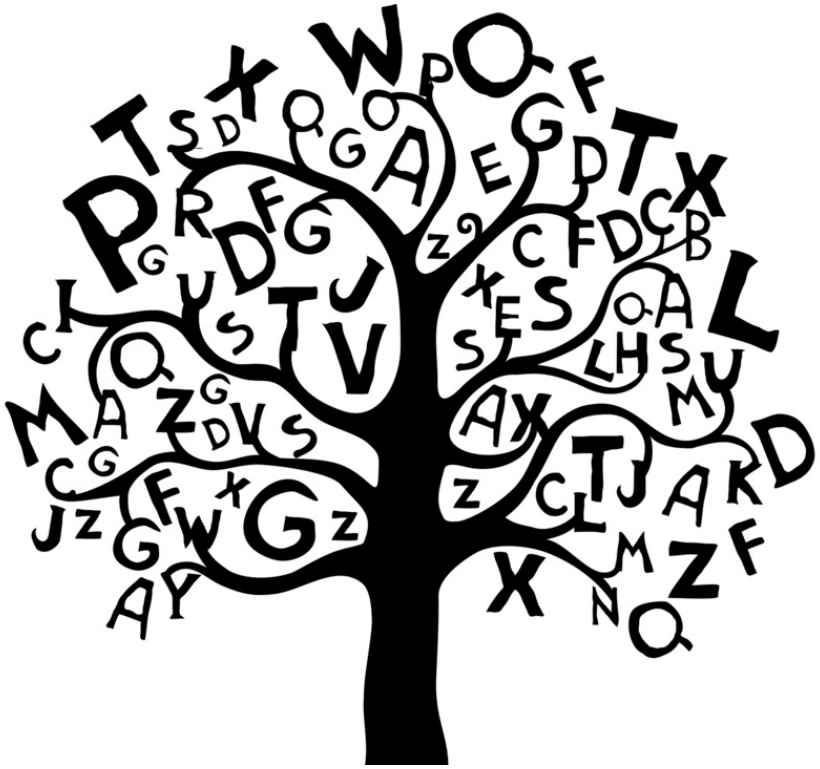
8

Syntax



Time flies like an arrow; fruit flies like a banana.

Oettinger (1966)

In an early observation on the difficulties of getting computers to process natural language, Anthony Oettinger used the example above to illustrate

how we tend to interpret sentences based on an expected structure and when we arrive at a problematic interpretation, we are able to go back and try to use a different structure. This process brings to light the importance of recognizing the underlying structure of sentences in order to make sense of them. If we keep thinking that the structure of the second expression is the same as the first in the example, we will definitely miss something. (For a helpful analysis, see Figure 8.9, on page 122.)

In Chapter 7, we moved from the general categories of traditional grammar to more specific methods of describing the structure of phrases and sentences. When we concentrate on the structure and ordering of components within a sentence, we are studying the **syntax** of a language. The word "syntax" comes originally from Greek and literally means "a putting together" or "arrangement." In earlier approaches, there was an attempt to produce an accurate description of the sequence or ordering "arrangement" of elements in the linear structure of the sentence. In more recent attempts to analyze structure, there has been a greater focus on the underlying rule system that we use to produce or "generate" sentences.

Syntactic rules

When we set out to provide an analysis of the syntax of a language, we try to adhere to the "all and only" criterion. This means that our analysis must account for *all* the grammatically correct phrases and sentences and *only* those grammatically correct phrases and sentences in whatever language we are analyzing. In other words, if we write rules for the creation of well-formed structures, we have to check that those rules, when applied logically, won't also lead to ill-formed structures.

For example, we might say informally that, in English, we put a preposition (*near*) before a noun (*London*) to form a prepositional phrase (*near London*). This will describe a large number of phrases, but does it describe all (and only) the prepositional phrases in English? Note that, if we use this as a rule of the grammar to create structures involving a preposition and a noun, we will end up producing phrases like **near tree* or **with dog*. These don't seem to be well-formed English structures, so we mark them with an asterisk *, indicating that they are ungrammatical.

We clearly need to be more careful in forming the rule that underlies the structure of prepositional phrases in English. We might have more success with a rule stating that we put a preposition before a noun phrase (not just a noun). In Chapter 7, we saw that a noun phrase can consist of a proper noun (*London*), a pronoun (*me*) or the combination of an article (*a, the*) with a noun (*tree, dog*), so that the revised rule can be used to produce these well-formed structures: *near London, with me, near a tree, with the dog*.

A Generative Grammar

When we have an effective rule such as "a prepositional phrase in English consists of a preposition followed by a noun phrase," we can imagine an extremely large number of English phrases that could be produced using this rule. In fact, the potential number is unlimited. This reflects another goal of syntactic analysis, which is to have a small and finite (i.e. limited) set of rules that will be capable of producing a large and potentially infinite (i.e. unlimited) number of well-formed structures. This small and finite set of rules is sometimes described as a **generative grammar** because it can be used to "generate" or produce sentence structures and not just describe them.

This type of grammar should also be capable of revealing the basis of two other phenomena: first, how some superficially different phrases and sentences

are closely related and, second, how some superficially similar phrases and sentences are in fact different.

Deep and Surface Structure

Our intuitions tell us that there must be some underlying similarity involving these two superficially different sentences: *Charlie broke the window* and *The window was broken by Charlie*. In traditional grammar, the first is called an active sentence, focusing on what *Charlie* did, and the second is a passive sentence, focusing on *The window* and what happened to it. The distinction between them is a difference in their **surface structure**, that is, the different syntactic forms they have as individual English sentences. However, this superficial difference in form disguises the fact that the two sentences are closely related, even identical, at a less superficial level.

This other "underlying" level, where the basic components (Noun Phrase + Verb + Noun Phrase) shared by the two sentences can be represented, is called their **deep structure**. The deep structure is an abstract level of structural organization in which all the elements determining structural interpretation are represented. That same deep structure can be the source of many other surface structures such as *It was Charlie who broke the window* and *Was the window broken by Charlie?*. In short, the grammar must be capable of showing how a single underlying abstract representation can become different surface structures.

