



3

Syntax: Infinite Use of Finite Means

To grammar even kings bow.

J. B. MOLÈIRE, *Les Femmes Savantes*, II, 1672

It is a remarkable fact that any speaker of a human language can learn and store in his or her mental lexicon thousands of words, each of which is an arbitrary pairing of sound and meaning. Even more astonishing is our ability to combine these words to produce and understand an infinite number of novel sentences, as we showed with the following sentence:

This is the dog that worried the cat that killed the rat that ate the malt that lay in the house that Jack built . . .

To further illustrate, consider the following:

Snorlax is asleep.

The monster is asleep.

The friend of the monster is asleep.

The rightmost person in the first row is asleep.

The person immediately to the left of the rightmost person in the front row is asleep.

The person behind the person immediately to the left of the rightmost person in the first row is asleep.

Snorlax is asleep.

Pikachu noticed that Snorlax is asleep.

Nobody cares that Pikachu noticed that Snorlax is asleep.
 Squirtle knows that nobody cares that Pikachu noticed that Snorlax is asleep.

We can do this because we know (a finite number of) rules, which can be applied repeatedly. All spoken language is governed by rules—the set of rules is called a **grammar**. Every speaker has a mental grammar of the rules of his or her language that he or she follows in producing, understanding, and making judgments of well-formedness (grammaticality) about his or her language.

If we modify the order of words or omit some of the words, the sentences sound “weird” or “odd.” (Recall that the asterisk or star preceding a sentence is the linguistic convention indicating that the sentence is ungrammatical or ill-formed according to the rules of the grammar.)

- *Asleep is Homer.
- *Professor the is asleep.
- *Rightmost person the in the first row is asleep.
- *Homer asleep.
- *Right most person front row is asleep.

The oddness of these sentences indicates that some rule of the language has been violated. The sentences are ungrammatical.

To further illustrate this idea let’s look at a simple made-up rule of English that we’ll call the “everybody knows” rule:

Rule: If *S* is a sentence of English then *Everybody knows that S* is a sentence of English.

This rule can be iterated (repeated) any number of times to produce an arbitrary number of new sentences.

Snorlax is asleep.
 Everybody knows that Snorlax is asleep.
 Everybody knows that everybody knows that Snorlax is asleep.
 Everybody knows that everybody knows that everybody knows that Snorlax is asleep.

This simple rule in the mind of a speaker enables him or her to produce and understand a potentially infinite number of sentences. The “everybody knows” rule describes (generates) an infinite set of sentences. Any sentence that conforms to the rule is judged well-formed and any sentence that does not conform to the rule is judged ungrammatical, such as the following:

- *Knows everybody that Snorlax is asleep.

Given any sentence a speaker could create another sentence by adding a (nother) prepositional phrase, relative clause, or by embedding one sentence inside another as in the “everybody knows” examples. Or simply by adding another adjective:

The kindhearted boy had many girlfriends.
 The kindhearted, intelligent boy had many girlfriends.
 The kindhearted, intelligent, handsome boy had many girlfriends.

All languages have mechanisms of this sort that make the number of sentences limitless. Like words, discussed in the previous chapter, sentences are composed of finitely many discrete units that are combined by rules. Thus languages make infinite use of finite means. In this respect knowledge of language is like knowledge of integers. There is no limit to the number of even integers you could enumerate: 2, 4, 6, 8, 10, ... Clearly, you didn't memorize all of them. Rather, you know a rule that allows you to produce new integers from old ones.

Rule: If E is an integer, $E+2$ is an integer.

This ability to make infinite use of finite means shows the creative nature of human linguistic knowledge—not creative in the sense that we are all accomplished poets, but creative in that none of us is limited to a fixed repertoire of expressions. Rather, we can exploit the resources of our language and grammar to produce, understand and make judgments about a limitless number of sentences embodying a limitless range of ideas and emotions.

The part of grammar that represents a speaker's knowledge of sentences and their structures is called **syntax**. The aim of this chapter is to first show you what syntactic structures look like and then to familiarize you with some of the rules that determine them. Most of the examples will be from the syntax of English, but the principles that account for syntactic structures are universal.

What the Syntax Rules Do

"Then you should say what you mean," the March Hare went on.

"I do," Alice hastily replied, "at least—I mean what I say—that's the same thing, you know."

"Not the same thing a bit!" said the Hatter. "You might just as well say that 'I see what I eat' is the same thing as 'I eat what I see'!"

"You might just as well say," added the March Hare, "that 'I like what I get' is the same thing as 'I get what I like'!"

"You might just as well say," added the Dormouse . . . "that 'I breathe when I sleep' is the same thing as 'I sleep when I breathe'!"

"It is the same thing with you," said the Hatter.

LEWIS CARROLL, *Alice's Adventures in Wonderland*, 1865

The rules of syntax combine words into phrases and phrases into sentences. Among other things, the rules define the correct word order for a language. For example, English is a Subject–Verb–Object (SVO) language. The English sentence in (1) is grammatical because the words occur in the right order; the sentence in (2) is ungrammatical because the word order is incorrect for English.

1. The President nominated a new Supreme Court justice.
2. *President the Supreme new justice Court a nominated.

The rules of the syntax also specify the **grammatical relations** of a sentence, such as **subject** and **direct object**. In other words, they provide information

about who is doing what to whom. This information is crucial to understanding the meaning of a sentence. For example, the grammatical relations in (3) and (4) are reversed, so the otherwise identical sentences have very different meanings.

3. Your dog chased my cat.
4. My cat chased your dog.

The word order of a sentence is crucial to its meaning. The sentences in (5) and (6) contain the same words, but the meanings are quite different, as the Mad Hatter points out.

5. I mean what I say.
6. I say what I mean.

Although the structure of a sentence contributes to its meaning, as illustrated in the examples 3–6, grammaticality and meaningfulness are not the same thing. Consider the following sentences:

Colorless green ideas sleep furiously.
A verb crumpled the milk.

Although these sentences do not make much sense, they are syntactically well-formed. They sound funny, but their funniness is different from what we find in the following strings of words, which are not syntactically well-formed:

*Furiously sleep ideas green colorless.
*Milk the crumpled verb a.

There are also sentences that we understand even though they are not well-formed according to the rules of the syntax. We can easily interpret Yoda's words to Luke Skywalker although the word order is incorrect for English.

“... when gone I am ... the last of the Jedi will you be”

To be a sentence, words must conform to specific patterns determined by the specific syntactic rules of the language.

Some sentences are grammatical even though they are difficult to interpret because they include nonsense words, that is, words with no agreed-on meaning. This is illustrated by the following lines from the poem “Jabberwocky” by Lewis Carroll:

’Twas brillig, and the slithy toves
Did gyre and gimble in the wabe

These lines are grammatical in the linguistic sense that they obey the word order and other constraints of English. Such nonsense poetry is amusing precisely because the sentences comply with syntactic rules and sound like English. Ungrammatical strings of nonsense words are not entertaining:

*Toves slithy the and brillig ‘twas
wabe the in gimble and gyre did

Grammaticality does not depend on the truth of sentences. If it did, lying would be easy to detect. Nor does it depend on whether real objects are being discussed or whether something is possible in the real world. Untrue sentences

can be grammatical, sentences discussing unicorns can be grammatical, and sentences referring to pregnant fathers can be grammatical.

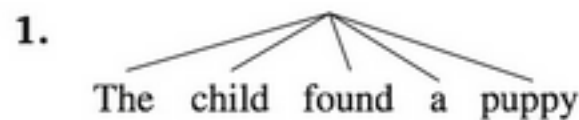
The ability to produce, understand, and judge the grammaticality of a sentence depends on whether it conforms to the unconscious rules of our mental grammar. This grammar is different from the prescriptive grammar rules that we are taught in school. We develop the mental rules of grammar long before we attend school, as we shall see in Chapter 9.

