* yields another good sentence of English:

[8] Old Sam sunbathed beside a road.

but replacing *stream* with *laughing* or *surreptitiously* does not.

[9a] *Old Sam sunbathed beside a laughing.*
[9b] *Old Sam sunbathed beside a surreptitiously.*

*Road* can replace *stream* in [4] because *road* and *stream* belong to the same category: they are both nouns. *Laughing* and *surreptitiously* cannot replace *stream* because they aren’t nouns; they belong to other categories (verb and adverb).

So we need to include information about grammatical categories in our diagrams and this is something we’ll look at in later chapters, especially Chapter 3. Together with information on how the words group into phrases, this will help to explain not only the facts about [6]–[9], but also facts about the functions of words (and phrases) in sentences.

The discussion so far suggests that diagram [5] is actually wrong as a structural description of sentence [4]. As soon as we want to explain even the simplest things about sentences, it’s necessary to go beyond the idea that sentences simply consist of words strung together in a line. We need to acknowledge that sentences have hierarchical structure.

---

**Establishing constituents**

I’ve been complaining in a rather general way about diagram [5]. What’s needed now is a more specific demonstration of just how it is wrong. I won’t give a complete analysis of sentence [4] here, but just a general introduction to the identification of constituents larger than the word.

Here’s one way of clearly establishing that [5] is wrong. If the sentence had the same (lack of) structure as an ordered sequence of numbers, we should be able to lop words off the end of the sentence and still be left with a good sentence
every time we did so. We can lop numbers off the end of a number sequence and still be left with a good (though shorter) number sequence: 1–2–3–4–5, 1–2–3–4, 1–2–3, 1–2, 1. Begin by removing first one word and then another from the end of sentence [4] until you’re left with just one word. Each time, write down the string that remains. In front of every string of words that seems to you not to constitute a complete and grammatical sentence, put an asterisk.

Assuming we all speak the same language, you should have a list of five strings marked in the following way:

[10] *Old Sam sunbathed beside a
[11] *Old Sam sunbathed beside
[12] Old Sam sunbathed
[13] *Old Sam
[14] *Old

Of the strings, only [12] could stand as a complete and well-formed sentence. [13] may not seem as odd as [10], [11], and [14] do, for reasons which will become apparent shortly. It should still be asterisked since it is not a complete sentence. What needs explaining is why string [12] is a good sentence while none of the others are.

In the first place, you should note that not all parts of a sentence are necessary in order for that sentence to be complete and well-formed. Consider [15].


[15] is a good sentence as it stands. But notice that we could add to it. For example, we could add the word *invitingly*, to produce another good sentence [16]:


In [16], then, we can say that *invitingly* is an optional part of the sentence: leaving it out gives us another (though shorter) complete and perfectly grammatical sentence, namely [15]. By contrast, *Martha* and *smiled* are obligatory.

The importance of this here is that I’ve referred to *invitingly* as a part, as a constituent, of sentence [16]: I have said that it is an optional constituent. Of course, it’s obvious that *invitingly* must be a constituent in sentence [16], since it is a word. But, to go back to sentence [4], we saw in [10]–[14] that we could omit the sequence of words *beside* plus *a* plus *stream*, leaving a perfectly good sentence. In other words, that sequence of words is optional. Notice, though, it’s only the sequence as a whole, as a single unit, that’s optional. None of the words in that sequence can be omitted individually – that’s what *[10] and *[11] show. So, just as I needed to refer to the single word *invitingly* and say it was an optional constituent in the structure of sentence [16], so I need to be able to refer to the sequence of words *beside + a + stream* and say of it that – as a
CHAPTER 1 SENTENCE STRUCTURE: CONSTITUENTS

unit – it is optional in the structure of sentence [4]. In doing so, I acknowledge that word-sequence as an identifiable part, as a constituent, of that sentence.

Sequences of words that can function as constituents in the structure of sentences are called phrases. Tree diagrams represent structure by marking which sequences of words in a sentence are its constituent phrases. So syntactic tree diagrams are, more specifically, called phrase markers.

I have shown that the sequence of words beside a stream is a constituent of sentence [4]. So [beside a stream] is a phrase. Having recognised it as a phrase, we must treat its words as parts, not directly of the sentence, but of the phrase itself. This phrase is intermediate between the sentence and its words, just as wheels are intermediate between the bicycle and its spokes. Since we can't omit any of those three words individually, it appears that, while the phrase as a whole is optional in the structure of the sentence, the words themselves are not optional in the structure of the phrase.