Structural Ambiguity

Let's say we have two distinct deep structures. One expresses the idea that "Annie had an umbrella and she bumped into a man with it." The other expresses the idea that "Annie bumped into a man and the man happened to be carrying an umbrella." Now, these two different versions of events can actually

be expressed in the same surface structure form: *Annie bumped into a man with an umbrella*. This sentence provides an example of **structural ambiguity**. It has two distinct underlying interpretations that have to be represented differently in deep structure. Note that this is not the type of ambiguity that we experience in hearing *Their child has grown another foot*, which illustrates lexical ambiguity mainly because the word *foot* has more than one meaning. (See Task H, page 120, for further analysis.)

The comedian Groucho Marx knew how to have fun with structural ambiguity. In the film *Animal Crackers*, he first says *I once shot an elephant in my pajamas*, then follows it with *How he got into my pajamas I'll never know*. In the non-funny interpretation, part of the underlying structure of the first sentence could be something like: "I shot an elephant (while I was) in my pajamas." In the other (ho, ho) interpretation, part of the underlying structure would be something like: "I shot an elephant (which was) in my pajamas." There are two different underlying structures with the same surface structure, revealed by syntactic analysis.

Syntactic Analysis

In syntactic analysis we use some conventional abbreviations for the parts of speech identified in Chapter 7. Examples are N (= noun), Art (= article), Adj (= adjective) and V (= verb). We also use abbreviations for phrases, such as NP (= noun phrase) and VP (= verb phrase). In English, the verb phrase (VP) consists of the verb (V) plus the following noun phrase (NP). We can take the simple sentence from Table 7.3 (page 98) and label the constituents using these categories, as in Figure 8.1.

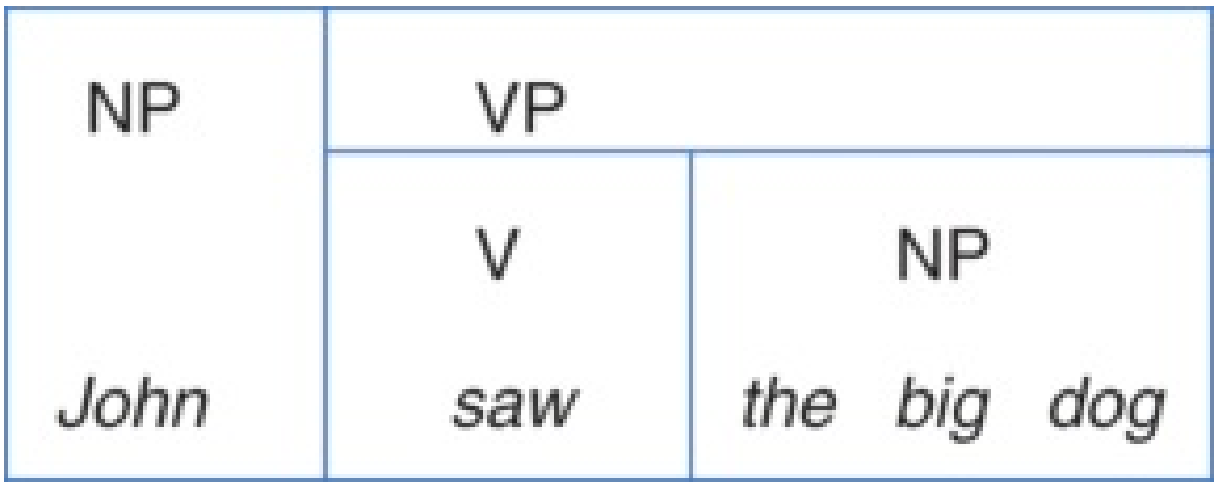


Figure 8.1

Figure 8.1 presents a static analysis of a single sentence. We would like to be able to represent the same syntactic information in a more dynamic format. One way of presenting the concept "consists of" is with an arrow (→), also interpreted as "rewrites as." The following rule states that a noun phrase (NP) such as *the dog* consists of or rewrites as (→) an article (*the*) and a noun (*dog*). This simple formula is the underlying structure of millions of different English phrases.

NP → Art N

However, it is not the only form a noun phrase can take. We want to be able to include another constituent (Adj) in the rule so that it is good for not only phrases like *the dog*, but also *the big dog*. This constituent is *optional* in a noun phrase, so we use round brackets to indicate that Adj is an optional constituent, as shown here:

NP → Art (Adj) N

Another common symbol is in the form of curly brackets **{}**. These indicate that *only one* of the elements enclosed within the curly brackets must be selected. We have already seen, in Figure 7.3, on page 96, that a noun phrase can

also contain a pronoun (*it*), or a proper noun (*John*). Using the abbreviations "Pro" (for pronoun) and "PN" (for proper noun), we can write three separate rules, as shown on the left, but it is more succinct to write one rule, on the right, using curly brackets.

NP → Art (Adj) N

NP → Pro

NP → PN

NP → {Art (Adj) N, Pro, PN}

Phrase Structure Rules

What we have started to create is a set of syntactic rules called **phrase structure rules**. As the name suggests, these rules state that the structure of a phrase of a specific type will consist of one or more constituents in a particular order.

The first rule in the following set of simple (and necessarily incomplete) phrase structure rules captures a very general rule of English sentence structure: "a sentence (S) rewrites as a noun phrase (NP) and a verb phrase (VP)." The second rule states that "a noun phrase rewrites as either an article plus an optional adjective plus a noun, or a pronoun, or a proper noun." In the third rule, a verb phrase rewrites as a verb plus a noun phrase.

S → NP VP

NP → {Art (Adj) N, Pro, PN}

VP → V NP

Lexical Rules

Phrase structure rules generate structures. In order to turn those structures into recognizable English, we also need **lexical rules** that specify which words can be used when we rewrite constituents such as PN. The first rule in the following set states that "a proper noun rewrites as *John* or *Mary*." (It is a very small world.)

PN → {*John, Mary*}	Art → {*a, the*}
N → {*girl, dog, boy*}	Adj → {*big, small*}
V → {*followed, helped, saw*}	Pro → {*it, you*}

We can rely on these rules to generate the grammatical sentences shown below in (1)–(6), but not the ungrammatical sentences shown in (7)–(12).