Sentence Structure

I really do not know that anything has ever been more exciting than diagramming sentences.

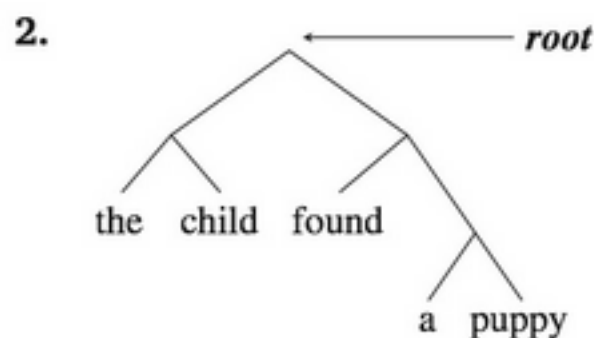
GERTRUDE STEIN, "Poetry and Grammar," 1935

The job of the linguist is to describe the structure of the sentences in a language in a way that matches the linguistic knowledge of its speakers. We can compare two competing hypotheses. The first says that a sentence consists simply of a string of words organized in a flat structure as in (1).



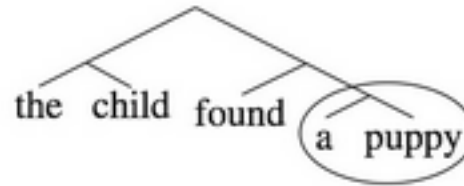
We have already seen that word order is an important aspect of syntactic knowledge and this simple diagram correctly captures the SVO word order of English: The subject (S) *the child*, comes before the verb (V) *found*, which comes before the object (O) *a puppy*.

Let us contrast this kind of description with another, one that says that sentences have a tree-like structure in which words are grouped together into natural units nested within other natural units in a hierarchical arrangement, as in (2).

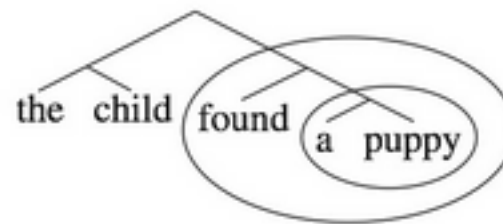


The "tree" in (2) is upside down with its "root" encompassing the entire sentence, "The child found a puppy," and its "leaves" being the individual words *the*, *child*, *found*, *a*, and *puppy*. The **tree diagram** in (2), embodies the hypothesis that these words are organized into subunits (or subtrees) and that speakers mentally represent sentences not as flat strings of words, but as complex structures with an internal organization. The subunits (or subtrees) of the sentence are called **constituents**.

In the tree diagram in (2), the words *a* and *puppy* form a constituent, as indicated below:

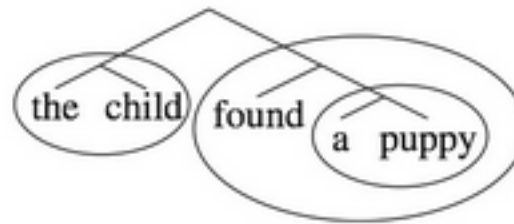


We can also represent constituents by using square brackets around the words [a puppy]. Constituents can be nested inside one another. So, [a puppy] occurs inside the constituent [found a puppy], as illustrated in the following tree.

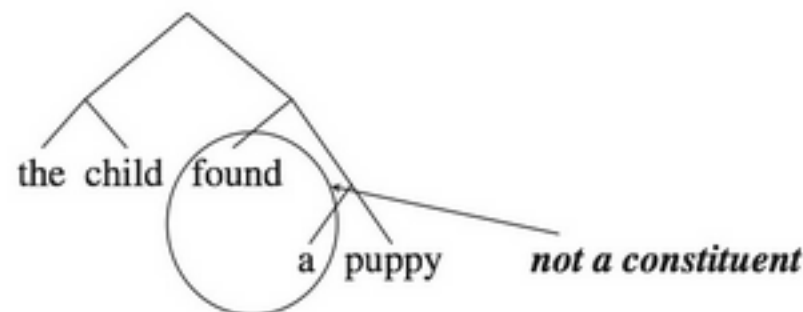


Using bracket notation, we would write this as [found [a puppy]].

There is one more constituent in the tree in (2). Do you know what it is? If you guessed [the child] you would be correct.



A constituent consists not just of the words, but of the subtree that branches into the words, and it ends at the **node** where the branches meet. A constituent corresponds to a node on the tree. And to be a constituent all the words under the node must be included. The words that form a constituent are contiguous (next to one another), but not all contiguous words form a constituent. In the following tree, the words *found a* are contiguous but they do not form a constituent. They are not contained exclusively under the same node.



We began our discussion with a simple sentence “The child found a puppy,” but this simple sentence belies a complex internal structure. The tree diagram in (2) groups the words of the sentence into the constituents *the child* and *found*

a puppy, corresponding to the subject and predicate of the sentence. A further division of the phrase *found a puppy* divides naturally into two branches, one for the verb *found* and the other for the direct object *a puppy*. This division conforms to our intuitions about the natural units of the sentence in a way that a different division, say, *found a* and *puppy*, would not.

Constituents and Constituency Tests

In addition to our intuitions of naturalness, various linguistic tests reveal the constituents of a sentence. The first test is the “stand alone” test. If a group of words can stand alone, for example, as an answer to a question, they form a constituent. So, in response to the question “What did the child find?” a speaker might answer *a puppy*, but not *found a*. *A puppy* can stand alone while *found a* cannot. We have a clear intuition that one of these is a meaningful unit and the other is just a list of words.

The second test is “replacement by a pronoun.” Pronouns can substitute for natural groups. In answer to the question, “Where did the child find *a puppy*?” a speaker can say, “I found *him* in the park.” Words such as *do* (which is not a pronoun per se) can also take the place of the entire predicate *found a puppy*, as in “The boy found a puppy and the girl *did* too.” If a group of words can be replaced by a pronoun or a word like *do*, it forms a constituent.

A third test of constituency is the “move-as-a-unit” test. If a group of words can be moved together and remain grammatical, they form a constituent. For example, if we compare the following sentences to the sentence “The child found a puppy,” we see that certain elements have moved:

It was *a puppy* that the child found.
A puppy was found by *the child*.

In the first example, the constituent *a puppy* has moved from its position following *found*; in the second example, the positions of *a puppy* and *the child* have been changed. In all such rearrangements, the constituents *a puppy* and *the child* remain intact. *Found a* does not remain intact, because it is not a constituent. Nor does *child found* for the same reason. Even though both these pairs of words occur next to each other in the original sentence *The child found a puppy*; they do not pass constituency tests, illustrating again that sentences are not simply string of words.

Some sentences have prepositional phrases in the predicate, for example:

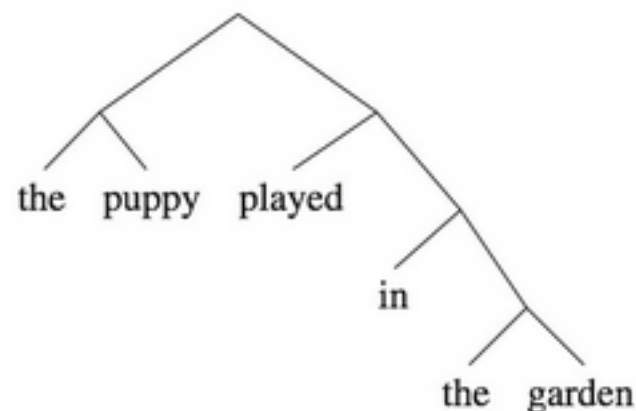
The puppy played in the garden.