In sentence [17] below, there are two separate sequences of words which can be omitted without affecting the grammaticality of the sentence. Can you identify them?

[17] The very muscular gentleman next to me lit a cigar.

[18], [19], and [20] are all perfectly good, complete sentences.

[18] The (...) gentleman next to me lit a cigar.
[19] The very muscular gentleman (...) lit a cigar.

So we need to be able to say that very muscular (omitted in [18] and [20]) and next to me (omitted in [19] and [20]) are optional constituents in the structure of sentence [17]. But they are not sentences and they are not words. They are phrases – elements of structure intermediate between sentence and word. Furthermore, we will see in due course that these phrases are immediate constituents, not of the sentence, but of yet further phrases within the sentence. They are phrases within phrases.

If a sequence of words can be omitted from a sentence leaving another good sentence, that's a good indication that the sequence is a phrase functioning as a constituent in the structure of the sentence. However, not all phrases are omissible. So we need to find a more general, systematic way of demonstrating that a given sequence of words is a phrase.

There are several different ways of doing this. Recall that we were never in doubt that invitingly was a constituent in [16]. It is a single word, after all. And we wanted to say of the sequence of words beside a stream that it had the same unitary character as a single word. This suggests that if you can replace a sequence of words in a sentence with a single word without changing the
overall structure of the sentence, then that sequence functions as a constituent of the sentence and is therefore a phrase. This test will confirm that *beside a stream* is functioning as a constituent in sentence [4]. For example, if the speaker of sentence [4] were in a position to point to the spot where Sam sunbathed, she could replace *beside a stream* by *here* or *there*:

[21] Old Sam sunbathed here/there.

Or she could be less specific, by replacing *beside a stream* with *somewhere*.

[22] Old Sam sunbathed somewhere.

Questions offer a clear example of this. We can form a question from [4] by replacing *beside a stream* with the question word *where* as in [23] and [24]:


Since we have used *where* to replace *beside a stream*, it’s natural that *beside a stream* should be a possible answer to the question. Answering such questions is a matter of replacing the question word with an informative phrase. So, answers to ‘WH’ questions (that is, questions that contain one of the question words *who*, *which*, *what*, *why*, *where*, *when*, *whose*, and *how*) are phrases.

All this justifies analysing *beside a stream* as a phrase. The question now is: How should we represent this phrase in terms of a phrase marker? As with the whole sentence, we need to know whether the words of the phrase are its immediate constituents, or whether it contains further phrases. There are just three phrase markers that could possibly represent the structure of *beside a stream*:

```
[25a] Phrase
   Phrase
   beside  a  stream

[25b] Phrase
   Phrase
   a  beside  stream

[25c] Phrase
   Phrase
   a  stream
```

Each gives a different analysis. Which do you think is the best representation of the structure of the phrase? In coming to a decision, ask yourself whether *a* belongs more with *beside* than with *stream* ([25a]), more with *stream* than with *beside* ([25b]), or whether it doesn’t seem to belong more with one than the other ([25c]). The question is: Does the phrase *beside a stream* include a further phrase? If it doesn’t, then [c] is right. But if it does, then either [a] or [b] is right – and the question is: which?

Now check that the tests mentioned above, replacement by a single word and the question test, confirm the analysis you have chosen.
Phrase marker [25c] says that the phrase does not contain any further phrase, that the words themselves are the immediate constituents of the phrase. According to [c], a does not belong more with either of the other words. Now, if [25c] is correct, [a] and [b] should seem equally bad. Well, I hope you agree that [a] is really bad. [a] suggests that we could find a single word to replace the supposed phrase *beside a*. It is difficult to imagine what word could replace that sequence. It seems incomplete and it’s impossible to say what it means. On the other hand, a *stream* does seem complete, it is fairly clear what it means, and we don’t have to rack our brains to find single words that could replace it – for example, *it, something, or one*. These yield good phrases: *beside it, beside something, and beside one*.

Notice, too, that if we were to change singular *stream* to plural *streams*, we would get the ungrammatical word-sequence *beside a streams* – unless we also omit a (to give *beside streams*). This strongly suggests that a belongs definitely with *stream* rather than with *beside*, that a is dependent on *stream*. Here, again, we are using the single word *streams* to replace the sequence *a stream*.

The question test, too, confirms that *a stream* is a phrase:

[26] Question: [a] Old Sam sunbathed beside what?
[b] What did old Sam sunbathe beside?