(1) *A dog followed the boy.*	(7) **Dog followed boy.*
(2) *You saw it.*	(8) **You it saw.*
(3) *John saw the big dog.*	(9) **John Mary small dog*
(4) *It followed Mary.*	(10) **Followed Mary the dog big.*
(5) *The small boy helped you.*	(11) **The helped you boy*
(6) *Mary helped John.*	(12 **Mary John helped.*

Tree Diagrams

One of the best ways to create a visual representation of underlying syntactic structure is through **tree diagrams**. We can use the symbols introduced earlier to label parts of the tree when we create a representation of how each part fits into

the underlying structure of phrases. The information in a phrase structure rule, on the left, can be expressed in a tree diagram, on the right, as in Figure 8.2.

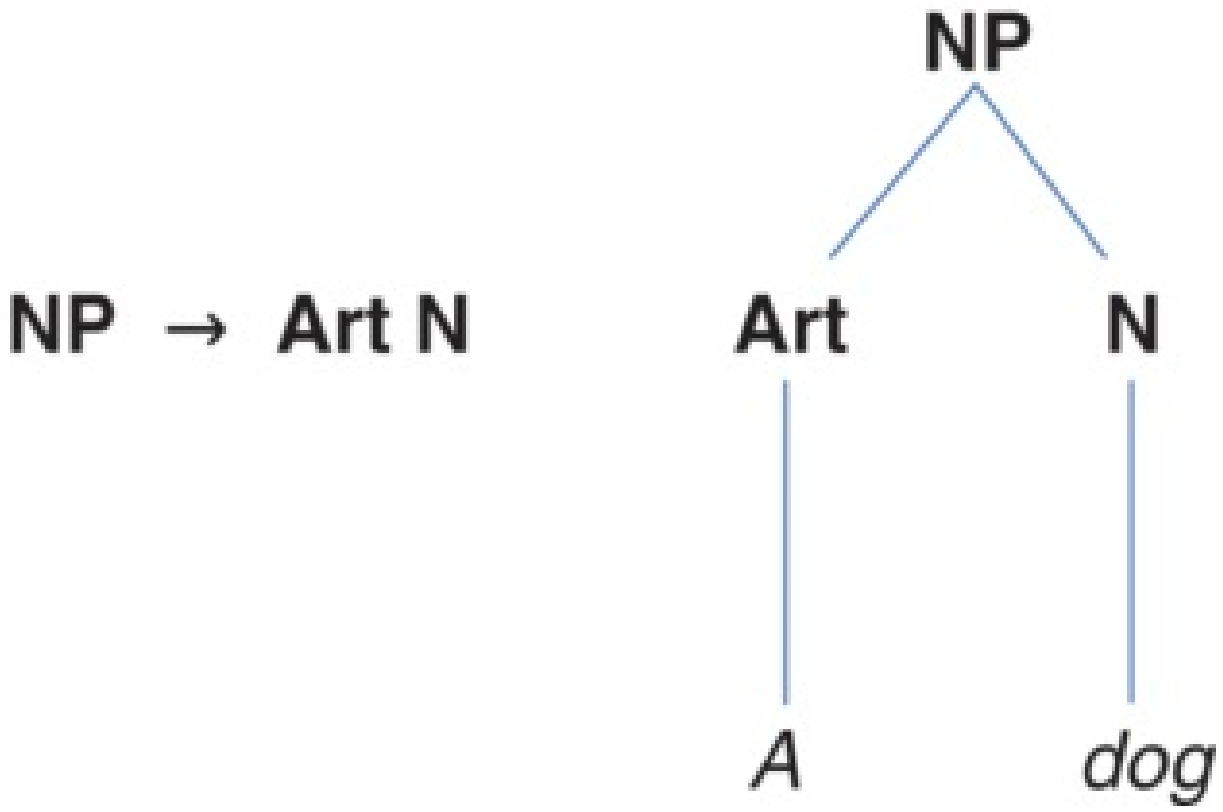


Figure 8.2

Although this kind of "tree," with its "branches," on the right, seems to grow down rather than up, it functions rather well as a diagram representing all the grammatical information found in the other analysis on the left. It also shows very explicitly that there are different levels in the analysis. That is, there is a level of analysis at which a constituent such as NP is represented and a different, lower, level at which a constituent such as N is represented.

We can use a similar tree diagram to represent the more complex structure of an English verb phrase (VP), as shown in Figure 8.3. Once again, this type of diagram provides a way of representing the hierarchical nature of underlying structure. In this hierarchy, the verb phrase (VP) in higher than and contains the

verb (V) and a noun phrase (NP). The noun phrase (NP) is higher than and contains the article (Art) and the noun (N).