We can use our tests to show that *in the garden* is also a constituent, as follows:

Where did the puppy play? *In the garden* (stand alone)
 The puppy played *there*. (replacement by a pronoun-like word)
In the garden the puppy played. (move as a unit)
 It was *in the garden* that the puppy played. (move as a unit)

The prepositional phrase in this example passes all three constituent tests. But in general a constituent need not pass all three tests. It is sufficient to pass one.

As before, our knowledge of the **constituent structure** of a sentence may be graphically represented by a tree diagram. The tree diagram for the sentence “The puppy played in the garden” is as follows:



The move-as-a-unit test can also tell us when what appears to be a constituent, such as a prepositional phrase, is in fact something different. The two phrases *ran up the hill* and *ran up the bill* are superficially quite similar, but we see in (3) and (4) that they behave quite differently. Consider first the expression *ran up the hill*, as in (3a). The rules of the syntax allow the word orders in (3b, c) as variants, revealing that *up the hill* is a constituent. By contrast, the expression *run up the bill* in (4a) does not have these same options, as shown in (4b, c), which means that *up the bill* is neither a prepositional phrase nor a constituent.

3. (a) Jack ran up the hill.
 (b) Up the hill Jack ran.
 (c) Up the hill ran Jack.
4. (a) Jack ran up the bill.
 (b) *Up the bill Jack ran.
 (c) *Up the bill ran Jack.

Structural Ambiguity



Hilary B. Price/King Features Syndicate

Syntactic trees reflect our judgments about the internal organization of sentences; flat structures do not. They can also account for other linguistic judgments, such as when a sentence is **ambiguous**. A sentence is ambiguous if it

has two or more meanings. Sometimes an ambiguity arises because a word has more than one meaning, as in the following sentence:

This will make you smart.

The two interpretations of this sentence are due to the two meanings of *smart*—“clever” and “burning sensation.” This is referred to as a **lexical ambiguity** and will be discussed further in Chapter 4. Other times multiple meanings arise because a sentence has more than one tree structure associated with it, resulting in a **structural ambiguity**. Each tree will correspond to one of the possible meanings of the sentence. For example, the sentence:

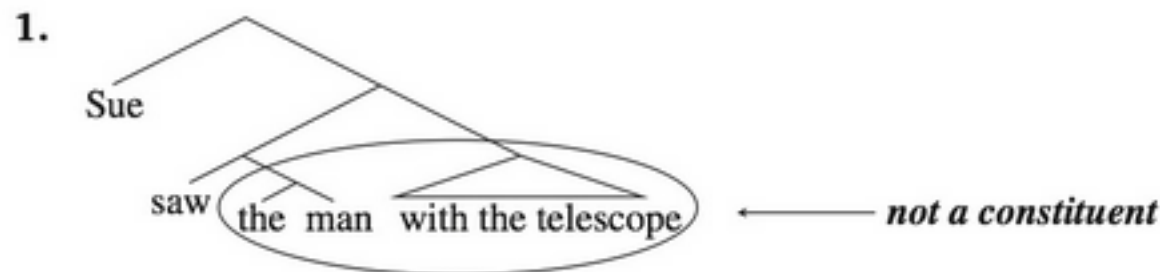
Sue saw the man with the telescope.

has two different meanings:

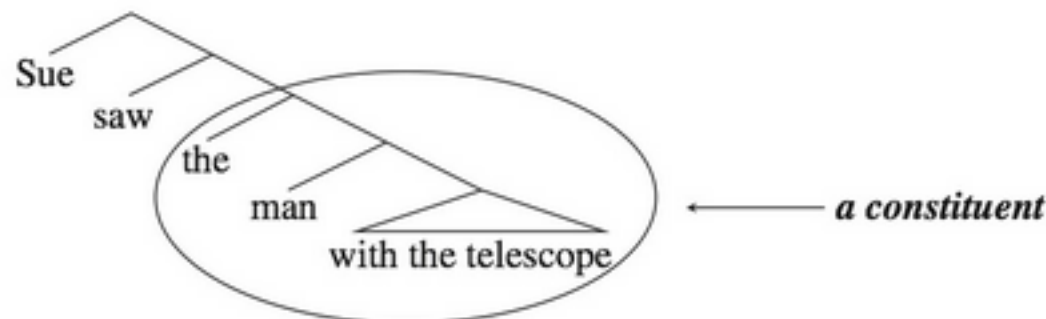
Meaning 1: The seeing is done with the telescope.

Meaning 2: The man is holding the telescope.

Notice that none of the individual words is ambiguous. The ambiguity is structural: The sentence has two different trees. Meaning 1 corresponds to the tree in (1), what we might call the instrumental meaning in which Sue is using the telescope to see the man. In this tree, the *the man* and the prepositional phrase *with the telescope* do not form a constituent.



Meaning 2 corresponds to the tree in (2). In this case, the phrases *the man* and *with a telescope* do form a constituent, reflecting the meaning in which the man is holding the telescope.



The availability of these two structures leads to a prediction: If we do a constituency test that forces *the man with the telescope* to be a constituent (e.g. move it as a unit), then meaning 1 (instrumental) should disappear and we should only have the meaning corresponding to the tree in (2). The following sentences confirm this prediction:

It was the man with the telescope that Mary saw.

The man with the telescope was seen by Mary.

What Mary saw was the man with the telescope.

None of these sentences has the interpretation in which the seeing is done with the telescope. In each, the only possible meaning is that the man is holding the telescope. This shows us that the structure of a sentence contributes importantly to its meaning, a point we will come back to in Chapter 4.

Structural ambiguities of the sort just discussed provide striking evidence in support of our hypothesis that sentences have a tree-like (hierarchical) structure, and against the idea that they are simply strings of words. The flat structure hypothesis could not explain how there can be two different meanings associated with *the same string of words*.

The cartoon at the head of this section illustrates both lexical and structural ambiguity. The lexical ambiguity is on the two meanings of *ring*; the structural ambiguity is whether *nose ring* is understood as a compound noun or a noun followed by a verb.

Syntactic Categories

There are ten parts of speech, and they are all troublesome.

MARK TWAIN, "The Awful German Language," in *A Tramp Abroad*, 1880

In the previous section, we illustrated how tree structures reflect our knowledge of the hierarchical organization of sentences. Speakers also have implicit knowledge of the categories of each of the subgroupings in a sentence.

Each grouping in the tree diagrams of "The child found a puppy" is a member of a large family of similar expressions. For example, *the child* belongs to a family that includes *the police officer*, *your neighbor*, *this yellow cat*, *he*, *John*, and countless others. We can substitute any member of this family for *the child* without affecting the grammaticality of the sentence, although the meaning of course would change.

A police officer found a puppy.
Your neighbor found a puppy.
This yellow cat found a puppy.

A family of expressions that can substitute for one another without loss of grammaticality is called a **syntactic category**, or more informally, a "part of speech." *The child*, *a police officer*, *John*, and so on belong to the syntactic category **noun phrase (NP)**. NPs may function as subjects or as objects in sentences. An NP often contains a *determiner* (such as *a* or *the*) and a noun, but it may also consist of a proper name (*Ann*), a pronoun (*I*), a noun without a determiner (*fish*), or even a clause or a sentence (*that dogs bark*). Even though a proper noun such as *John* and pronouns such as *he* and *him* are single words, they are technically NPs, because they pattern like NPs in being able to fill a subject, object or other NP slot.

John found the puppy.
He found the puppy.
Boys love puppies.
The puppy loved him.
The puppy loved John.

NPs can be quite complex, as illustrated by the sentence:

The girl that Professor Snape loved married the man of her dreams.

The NP subject of this sentence is *the girl that Professor Snape loved*, and the NP object is *the man of her dreams*. We know this because each of these lengthy expressions fills a slot otherwise occupied by a simpler NP as in *Mary loved John*.

Syntactic categories are part of a speaker's knowledge of syntax. That is, speakers of English know that only items (a), (b), (e), (f), and (g) in the following list are NPs even if they have never heard the term *noun phrase* before.