Answer: A stream.

Notice that there is no question to which *beside a* would be a suitable answer.

[27] provides further evidence that *a stream* forms a phrase, since it has been moved as a unit in forming a new construction.

A *stream* is what old Sam sunbathed beside.

It is worth noting, then, that the movement of a sequence of words in forming a construction indicates that the sequence is a phrase. As a further example, note the acceptability of moving *beside a stream* to the beginning of sentence [4]:

[28] Beside a stream, old Sam sunbathed.

In short, the various kinds of evidence discussed confirm that [25b] is the correct representation of the structure of our phrase. It shows a phrase within a phrase.

As an exercise, think of some other possible answers to the *what* question in [26]. They can be as different as you like from the answer already given, and they can be as long as you like. Provided they do not sound ungrammatical, every sequence of words you choose will be a phrase.

Here are some suggestions:

[29a] a large pile of Bokhara rugs
[29b] the magnolia bush at the bottom of his garden
[29c] an unreliable puppy that was taking the occasional nip at his toes.
All these are phrases. They could all serve as answers to the *what* question, and they are all replaceable by a single word. Furthermore, they all contain further phrases.

Earlier, when we were considering whether there was a single word that could be used to replace the sequence *beside a*, I mentioned meaning and implied that phrases form not only syntactic units (constituents in the structural form of sentences) but also semantic units. In other words, they form identifiable parts of the meaning of sentences; they form coherent units of sense. It is reasonable to ask what *beside a stream* and *a stream* mean, but it is not reasonable to ask what *beside a* means; it has no meaning.

Does the discussion so far suggest an explanation why [13] on page 11 seems more acceptable than those in [10], [11], and [14]? How, exactly?

I put an asterisk in front of [13] because it was not a complete sentence. However, it is a complete phrase, and in this it contrasts with the other strings. *Old Sam* could be replaced by a single word – *he, someone*, or even just *Sam* – making no difference to the overall structure of the sentence. Furthermore, *old Sam* could be used as an answer to the question *Who sunbathed beside a stream?*, where I have replaced the sequence *old Sam* with the single ‘WH’ word *who*.

‘Phrase’ and ‘constituent’

I have said that a *phrase* is a sequence of words that can function as a constituent in the structure of sentences. The important word here is ‘can’.

We have seen that *beside a stream, a stream*, and *old Sam* can function as constituents in sentence structure – and they do function as constituents in sentence [4] and many other sentences. They are therefore phrases. The fact that those word-sequences are constituents in sentence [4], however, doesn’t mean they function as constituents of every sentence in which they appear. Here, as an obvious example, is a sentence in which the word-sequence *old + Sam* is definitely not a constituent:

[30] Though he was old Sam did regular press-ups.

This is clear when we try to replace that sequence with a single word:

[31] *Though he was someone did regular press-ups.
[32] *Though he was who did regular press-ups?

Out of the context of any particular sentence, *old Sam* is a phrase. It is a phrase of English because it *can* be a constituent of an English sentence. But that word-sequence is not a constituent of *every* sentence in which it appears. It is not a constituent of sentence [30], for example.
So: although old Sam is indeed a phrase, it’s not a phrase that actually figures in the structure of [30]! As I mentioned in the Introduction, in a quite literal sense there’s more to syntax, and to your own understanding of sentences, than meets the eye. Hierarchical sentence structure is really quite abstract. It is not there visibly on the page. It’s in your head. **Your understanding of particular word-sequences is matter of how you structure them in your mind.** That is why syntax is interesting. And it is why we need to construct concrete phrase markers to represent this abstract mental structure.

Consider now sentence [33] and decide whether the sequence \( a + \text{stream} + \text{that} + \text{had} + \text{dried} + \text{up} \) is a constituent or not.

[33] Sam sunbathed beside a stream that had dried up.

That sequence of words would be a perfectly good answer to the question *What did old Sam sunbathe beside?* Furthermore, it’s replaceable by a single word while preserving the overall structure of the sentence. So it is a constituent of [33]. And, just as with \( a \text{ stream} \) in sentence [4], it forms a further phrase with *beside*. This further phrase can be represented as in [34]:

\[
\text{[34]}
\]

In [34] I have adopted the useful convention of using a **triangle** to represent a constituent when I am not concerned with its internal structure. For ease of reference, I have distinguished the phrases by letter.