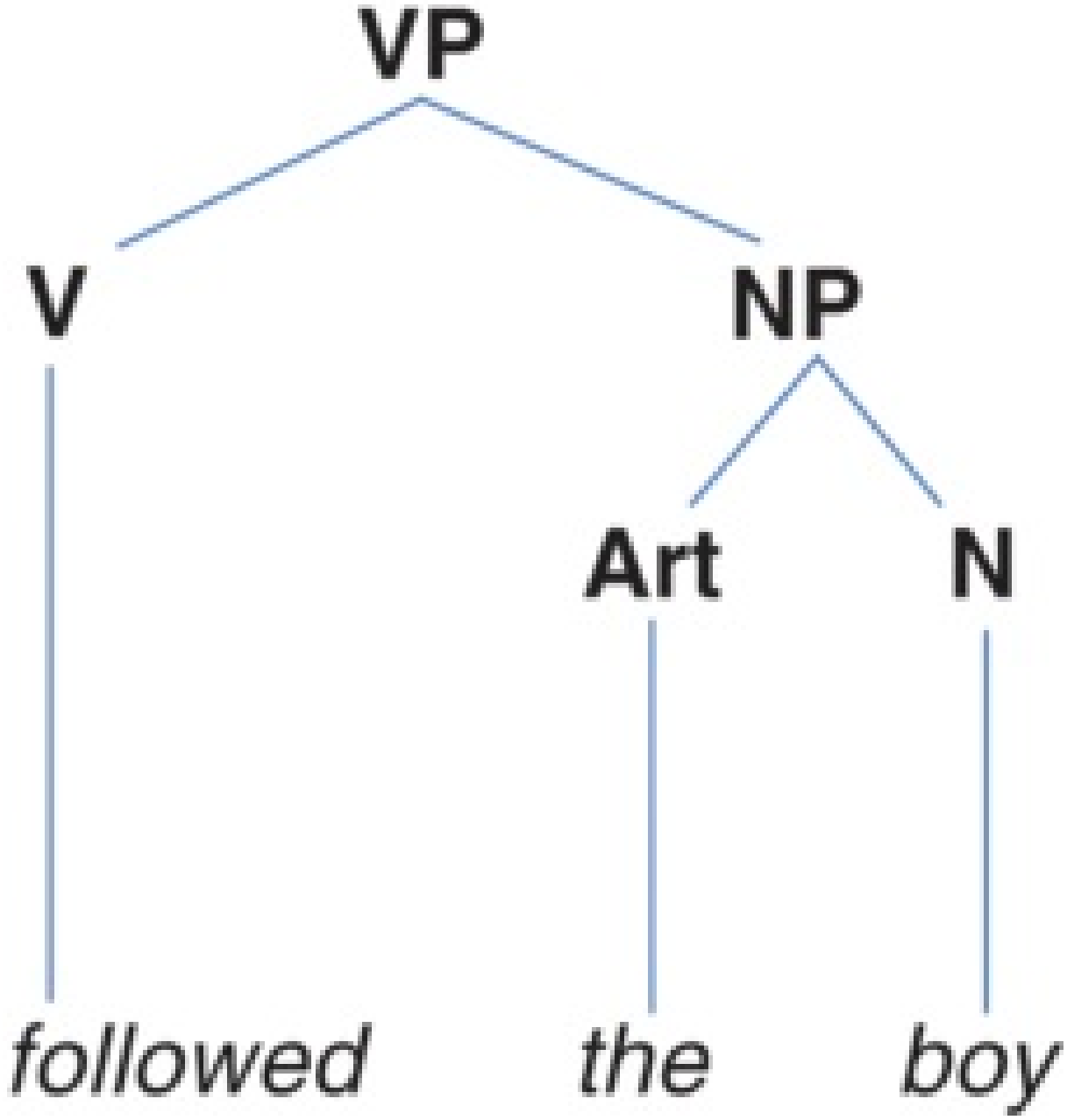
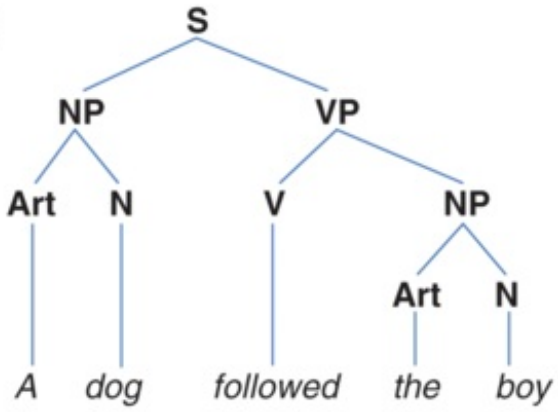


Figure 8.3

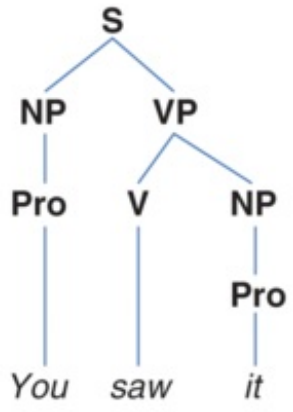
Tree Diagrams of English Sentences

We can now put together tree diagrams for whole sentences, hierarchically organized, as shown in Figure 8.4. Notice that essentially the same basic tree diagram structure is the foundation for all the different sentences (1)–(6), from page 112, with variable constituents included in each one.

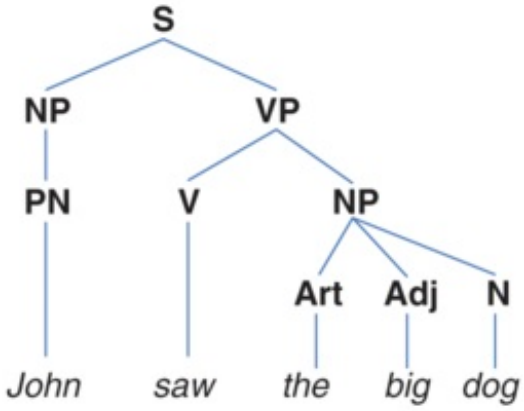
(1)



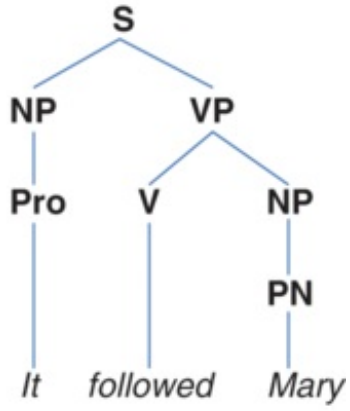
(2)



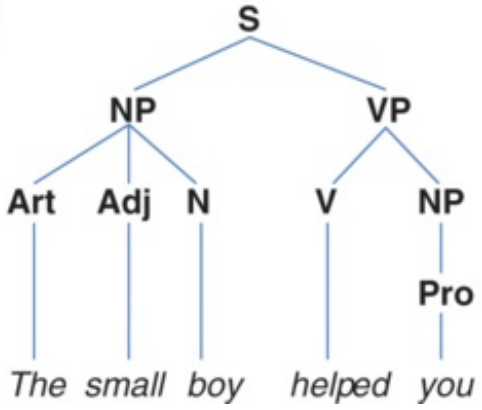
(3)



(4)



(5)



(6)