1. (a) a bird
 (b) the red banjo
 (c) have a nice day
 (d) with a balloon
 (e) the woman who was laughing
 (f) it
 (g) John
 (h) went

You can test this claim by inserting each expression into three contexts: *What/who I heard was _____*, *Who found _____?* and *_____ was seen by everyone*. For example, **Who found with a balloon?* is ungrammatical, as is **Went was seen by everyone*, as opposed to *Who found it?* or *John was seen by everyone*. Only NPs fit into these contexts because only NPs can function as subjects and objects.

There are other syntactic categories. The expression *found a puppy* is a **verb phrase (VP)**. A verb phrase always contains a **verb (V)**, and it may contain other categories, such as a noun phrase or **prepositional phrase (PP)**, which is a preposition followed by an NP, such as *in the park*, *on the roof*, and *with a balloon*. In (2) the VPs are those phrases that can complete the sentence "The child _____."

2. (a) saw a clown
 (b) a bird
 (c) slept
 (d) smart
 (e) ate the cake
 (f) found the cake in the cupboard
 (g) realized that the Earth was round

Inserting (a), (c), (e), (f), and (g) will produce grammatical sentences, whereas the insertion of (b) or (d) would result in an ungrammatical sentence. Thus, (a), (c), (e), (f), and (g) are verb phrases.

Lexical and Functional Categories



"Very traditional. He's the noun. She's the adjective."

© The New Yorker Collection 2003 William Haefeli from cartoonbank.com All Rights Reserved.

Syntactic categories include both phrasal categories such as NP, VP, AP (adjective phrase), PP (prepositional phrase), and AdvP (adverbial phrase), as well as lexical categories such as noun (N), verb (V), preposition (P), adjective (A), and adverb (Adv). Each lexical category has a corresponding phrasal category. Following is a list of phrasal categories and lexical categories with some examples of each type:

Phrasal categories

Noun Phrase (NP)	<i>men, the man, the man with a telescope sees, always</i>
Verb Phrase (VP)	<i>sees, rarely sees the man, often sees the man with a telescope</i>
Adjective Phrase (AP)	<i>happy, very happy, very happy about winning</i>
Prepositional Phrase (PP)	<i>over, nearly over, nearly over the hill</i>
Adverbial Phrase (AdvP)	<i>brightly, more brightly, more brightly than the Sun</i>

Lexical categories

Noun (N)	<i>puppy, boy, man, soup, happiness, fork, kiss, pillow</i>
Verb (V)	<i>find, run, sleep, throw, realize, see, try, want, believe</i>
Preposition (P)	<i>up, down, across, into, from, by, with, over</i>
Adjective (A)	<i>red, big, happy, candid, hopeless, fair, idiotic, lucky</i>
Adverb (Adv)	<i>again, always, brightly, often, never, very, fairly</i>

Many of these categories may already be familiar to you. Other categories may be less familiar such as the category **determiner (Det)**, which includes the articles *a* and *the*, as well as **demonstratives** such as *this*, *that*, *these*, and *those*, and "quantifiers" such as *each* and *every*. Another less familiar category

Our knowledge of syntactic categories is also revealed through our intuitions about nonsensical sentences. Recall the sentences in (1) and (2) below. Although neither of these sentences makes sense, we have a clear intuition that (1) is grammatical in a way that (2) is not.

1. Colorless green ideas sleep furiously.
2. *Sleep colorless green furiously ideas.

This is because sentence (1) obeys the word order constraints of English while sentence (2) does not. In other words, we recognize the category of each of the words: *Colorless* is an adjective, *ideas* is a noun, *colorless green ideas* is a noun phrase, and *sleep furiously* is a verb phrase, and know that they fit properly together in (1) but not in (2). We are not taught these categories nor their word order. We know this implicitly before we go to school. They are part of our grammar that we develop as a child growing up (see Chapter 9).

In these sentences, we can identify when the order is correct and when it is not, even though the meanings of the different words and constituents do not jibe. An idea cannot be green or colorless, (except in a metaphorical sense), but even if ideas had color we would say *colorless green ideas* and not *ideas green colorless* or *green ideas colorless*.

Similarly, we may not be able to make sense of Lewis Carroll's Jabberwocky, but we can identify the words *brillig*, *slithy*, and *mimsy* as adjectives, *toves*, *wabe*, *borogoves*, and *momeraths* as nouns, and *outgrabe*, *gyre*, and *gimble* as verbs, all based on their position in the sentences.

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogoves,
And the momeraths outgrabe.

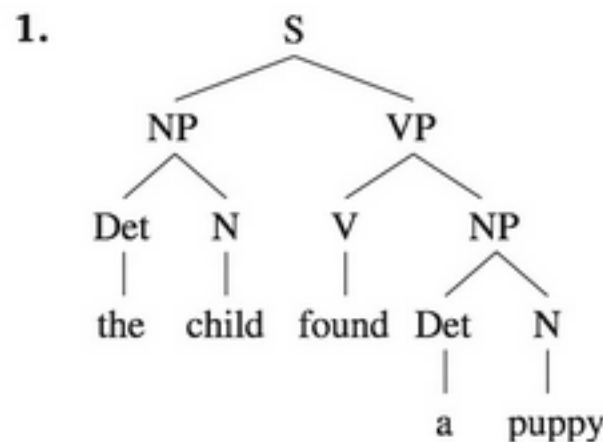
Speakers know the syntactic category of the various constituents and how they are ordered with respect to one another. They also know how to group words into units—*constituents*. This knowledge is graphically represented in tree structures that reveal the grammatical organization of the words of a sentence. Tree structures also explain how the grouping of words in a sentence relates to its meaning, such as when a sentence or phrase is ambiguous. And even when the meaning is nonsensical the structure must obey the syntactic rules of the language. The rules of syntax also permit speakers to produce and understand a limitless number of sentences never produced or heard before—the *creative aspect of linguistic knowledge*, illustrated at the beginning of this chapter. A major goal of linguistics is to show clearly and explicitly how syntactic rules account for what speakers implicitly know about their language.

Phrase Structure Trees

Who climbs the Grammar-Tree distinctly knows
Where Noun and Verb and Participle grows.
JOHN DRYDEN, "The Sixth Satyr of Juvenal," 1693

Now that you know something about constituent structure and grammatical categories, you are ready to learn how the phrases and sentences of a language are constructed. We will begin by illustrating trees for simple phrases and then proceed to more complex structures. The trees that we will build here are more detailed than those we saw in the previous sections, because the branches of the tree will have category labels identifying each constituent. In this section, we will also introduce the kind of syntactic rules that **generate** (a technical term for describe or specify) the different structures.

The tree diagram in (1) provides labels for each of the constituents of the sentence “The child found a puppy.” These labels show that the entire sentence belongs to the syntactic category of S (because the S-node encompasses all the words). It also reveals that *the child* and *a puppy* belong to the category NP, that is, they are noun phrases, and that *found a puppy* belongs to the category VP or is a verb phrase, consisting of a verb and an NP. It also shows the syntactic category of each of the words in the sentence.



A tree diagram with syntactic category information is called a **phrase structure tree** or a **constituent structure tree**. Phrase Structure trees (PS trees) represent three aspects of a speaker’s syntactic knowledge:

1. The linear order of the words in the sentence
2. The identification of the syntactic categories of words and groups of words
3. The hierarchical structure of the syntactic categories (e.g., an S is composed of an NP followed by a VP, a VP is composed of a V that may be followed by an NP, and so on).

The syntactic category of each word is listed in our mental dictionaries, as we will discuss in more detail in Chapter 4. This lexical information guides the syntax of the language. Words appear in trees under labels that correspond to their syntactic category. Nouns are under N, determiners under Det, verbs under V, and so on.