The question I want you to consider now is this: Does the sequence \( \text{beside} + a + \text{stream} \) – which formed a constituent in sentence [4] – form a constituent in sentence [33]? And if not, why not? The phrase marker [34] should help you to answer this.

You have probably guessed that the answer is ‘No’: \( \text{beside} + a + \text{stream} \) is not a constituent in [33]. Why not? Well, we agreed that in [33]/[34] \( a + \text{stream} \) is part of a larger phrase, but that larger phrase is not \( \text{beside} + \text{a stream} \) – it’s \( \text{a stream that had dried up} \). *Beside* forms a phrase, not with \( a + \text{stream} \), but with the sequence \( \text{a stream that had dried up} \). The words \( a \) and \( \text{stream} \) are part of PHRASE-b. If an element (word or phrase) is part of a phrase, it can only relate to other elements within that same phrase. If we wanted to say that \( \text{beside} + \text{a stream} \) formed a phrase in [33], we would be forced to represent the complete phrase \( \text{beside a stream that had dried up} \) as in [35]:

\[
\text{[35]}
\]
But [35] is wrong; it fails to represent \textit{a stream that had dried up} as a phrase. The moral is that an element can belong directly only to one phrase at a time. I say ‘directly’ since in [34], for example, \textit{a stream} belongs both to PHRASE-b (directly) and to PHRASE-a (indirectly). It is, in fact, impossible to draw a phrase marker that says of \textit{a stream} that it simultaneously forms a phrase directly with \textit{beside} and with \textit{that had dried up}.

You may be uncertain whether or not a given sequence of elements is represented as a phrase by a phrase marker. Before explaining this, I need to introduce some terminology that helps in finding our way around phrase markers. Here goes.

Any point in a phrase marker that could branch and bear a label is called a ‘\textit{node}’. In phrase marker [34] there are two nodes, labelled ‘PHRASE-a’ and ‘PHRASE-b’. A node is said to \textit{dominate} everything that appears below it and joined to it by a line. Thus the node labelled ‘PHRASE-a’ dominates all the following elements: \textit{beside}, PHRASE-b, \textit{a}, \textit{stream}, \textit{that}, \textit{had}, \textit{dried}, and \textit{up}. A node is said to \textit{immediately dominate} another element when there are no intervening elements. Thus PHRASE-a in [34] immediately dominates just \textit{beside} and PHRASE-b. PHRASE-a dominates \textit{stream} but it does not immediately dominate it, because the node labelled ‘PHRASE-b’ intervenes.

Using this terminology, I can now show how to decide whether a sequence of elements is represented as a constituent in a phrase marker. In a phrase marker, a sequence of elements is represented as a \textit{constituent} if there is a node that dominates all those elements and no others. In other words, if you can trace just the elements under consideration (i.e. all those elements and only those elements) up to a single node, then those elements are represented as a constituent (a phrase).

Look at [34]. The sequence \textit{a} + \textit{stream} + \textit{that} + \textit{had} + \textit{dried} + \textit{up} is represented as a constituent because the elements (words, in this case) can all be traced back to a single node that does not dominate any other element, namely, PHRASE-b. The sequence \textit{beside} + \textit{a}, on the other hand, is not represented as a constituent because the only node that dominates both of those words (namely, PHRASE-a) dominates other elements as well (namely, \textit{stream}, \textit{that}, \textit{had}, \textit{dried}, and \textit{up}). Similarly, in the incorrect phrase marker [35], \textit{a stream that had dried up} is not represented as a constituent because there is no node that dominates all and only those words. The only node that dominates all of them is PHRASE-a, but PHRASE-a doesn’t dominate only those words, it also dominates \textit{beside}.
I’ve given two examples in which a sequence of words functioning as a constituent in one sentence does not function as a constituent in another. Here, as a final example, is what is known as a structurally ambiguous sentence. On one interpretation, the sequence old + Sam does function as a constituent but on the other interpretation it doesn’t:

[36] Heseltine asked how old Sam was.

Try to identify the two meanings of [36]. A good way of doing this is to decide on the exact question which Heseltine is reported in [36] to have asked. You may find it helpful to make a written note of the two questions.

Having identified the two meanings in the way suggested, you should not have much difficulty in deciding which interpretation demands that the sequence does form a constituent and which demands that it does not.

The two quite different questions that could have been asked by Heseltine are [a] How old is Sam? and [b] How is old Sam? As these different questions show, on the first interpretation, [a], old belongs with how to form the phrase how old. In this question, the phrase as a unit has been moved from its position at the end of the sentence (Sam is how old?). On this interpretation, since old forms a constituent with how, it simply cannot also form a constituent with Sam. It is on the second interpretation, [b], that old and Sam go together, forming a phrase. This example illustrates how deciding what phrases there are in the sentence is a crucial part of deciding what the sentence actually means.