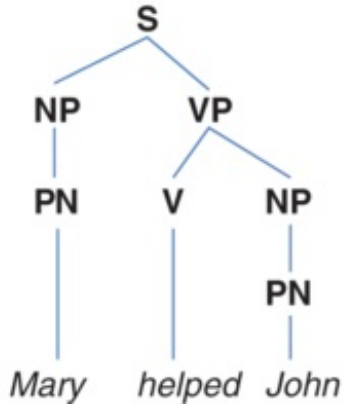


Figure 8.4

Just Scratching the Surface

At the bottom of all the trees in Figure 8.4 are surface structure variations of a single underlying deep structure, revealing the generative power of the phrase structure rules involved. There are other phrase structure rules involved in the composition of more complex sentences. Some are presented in Task C on page 117 and Task I on page 120 for English, and Tasks E and F, on pages 118–119 for other languages. As we try to develop better ways of analyzing the syntactic structure of complex sentences, we inevitably need a larger analytic framework. (We have barely scratched the surface structures.) However, having explored some of the basic issues, terminology, representations and methods of syntactic analysis in order to talk about basic structures in the English language, we will now move on to consider how we might incorporate the analysis of meaning in the study of language.

Study Questions

1 What was the original literal meaning of *syntax* in Greek?

2 What is wrong with the following rule of English syntactic structure? "A prepositional phrase rewrites as a preposition followed by a noun."

3 Which of the following expressions are structurally ambiguous and in what way?

(a) *These are designed for small boys and girls.*

(b) *The parents of the bride and groom were waiting outside.*

(c) *How come a bed has four legs, but only one foot?*

(d) *We met an English history teacher.*

(e) *Flying planes can be dangerous.*

(f) *The students complained to everyone that they couldn't understand.*

4 What part of speech is *lovely* in the following sentence?

We saw a lovely rainbow yesterday.

5 How many noun phrases are there in the following sentence?

George saw a small dog in the park near the fountain and it followed him when he left the park.

6 Which part of the following sentence is the VP?

None of the people in the building supported the proposed rent increase.

7 Which of the following expressions would be generated by this phrase structure rule: NP → {Art (Adj) N, Pro, PN}?

(a) *a lady*	(c) *her*	(e) *the widow*
(b) *the little girl*	(d) *Annie*	(f) *she's an old woman*

8 What kind of generative rule is this: N → {*girl, dog, boy*}?

9 Do phrase structure rules represent deep structure or surface structure?

10 Complete the following tree diagrams.

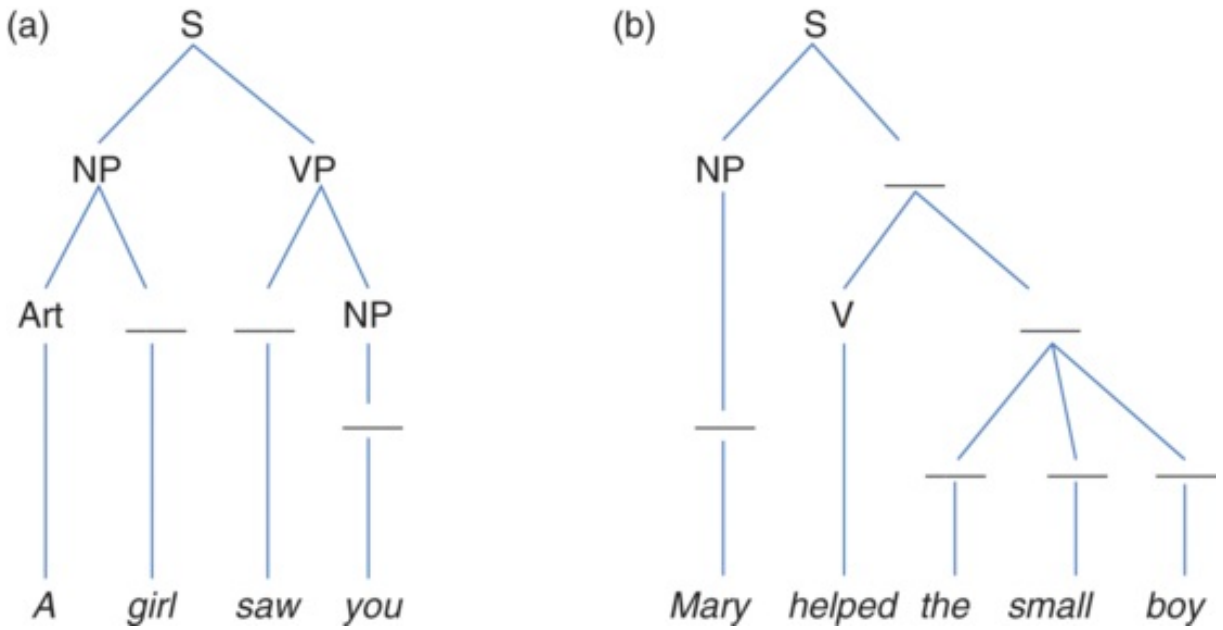


Figure 8.5

Tasks

A What is the distinction made between “competence” and “performance” in the study of syntax?

B What is meant by the expression “an embedded structure”? Were there any examples in this chapter?

C In some versions of syntactic analysis there are also “movement” rules that move parts of structures to different positions. For example, the statement *You can see it* becomes the question *Can you see it?* by moving one element (*can*) to the front. This element is an auxiliary (or “helping”) verb, as are *could, should, will, would*. They attach to verbs (*follow, help, see*) in the basic tree, as on the left in Figure 8.6, and are moved to the front to create a new tree, as on the right in Figure 8.6. In some descriptions, this change is called “inversion.”