The larger syntactic categories such as VP consist of all the syntactic categories and words below that node in the tree. The VP in the PS tree above consists of syntactic category nodes V and NP and the words *found*, *a*, and *puppy*. Because *a puppy* can be traced up the tree to the node NP, this constituent is a noun phrase. Because *found* and *a puppy* can be traced up to the node VP, this constituent is a verb phrase. In discussing trees, every higher node is said to **dominate** all the categories beneath it. S dominates every node. A node is said to **immediately dominate** the categories one level below it. VP immediately dominates V and NP, the categories of which it is composed. Categories that

are immediately dominated by the same node are **sisters**. V and NP are sisters in the phrase structure tree of “the child found a puppy.”

PS trees are also useful for defining various grammatical relations in a precise way. For example, the subject of a sentence is the NP immediately dominated by S (*the child* in the tree in (1)) and the direct object is the NP immediately dominated by VP (*the puppy* in the tree in (1)).

Phrase Structure Rules

The information shown in a PS tree can also be represented by another formal device: phrase structure (PS) rules. PS rules capture the knowledge that speakers have about the possible structures of a language. Just as a speaker cannot have an infinite list of sentences in his or her head, so he or she cannot have an infinite set of PS trees in his or her head. Rather, a speaker’s knowledge of the permissible and impermissible structures must exist as a finite set of rules that characterize a tree for any sentence in the language. To express the structure given above, we need the following PS rules:

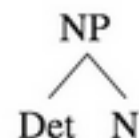
1. $S \rightarrow NP VP$
2. $NP \rightarrow Det N$
3. $VP \rightarrow V NP$

You can think of PS rules as templates that a tree must match to be grammatical. They express the regularities of the language and make explicit a speaker’s knowledge of the order of words and the grouping of words into syntactic categories. For example in English an NP may contain a determiner followed by a noun. This is represented by rule 2. This rule conveys two facts:

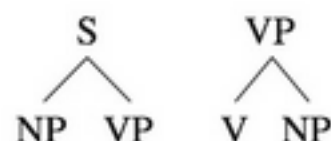
A noun phrase may contain a determiner followed by a noun in that order.

A determiner followed by a noun is a noun phrase.

Phrase structure rules specify the well-formed structures of a language precisely and concisely. To the left of the arrow is the dominating category NP. The categories that it immediately dominates appear on the right side, in this case Det and N. The right side of the arrow also shows the linear order of these components. Thus, the subtree for the English NP looks like this:



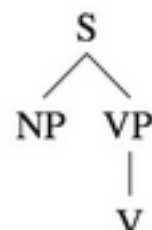
Rule 1 says that a sentence (S) contains (immediately dominates) an NP and a VP in that order. Rule 3 says that a verb phrase consists of a verb (V) followed by an NP. These rules are general statements and do not refer to any specific VP, V, or NP. The subtrees represented by rules 1 and 3 are as follows:



A VP need not contain an NP object, however. It may include a verb alone, as in the following sentences:

The woman laughed.
The man danced.
The horse galloped.

These sentences have the structure:



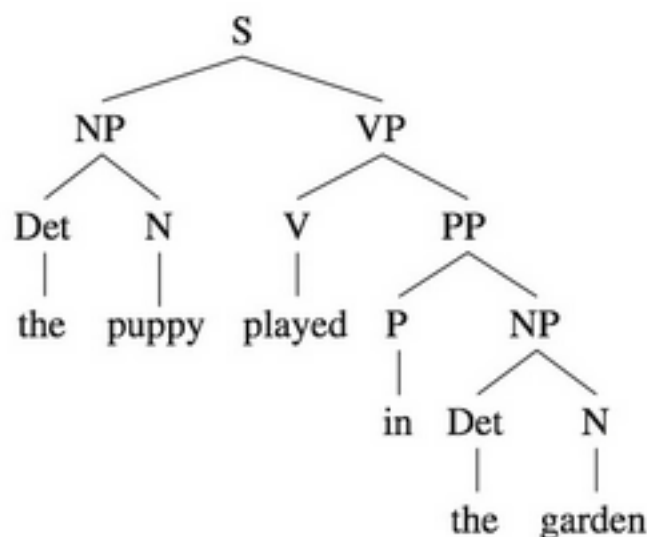
Thus, a tree may have a VP that immediately dominates only V, as specified by rule 4, which we include in our grammar:

4. $VP \rightarrow V$

The following sentences contain prepositional phrases following the Verb:

The puppy played in the garden.
The boat sailed up the river.
A girl laughed at the monkey.
The sheepdog rolled in the mud.

The PS tree for such sentences is



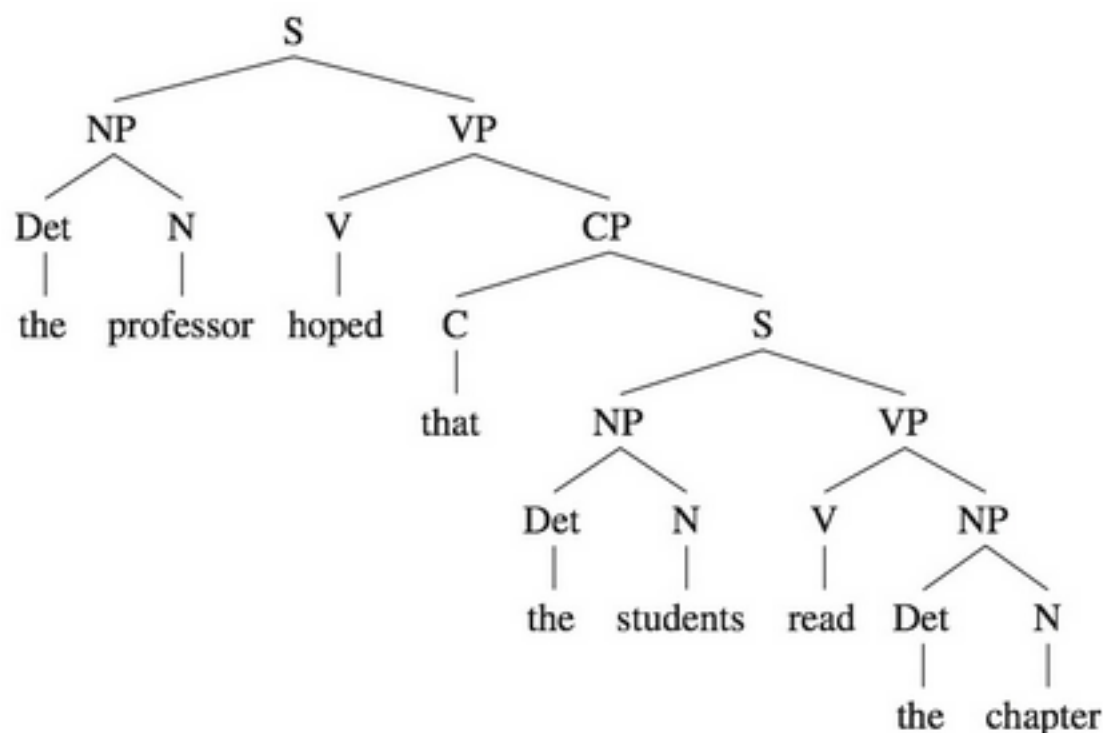
To generate structures of this type we need two additional PS rules as in 5 and 6.

5. $VP \rightarrow V PP$

6. $PP \rightarrow P NP$

Another option open to the VP is to contain or *embed* a sentence. For example, the sentence “The professor hoped that the students read the chapter” contains the sentence “the students read the chapter.” Preceding the **embedded sentence** is the word *that*, which belongs to the category of complementizers

C(omp), a functional category like T(ense) and Det. Here is the structure of such sentence types:



To allow such embedded sentences, we need to add these two new rules to our set of phrase structure rules.