Most people, when presented with a sequence of words out of the context of any sentence, have feelings as to whether that sequence could function as a constituent in a sentence (i.e. whether it is a phrase) – at least once they start thinking about it (as you are being encouraged to do here). It is usually simply a matter of deciding whether it seems to you to form a unit of sense. In the main, this is a reliable guide as to whether that sequence actually is a constituent in a sentence to be analysed, though, as we have seen from the last three examples, not one hundred per cent reliable. And, even in the context of a sentence, you will find that you do have an intuitive feeling as to which sequences are functioning as its constituents. In this chapter I have considered various kinds of evidence for constituents – omission, replacement by a single word, the question test, movement, the sense test. These are useful in confirming your intuitions, and in checking on cases where you are in doubt – one’s first intuitions are not always strong and not always reliable.
Exercises

1. Look again at the discussion on page 17 above and then, on the basis of the tree diagram below, say which of the following sequences are constituents of A.

   (1) $c + d$.  (2) $a + b + c$.  (3) $c + d + e + f$.  (4) $e + f$.  (5) $e + f + g + h$.
   (6) $g + h$.  (7) $E + C$.  (8) $D + E$.  (9) $F + g + h$.

   ![Tree Diagram](image)

2. In tree diagram (a) above, what are the immediate constituents of:
   (1) A?  (2) B?  (3) C?

3. (a) Draw a phrase marker for the phrase *their rather dubious jokes* which shows that it contains the further phrase *rather dubious jokes*, which in turn contains *rather dubious* as a phrase.
   (b) *Men from the Ministry* is a phrase which contains *from the Ministry and the Ministry* as phrases. Draw a phrase marker for the whole phrase.

4. Decide whether the italicised strings in the following sentences are constituents of those sentences or not. Note that (g) is ambiguous; as with the ambiguous example discussed in this chapter, you should identify the two interpretations and say on which interpretation the italicised sequence forms a constituent.
   (a) John considered *visiting his great aunt*.
   (b) Maria simply gazed *at the bollard she had just demolished*.
   (c) Maria simply gazed *at the bollard* she had just demolished.
   (d) *In the machine* the gremlin could be heard juggling with ball-bearings.
   (e) *In the machine the gremlin* could be heard juggling with ball-bearings.
   (f) Rory put *a silencer on the gun*.
   (g) Sam managed to touch *the man with the umbrella*.

5. In the light of the discussion of this chapter, how many constituents can you identify in sentence (a) given that the much shorter (b) is a grammatical sentence? (Don’t attempt a complete analysis of sentence (a) – the fact that sentence (b) is well-formed doesn’t provide enough information for that.)
(a) Being of a cautious disposition, Timothy very wisely avoided the heavily built man whenever he drank at the Wrestler’s Arms.

(b) Timothy avoided the man.

6. I’ve not yet provided a complete analysis of sentence [4]. We have agreed that old Sam, beside a stream, and a stream are among its constituent phrases. So we can at least draw an incomplete phrase marker for it, as in (a):

```
(a) Sentence
     /\   \n    /    \   \n   Phrase       Phrase
   /\   \     /\   \     /\   \     /\   \     /\   \   
old Sam sunbathed beside a stream
```

We know that the complete string constitutes a sentence. In a complete phrase marker, then, all the elements must be joined up to the Sentence node in some way. The question is: How? There are three ways in which this could be done. Each way offers a different analysis of the sentence – a different analysis of how sunbathed fits into the structure and a different account of the immediate constituents of the sentence. Draw the three different phrase markers and explain in words (using ‘constituent’ and ‘immediate constituent of the sentence’) what different claims are made about the structure of the sentence by each phrase marker. (Make sure the phrases we have already acknowledged remain represented as phrases in your complete phrase markers!) I’m not here asking you to choose which analysis you think is best – though I hope you have views on the matter. In fact, all three analyses have been proposed at one time or another, though one of them is most generally accepted these days and it is this that I shall adopt in the next chapter.

Discussion of exercises

1. (1) Yes. Both c and d – and only c and d – can be traced back to node E.

   (2) No. D dominates a and b but not c. Node B does dominate a, b and c, but it also dominates d; so there is no node that dominates all and only a, b, and c.