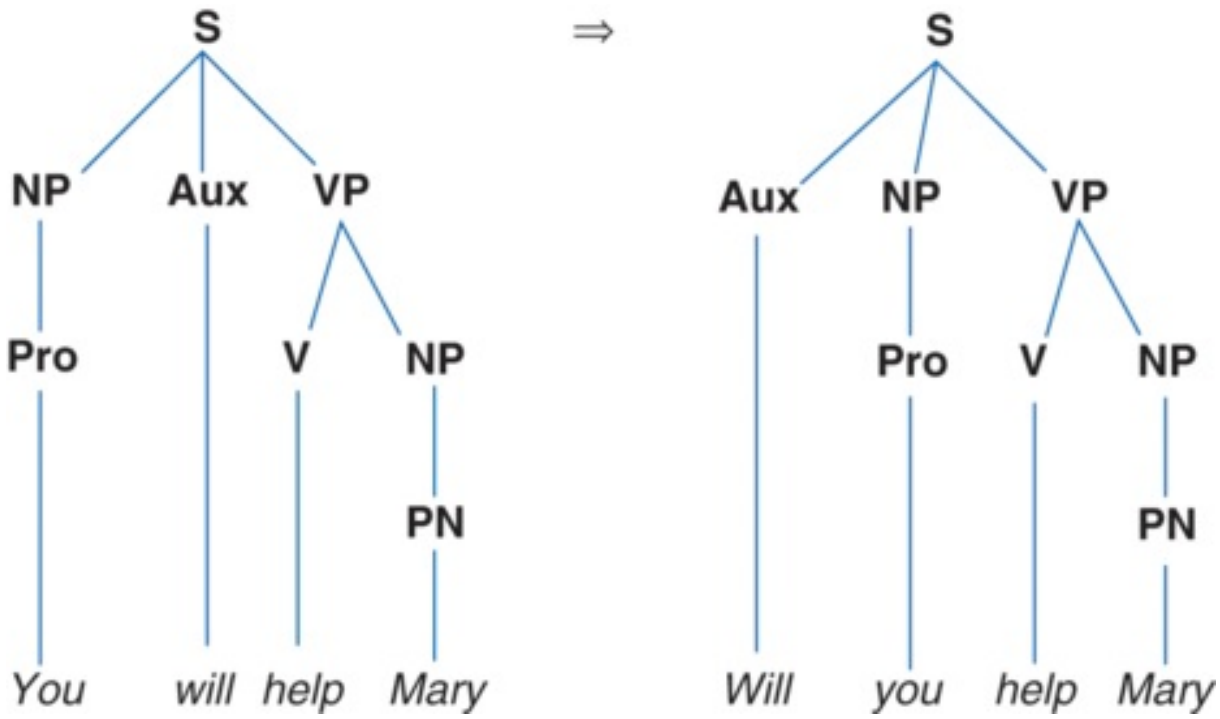


Figure 8.6

A special arrow (⇒) is used to indicate that a constituent can be moved, as shown in this rule for **Aux-movement**: NP Aux VP ⇒ Aux NP VP.

Which of these structures would result from applying the Aux-movement rule?

(1) *John will follow Mary.*

(2) *Can you see the dog?*

(3) *Could it follow you?*

(4) *The girl helped you.*

(5) *Could you help the dog?*

(6) *Mary should see it.*

(7) *Will the boy see you?*

(8) *Would John help the girl?*

(9) *It can see you.*

(10) *Can't you follow it?*

D In spoken English, the sequence *want to* is sometimes contracted to *wanna*, as in *I don't wanna go* or *What do you wanna do tonight?*. However, as illustrated in the following set of sentences, there are some structures where *want to* cannot be contracted. English-speaking children know how to use *wanna* in the right places (and none of the wrong places) at a very early age. Can you work out what it is that they know about using *wanna*?

(1) *Who do you* ***want to*** *or* ***wanna*** *visit?*

(2) *Who would you* ***want to*** *or* ***wanna*** *go out with?*

(3) *How many of your friends do you* ***want to*** *or* ***wanna*** *invite to the wedding?*

(4) *Who do you* ***want to*** *(*wanna) win the game?*

(5) *Who would you* ***want to*** *(*wanna) look after your pets?*

(6) *How many of your friends do you* ***want to*** *(*wanna) stay with us?*

E The following simplified set of phrase structure rules describes part of the syntax of a language called Ewe, spoken in West Africa. Based on these rules, which of the following sentences (1)–(10) should have an asterisk * before them?

S→ NP VP	N → {*oge, ika, amu*}
NP→ N (Art)	Art → *ye*
VP→ V NP	V → {*xa, vo*}

(1) *Oge xa ika*

(6) *Vo oge ika*

(2) *Ye amu vo oge*

(3) *Ika oge xa ye*

(4) *Oge ye vo ika ye*

(5) *Amu xa oge*

(7) *Amu ye vo ika*

(8) *Ye ika xa ye oge*

(9) *Xa amu ye*

(10) *Oge ye xa amu*

F Using these simple phrase structure rules for Scottish Gaelic, identify (with *) the ungrammatical sentences and draw tree diagrams for the grammatical sentences.

S→ V NP NP

Art→ *an*

N→ *cu, {duine, gille}*

PN→ {*Calum, Mairi, Tearlach*}

NP→ {Art N (Adj), PN}

Adj→ {*ban, beag, mor*}

V→ {*bhuail, chunnaic, fhuair*}

---

(1) *Calum chunnaic an gille.*

(2) *Bhuail an beag cu Tearlach.*

(3) *Bhuail an gille mor an cu.*

(4) *Chunnaic Tearlach an gille.*

(5) *Ban an cu an duine beag.*

(6) *Fhuair Mairi an cu ban.*

G The basic structure of a sentence in Tamasheq, spoken in north-west Africa, is illustrated as (1) in the sentences on page 120, but an emphasized element can be

moved to front position, as shown in the other examples. All these examples are from Sudlow (2001: 47), with minor changes.

(i) After looking at the syntactic structure of each Tamasheq sentence, can you add these English translations to appropriate places in the chart?

"It isn't men who cook porridge."
"Porridge, men aren't the ones who cook it."
"Men don't cook porridge?"
"Men aren't the ones who cook porridge."

(ii) Using information from Chapters 7 and 8, can you decide which of these languages has the same basic sentence structure as Tamasheq, as shown in example (1): English, Ewe, Gaelic, Japanese, Latin?