7. $VP \rightarrow V CP$

8. $CP \rightarrow C S$

CP stands for complementizer phrase. Rule 8 says that CP contains a complementizer such as *that* followed by the embedded sentence. Other complementizers are *if* and *whether* in sentences such as

I don't know whether I should talk about this.

The teacher asked if the students understood the syntax lesson.

which have structures similar to the one above.

Here are the PS rules we have discussed so far. The rules have been slightly renumbered to keep all the VP rules together. We will introduce some other rules later.

1. $S \rightarrow NP VP$
2. $NP \rightarrow Det N$
3. $VP \rightarrow V NP$
4. $VP \rightarrow V$
5. $VP \rightarrow V PP$
6. $VP \rightarrow V CP$
7. $PP \rightarrow P NP$
8. $CP \rightarrow C S$

Building Phrase Structure Trees

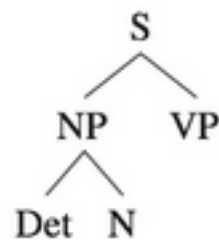
Everyone who is master of the language he speaks . . . may form new . . . phrases, provided they coincide with the genius of the language.

JOHANN DAVID MICHAELIS, "Dissertation," 1739

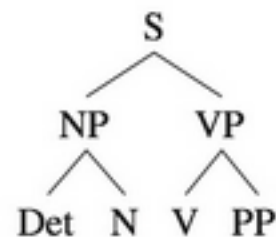
The phrase structure rules can be used as a guide for building trees that follow the structural constraints of the language. In so doing, certain conventions are followed. The S occurs at the top or "root" of the tree (remember the tree is upside down). So, first find the rule with S on the left side of the arrow (rule 1) and put the categories on the right side below the S, as shown here



Continue by matching any syntactic category at the bottom of the partially constructed tree to a category on the left side of a rule, then expand the tree with the categories on the right side. For example, we may expand the tree by applying the NP rule to produce:

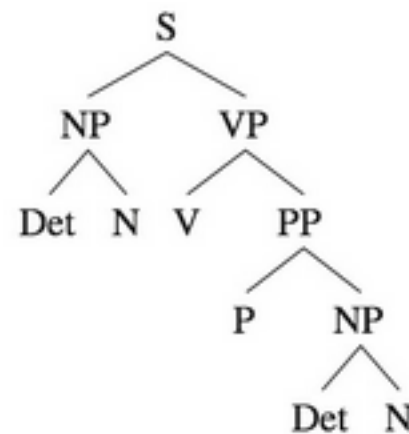


The categories at the bottom are Det, N, and VP, but only VP occurs to the left of an arrow in the set of rules and so needs to be expanded using one of the VP rules. Any one of the rules will work. The order in which the rules appear in the list of rules is irrelevant. (We could have begun by expanding the VP rather than the NP.) Suppose we use rule 4 next. Then, the tree has grown to look like this:



We continue in this way until all phrasal categories are expanded, that is, none of the categories at the bottom of the tree appears on the left side of any rule. The PP must expand into a P and an NP (rule 7), and the NP into a Det and an N. (Proper names and pronouns which are NPs and not nouns are an

exception to the “full expansion convention.”) We can use a rule as many times as it can apply. In this tree, we used the NP rule twice. After we have applied all the rules that can apply, the tree looks like this:



By following these conventions, we generate only trees specified by the PS rules, and hence only trees that conform to the syntax of the language. By implication, any tree not so specified will be ungrammatical, that is, not permitted by the syntax. At any point during the construction of a tree, any rule may be used as long as its left-side category occurs somewhere at the bottom of the tree. By choosing different VP rules, we could specify different structures corresponding to sentences such as:

- The boys left. (VP $\rightarrow$ V)
- The wind blew the kite. (VP $\rightarrow$ V NP)
- The senator hopes that the bill passes. (VP $\rightarrow$ V CP)

Because the number of possible sentences in a language is infinite, there are also an infinite number of trees. However, all trees are built out of a finite set of phrase structure rules.

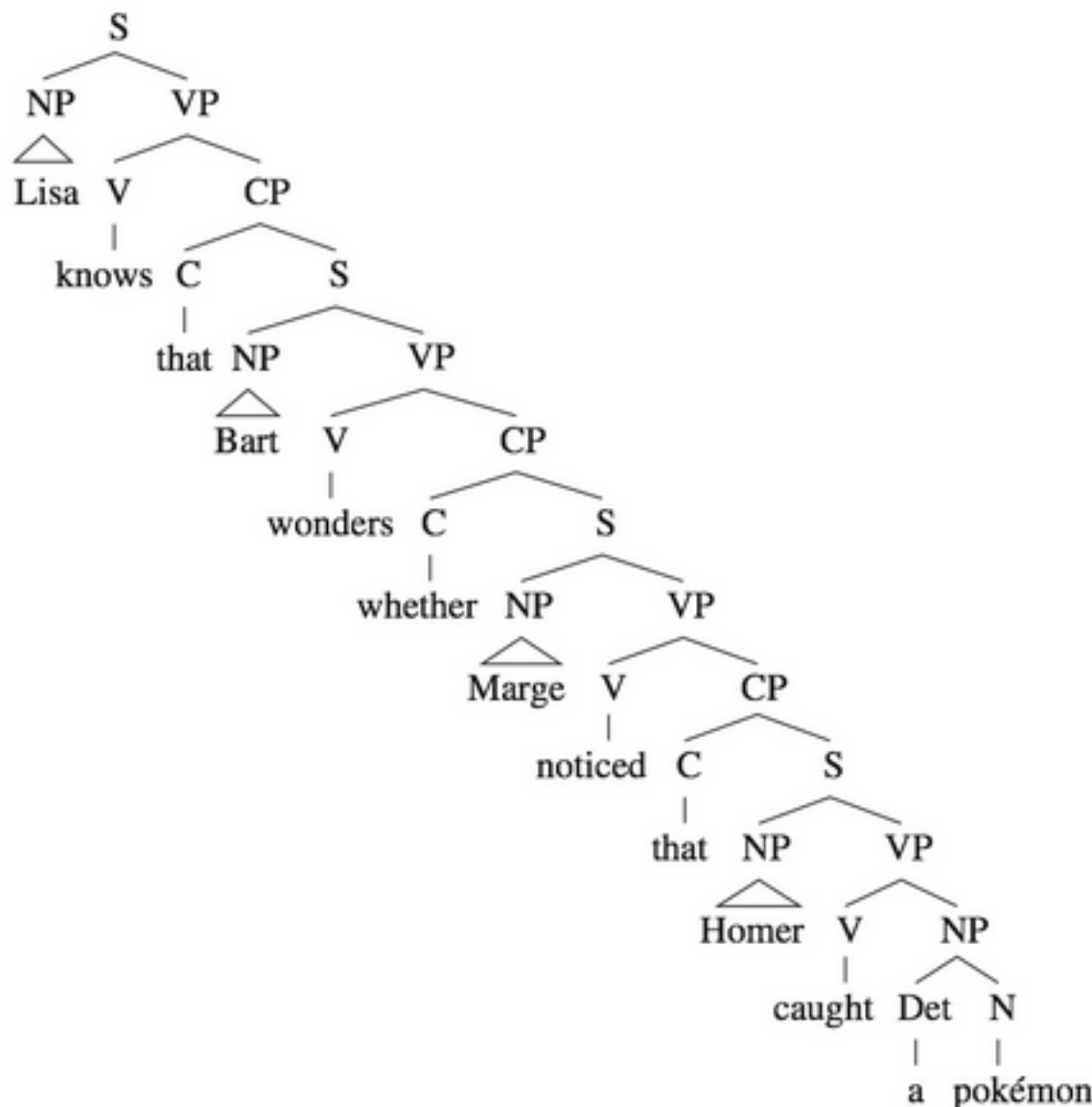
The Infinity of Language: Recursive Rules

Though incomplete, the set of PS rules we have introduced thus far is sufficient to illustrate the mechanisms by which languages generate a limitless number of sentences. Consider the following set of sentences, similar to those discussed at the beginning of this chapter.