   (3) No. No single node that dominates all and only c, d, e, and f. Only A dominates them all, but A dominates a, b, g, and h too.

   (4) Yes. e and f (and only e and f) can be traced back to the single node F.

   (5) Yes. They alone can all be traced back to C.

   (6) No. No. (8) Yes. (9) Yes.

2. (1) B and C. (2) D and E. (3) F, g, and h.
EXERCISES

3. (a) Yes. It could be replaced by it and by what in forming the question What did he consider?, to which visiting his great aunt is a possible answer. (Note also that the sequence moves as a unit in forming the construction Visiting his great aunt is what he considered.)

(b) Yes. (cf. she simply gazed at it. What did she gaze at? Answer: the bollard she had just demolished.)

(c) No. In (b) above, the sequence the + bollard was shown to be part of the phrase the bollard she had just demolished; it cannot then form a constituent with from. (See the discussion of beside a stream that had dried up [33] in the chapter, pp. 16–17.)

(d) Yes. It could be replaced by there or somewhere. Furthermore, in the machine is a good answer to the question Where could the gremlin be heard juggling with ball-bearings? Finally, the sequence could be omitted leaving a well-formed sentence.

(e) No. There is no question that In the machine the gremlin could possibly be an answer to. Who/What could be heard . . . ? could receive the gremlin as a possible answer; Where could the gremlin be heard . . . could receive In the machine. Each of these, then, are phrases. But there is no single question word that covers both where and what. So here we have a sequence of phrases here but those two phrases don’t make up a further phrase.

(f) No. Consider the oddity of *Rory put it and *Rory put something. And the oddity of *What did Rory put? (Answer: *A silencer on the gun.)

(g) On one interpretation the sequence is a constituent, cf. Sam managed to touch him and Who did Sam manage to touch? (Answer: The man with the umbrella.) On the other interpretation, it is not a single phrase but a sequence of two phrases. Cf. Sam managed to touch him with an umbrella, Who did Sam manage to touch with an umbrella? (Answer: the man.)

5. The fact that (b) is a well-formed sentence allows us to infer that every sequence of words omitted from (a) in order to form (b) can be counted as a constituent of (a). These are:

Being of a cautious disposition
very wisely
heavily built
whenever he drank at the Wrestler’s Arms.
There are other constituents in the (a) sentence, of course, and the constituents listed here themselves contain further phrases as constituents.

6. Here are the three complete phrase markers. New bits are in bold.

(a) represents *sunbathed* as forming a constituent with *beside a stream*, and divides the sentence into just two immediate constituents: *old Sam* and *sunbathed beside a stream*. (b) also divides the sentence into two, but this time the two parts are *old Sam sunbathed* and *beside a stream*. Phrase marker (c) represents the sentence as having three immediate constituents, *old Sam* and *sunbathed* and *beside a stream*; it says that *sunbathed* forms a constituent neither with *old Sam* nor with *beside a stream*.

In attempting to represent what phrase marker (a) represents, you may have been tempted simply to draw an extra line out from the phrase node dominating *beside a stream* as (d):
But (d) is incorrect. Can you see why? (Check the discussion on page 17.) Although it associates sunbathed with beside a stream, it fails to represent beside a stream as a phrase in its own right, independently of sunbathed. It fails to do this because there is no node that dominates all and only beside + a + stream. (The only node that dominates them all dominates sunbathed as well.) Check that you have not succumbed to a similar temptation in connection with (b).

Further exercises

1. The structural ambiguity of [36] in the text is a matter of whether old Sam or how old is a constituent. All the following are structurally ambiguous. In each case, identify the source of the ambiguity in terms of two different constituent analyses, as I have just done with [36].

   (1) This story shows what evil men can do.
   (2) They only sell rotten fruit and vegetables.
   (3) More interesting meals would have been appreciated.
   (4) We need an agreement between workers on overtime.
   (5) Bill asked the man who he had seen.

2. Draw a phrase marker for the phrase no previous experience of syntax, showing that it contains the phrase previous experience of syntax as a constituent, which in turn has the phrase experience of syntax as a constituent, which in turn has the phrase of syntax as a constituent (which, of course, is made up by of and syntax).

3. The new students are very worried is a sentence. Assume that it has two phrases as immediate constituents: the new students and are very worried. Furthermore, assume that the new students consists of the word the and the phrase new students. And that are very worried consists of the word are and the phrase very worried. Try drawing the phrase marker for the sentence in the light of all that.