(1)

war	*səkədiwan*	*meddan*	*asink*	"Men don't cook porridge."
(not)	(cook)	(men)	(porridge)	

(2)

meddan a waren isəkədiw asink	_____

(3)

asink, meddan a waren t-isəkədiw	_____

(4)

wadde medan a isakadawan asink _____

(5)

meddan war səkədiwan asink? _____

H Which of the following two tree diagrams could be used to represent the underlying structure of the sentence: *George saw the boy with a telescope*?

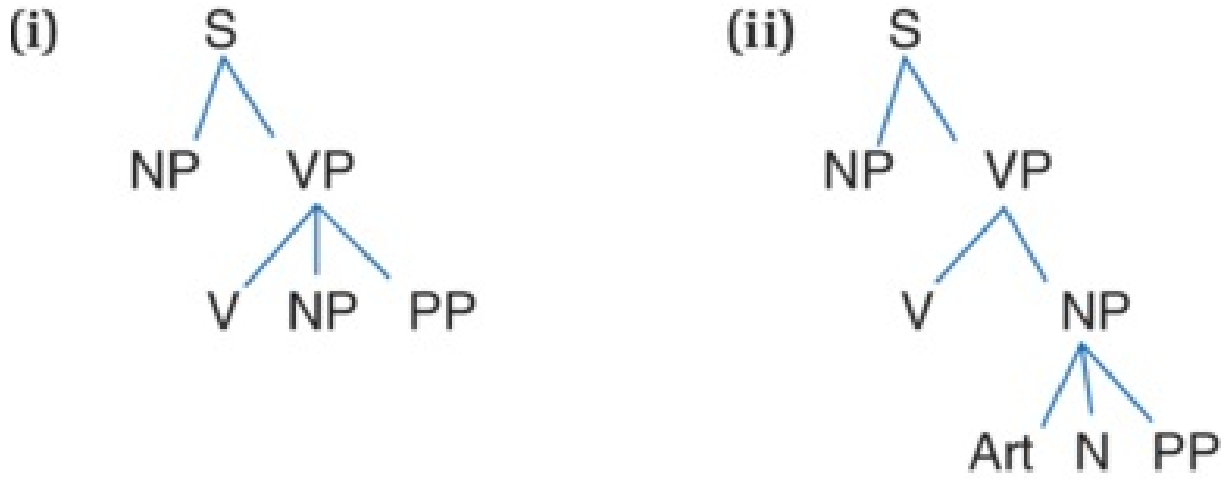


Figure 8.7

I The concept of recursion is used in syntax to describe the repeated application of a rule to the output of an earlier application of the rule. For example, we can use the terms "complementizer" (**C**) for the English word *that*, and "complement phrase" (**CP**) for *that Mary helped you* as part of the sentence *Cathy knew <u>that Mary helped you</u>*. In the complement phrase, the part *Mary helped you* represents a sentence (**S**), so there must be a rule: **CP → C S**, or "a complement phrase rewrites as a complement and a sentence."

This provides us with a small set of rules incorporating recursion. (Note that when you reach the end of this set of rules, you can keep going back to the beginning and repeating the sequence. That is the essence of recursion.)

S	→ NP VP
VP	→ V CP
CP	→ C S

Using these rules, fill in the missing elements in the tree diagram in Figure 8.8.