1. Homer caught a pokémon.
2. Marge noticed that Homer caught a pokémon.
3. Bart wonders whether Marge noticed that Homer caught a pokémon.
4. Lisa knows that Bart wonders whether Marge noticed that Homer caught a pokémon.

We see that sentence 1 is embedded inside sentence 2, sentence 2 inside sentence 3, sentence 3 inside sentence 4. We could continue this process indefinitely. This is made possible by the fact that rule phrase structure rule 6 (VP $\rightarrow$ V CP) in combination with rules 8 (CP $\rightarrow$ C S) and 1 (S $\rightarrow$ NP VP) form a **recursive** set, in which the symbols S and VP occur on both the left and right side of the rules. Therefore, the rules allow S to contain VP, which in turn contains CP, which in

turn contains S, which in turn again contains VP and so on, potentially without end. Recursive rules are of critical importance because they allow the grammar to generate an infinite set of sentences. The PS tree for sentence 4 illustrates the application of these rules (here we use triangles under the NPs to indicate that proper names are full NPs, not nouns):



The property of recursion also illustrates the difference between competence and performance, discussed in Chapter 1. All speakers of English (and all other languages) have as part of their linguistic competence—their mental grammars—the ability to embed phrases within each other ad infinitum. However, as the structures grow longer, they become increasingly more difficult to produce and understand. This can be due to short-term memory limitations, muscular fatigue, breathlessness, boredom, or any number of performance factors. (We will discuss performance factors more fully in Chapter 10.) Nevertheless, these very long sentences would be well-formed according to the rules of the grammar.

Below, we will see other examples of recursive rule sets such as the one responsible for the potentially infinite number of prepositional phrases in sentences like:

The person behind the person immediately to the left of the rightmost person in the first row is asleep.

The Internal Structure of Phrases

I really do think that science has an internal structure, and it makes sense, and we can test it.

LISA RANDALL, Theoretical Physicist

In the previous sections, we focused on the hierarchical organization of sentences into phrasal categories such as NP, VP, and PP. In this section, we will look at the internal structure of phrases themselves.

Heads, Complements, and Selection

One of the striking things we observe when we consider the various phrase structure rules given above (and the subtrees they generate) is that they have a similar organization. Consider the following examples of each of the phrasal categories we have discussed:

NP: the *mother* of James Whistler

VP: *sing* an aria

PP: *over* the hill

For completeness, we add the category AP (adjective phrase), illustrated by the example

AP: *wary* of snakes

generated by the following rule:

9. AP → A PP

As we noted in our discussion of grammatical categories, the core of every phrase is a lexical category of its same syntactic type (italicized), which is its **head**; for example, the NP *the mother of James Whistler* is headed by the noun *mother*; the VP *sing an aria* is headed by the verb *sing*; the AP *wary of snakes* is headed by the adjective *wary*; the PP *over the hill* is headed by the preposition *over*. Loosely speaking, the entire phrase refers to whatever the head refers to. For example, the VP *sing an aria* refers to a “singing” event; the NP *the mother of James Whistler* to someone’s mother.

In addition to the head, the phrasal categories may contain other categories such as NP, PP or CP. These sister categories are called **complements**. A complement is a phrasal category that occurs next to a head, and only there, and which elaborates on the meaning of the head. The complements are underlined: For example, the head N *mother* takes the PP complement of James Whistler; the head V *sing* takes the NP complement an aria; the head A(djective) *wary* takes the PP of snakes, and the P(reposition) *over* takes the NP the hill as complement.

Selection

Complements are not always present in the phrase structure. They are optional; only the head is obligatory. The choice of complement type for any particular phrase depends on the specific properties of the head of that phrase. For

example, verbs select different kinds of complements: *find* is a transitive verb and requires an NP complement (direct object), as in *The boy found the ball*, but not **The boy found*, or **The boy found in the house*. Some verbs such as *eat* are optionally transitive. *John ate* and *John ate a sandwich* are both grammatical. *Sleep* is an **intransitive verb**; it cannot take an NP complement:

Michael slept.

*Michael slept the baby.

Some verbs, such as *think*, may select both a PP and a sentence complement (underlined):

Let's think about it.

I think a girl won the race.

Other verbs, such as *tell*, select an NP and a sentence:

I told the boy a girl won the race.

Yet other verbs such as *feel* select either an AP or a sentence complement:

Paul felt strong as an ox.

He feels he can win.

Categories besides verbs also select their complements. For example, the noun *belief* selects either a PP or a sentence, while the noun *sympathy* selects a PP, but not a sentence, as shown by the following examples:

the belief in freedom of speech

the belief that freedom of speech is a basic right

their sympathy for the victims

*their sympathy that the victims are so poor

Adjectives can also have complements. For example, the adjectives *tired* and *proud* select PPs:

tired of stale sandwiches

proud of her children

The information about the complement types selected by particular verbs and other lexical items is called **C-selection** or **subcategorization**, and is included in the lexical entries of the items in our mental lexicons. (C stands for “category.”)

A verb also includes in its lexical entry a specification that imposes certain semantic requirements its subjects and complements, just as it selects for syntactic categories. This kind of selection is called **S-selection**. (S stands for “semantic.”) For example, the verb *murder* requires its subject and object to be animate, while the verb *quaff* requires its subject to be animate and its object liquid. Verbs such as *like*, and *hate* select animate subjects. The following sentences violate S-selection and can only be used in a metaphorical sense. (We will use the symbol “!” to indicate a semantic anomaly.)

!Golf plays John.

!The beer drank the student.

!The tree liked the boy.

The famous sentence *Colorless green ideas sleep furiously* cited above is anomalous because (among other things) S-selection is violated (e.g., the verb *sleep* requires an animate subject). In Chapter 4, we will discuss the semantic relationships between a verb and its subject and objects in far more detail.

The well-formedness of a phrase depends, then, on at least two factors: whether the phrase conforms to the structural constraints of the language as expressed in the PS rules, and whether it obeys the selectional requirements of the head—both syntactic (C-selection) and semantic (S-selection).