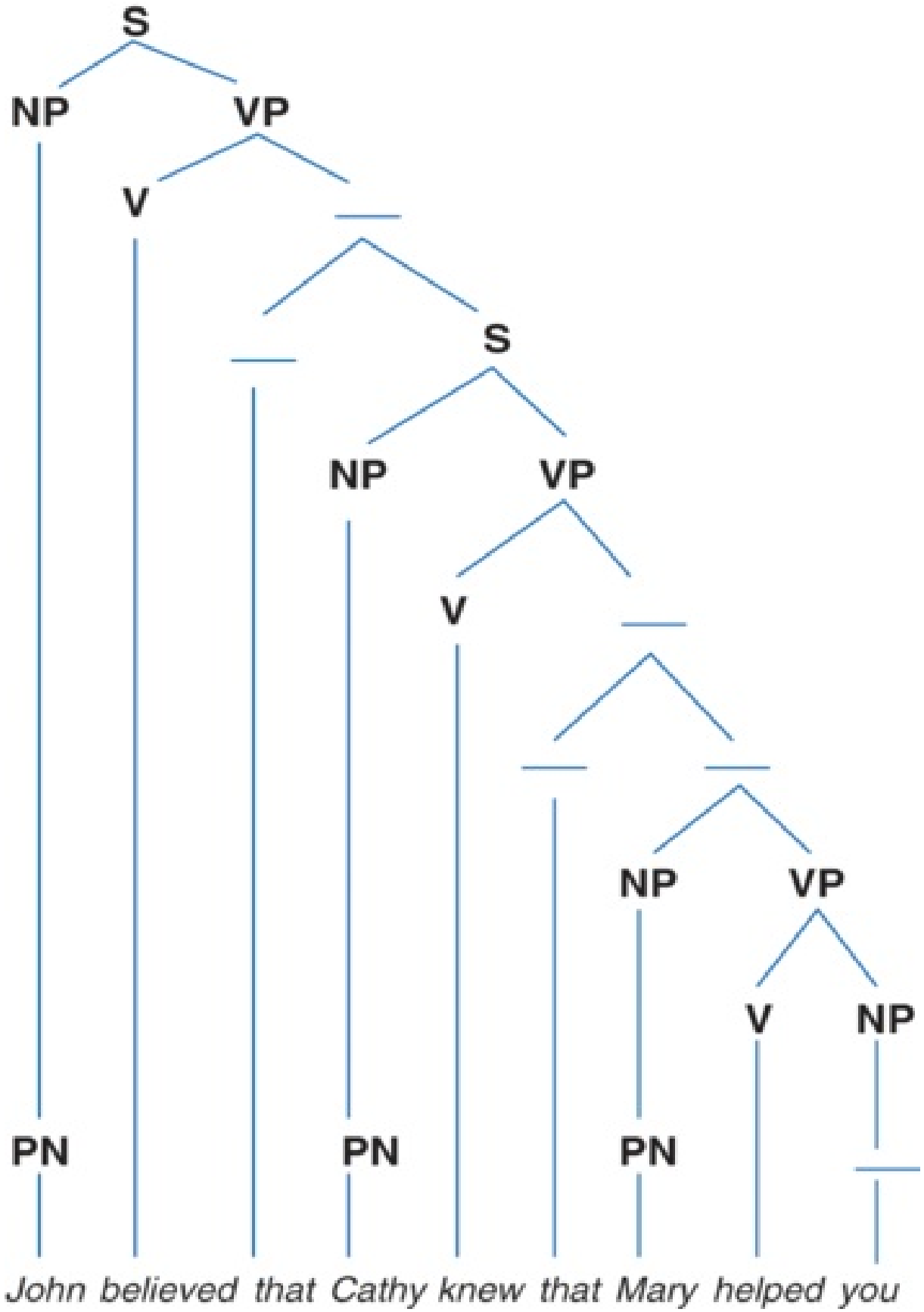

S
NP
VP
V
S
NP
VP
V
NP
VP
V
NP
PN
PN
PN
John believed that Cathy knew that Mary helped you

Figure 8.8

Discussion Topics/Projects

I There is a principle of syntax called "structure dependency" that is often used to show that the rules of language structure depend on hierarchical organization and not on linear position. For example, someone trying to learn English might be tempted to think that questions of the type in (2) are formed simply by moving the second word in a statement (1) to become the first word of a question (2).

---

(1) *Shaggy **is** tired.*	(2) ***Is** Shaggy tired?*
*You **will** help him.*	***Will** you help him?*

Using the sentences in (2)–(6), try to decide if this is the best way to describe how all of these English questions are formed and, if it is not, try to formulate a better rule.

(3) *Are the exercises in this book too easy?*

(4) *Is the cat that is missing called Blackie?*

(5) *Will the price of the new book you've ordered be really expensive?*

(6) *Was the guy who scored the winning goal in the final playing for love or money?*

(For background reading, see chapter 3 of Fromkin, Rodman and Hyams, 2014.)

II We could propose that passive sentences (*George was helped by Mary*) are derived from active structures (*Mary helped George*) via a rule such as the

following:

(active) NP_1 V NP_2 => NP_2 *be* V-*ed by* NP_1 (passive)

Note that the tense, past or present, of the V (e.g. *helped*) in the active structure determines the tense of *be* in the passive structure (e.g. ***was*** *helped*). Which of the following active sentences can be restructured into passive sentences using this rule? What prevents the rule from working in the other cases?

(1) *The dog chased the cat.*

(2) *Snow White kissed Grumpy.*

(3) *He loves them.*

(4) *Betsy borrowed some money from Christopher.*

(5) *The team played badly.*

(6) *The bank manager laughed.*

(7) *They have two children.*

(8) *The duckling became a swan.*

(9) *Someone mentioned that you played basketball.*

(10) *The police will arrest violent demonstrators.*

(For background reading, see Morenberg, 2013.)

Note: The different underlying structures in Oettinger's (1966: 168) example, *Time flies like an arrow; fruit flies like a banana*, cited at the beginning of this chapter, can be represented in the following tree diagrams. The different structures depend on some lexical ambiguity since *flies* is a verb in the first part and a noun in the second part. Also *like* is a preposition in the first part and a verb in the second part.

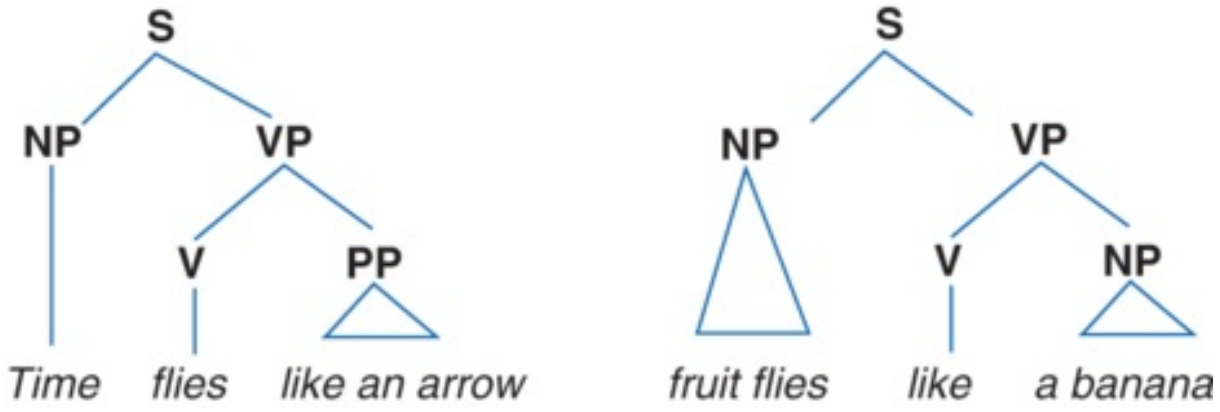


Figure 8.9

Further Reading

Basic Treatments

Miller, J. (2008) *An Introduction to English Syntax* (2nd edition) Edinburgh University Press

Thomas, L. (1993) *Beginning Syntax* Blackwell

More Detailed Treatments

Morenberg, M. (2013) *Doing Grammar* (5th edition) Oxford University Press

Tallerman, M. (2014) *Understanding Syntax* (4th edition) Routledge

Specifically on English Syntax

Jonz, J. (2013) *An Introduction to English Sentence Structure* Equinox Publishing

Radford, A. (2009) *An Introduction to English Sentence Structure* Cambridge University Press

On Generative Grammar

Baker, M. (2001) *The Atoms of Language: The Mind's Hidden Rules of Grammar* Basic Books

On Structural Ambiguity

Pinker, S. (1994) *The Language Instinct* (chapter 4) William Morrow

Tree Diagrams

Carnie, A. (2012) *Syntax* (3rd edition) Wiley-Blackwell

Other References

Fromkin, V., R. Rodman and N. Hyams (2014) *An Introduction to Language* (10th edition) Wadsworth

Sudlow, D. (2001) *The Tamasheq of North-East Burkina Faso* R. Köppe Verlag