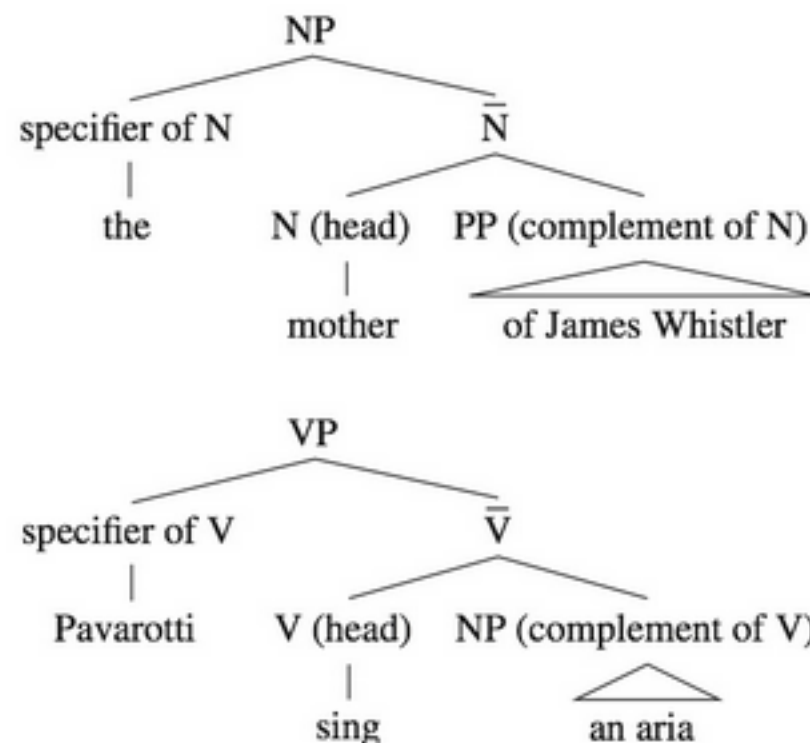
The Three Levels of Phrases

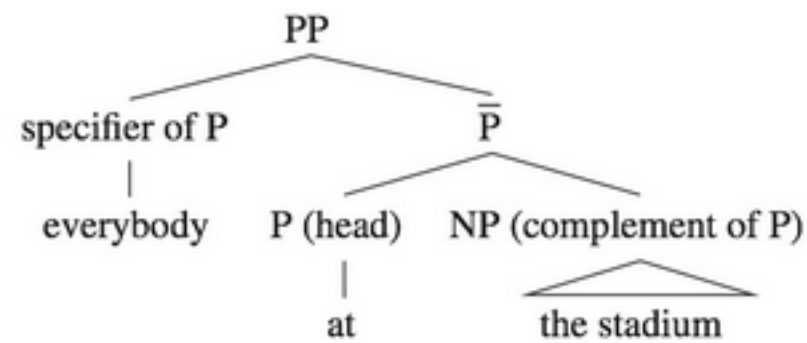
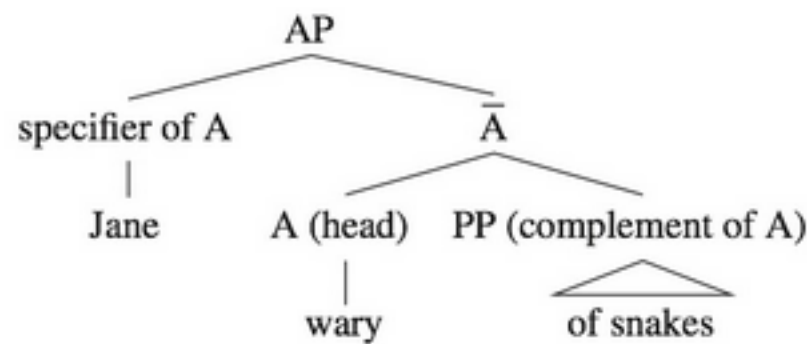
In addition to the head and its complements, a phrase may have an element preceding the head. These elements are called **specifiers**. For example, in the NP *the mother of James Whistler*, the determiner *the* is the specifier of the NP. In English, possessives may also be specifiers of NP, as in *Nellie's ball*. The specifier position may also be empty, as in the NP *dogs with bones*. PPs, APs, and VPs also have specifiers, but for various reasons they are harder to see. They usually show up when the phrase is embedded in another sentence, as in

- a. Betty made [Jane wary of snakes].
- b. I heard [Pavarotti sing an aria].
- c. I saw [everyone at the stadium].

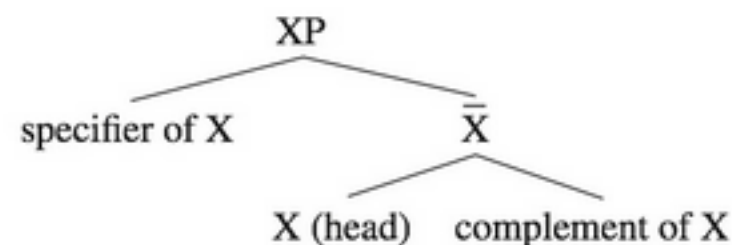
In (a) *Jane* is the specifier of the AP *wary of snakes*, in (b) *Pavarotti* is the specifier of the VP *sing an aria*, and in (c) *everyone* is the specifier of the PP *at the stadium*. *Specifier* is a purely structural notion. In English, it is the first position in the phrase, if it is present at all, and a phrase may contain at most one specifier.

Unlike complements, specifiers are not sisters of the head, but rather sisters of the phrase formed by the head and the complement. These observations tell us that all of the phrasal categories, NP, VP, AP, and PP, have a similar three-tiered structure, as follows:





To capture the generalization that each phrasal category has the same internal structure, we substitute X in place of N, V, P, A and we get the following tree:



This three-tiered structure, referred to as **X-bar ($\bar{X}$) schema**, is a template or blueprint that specifies how the phrases of a language are organized, or alternatively, how PS rules are formed. The X-bar schema “stands for” the various phrasal categories given above (and others we will see later) and applies to all syntactic phrases. The parentheses around the specifier and complement indicate that these expansions are optional and depend on the selectional properties of the head. The head is the only obligatory category of a phrase. The “bar” category is an intermediate level category necessary to account for certain syntactic phenomena that we’ll see shortly.

Assuming X-bar schema we must modify our PS rules to incorporate the three tiers. Here are the revised rules for NP:

2a: $NP \rightarrow (Det) \bar{N}$

2b: $\bar{N} \rightarrow N (XP)$

Under the new rules, NP expands as an optional Det and $\bar{N}$ and $\bar{N}$ expands as N and an optional complement of any category (XP). These rules will generate the PS tree for phrase *the mother of Whistler*, illustrated on the previous page, where XP stands for the PP *of Whistler* as complement to the head N *mother*.

We have several VP rules in our list showing the different complements to V (rules 3–6). X-bar allows us to collapse these VP rules as follows:

3a: $VP \rightarrow (\text{Spec}) \bar{V}$

3b: $\bar{V} \rightarrow V (XP)$

Under the new rules, VP expands to include an optional specifier (*Pavarotti* in sentence b above), and $\bar{V}$, which in turn contains V and an optional complement of any category such as NP as direct object in *found a puppy*. Here is the revised set of rules. Notice that the rules have been renumbered and are more compact:

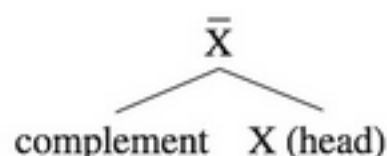
1. $S \rightarrow NP VP$
2. $NP \rightarrow (\text{Det}) \bar{N}$
3. $\bar{N} \rightarrow N (XP)$
4. $VP \rightarrow (\text{Spec}) \bar{V}$
5. $\bar{V} \rightarrow V (XP)$
6. $PP \rightarrow P NP$
7. $CP \rightarrow C S$
8. $AP \rightarrow A PP$

Our PS rules for PP, and AP (rules 6 and 8) also adhere to X-bar (e.g., $PP \rightarrow (\text{Spec}) \bar{P}$ etc.) but we omit the details. We will revisit the rules for S (rule 1) and CP (rule 7) below.

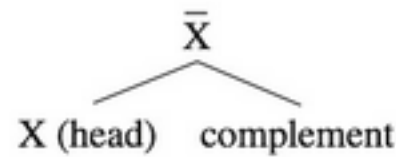
The X-bar schema is hypothesized to be part of Universal Grammar. As such, all languages have phrases that consist of heads, specifiers, and complements that relate to each other as just described. However, the order of the head and complement may differ in different languages. In English, for example, we see that the head comes first, followed by the complement. In Japanese, complements precede the head, as shown in the following examples:

Taro-ga	inu-o	mitsuketa	
Taro-subject marker	dog-object marker	found	“Taro found a dog”
Inu-ga	niwa-de	asonde	iru
Dog-subject marker	garden-in	playing	is
			“The dog is playing in the garden”

In the first sentence, the direct object complement *inu-o*, “dog,” precedes the head verb *mitsuketa*, “found.” In the second, the NP complement *niwa*, “garden,” precedes the head preposition *de*, “in.” English is a VO language, meaning that the verb ordinarily precedes its object. Japanese is an OV language, and this difference is reflected in the head/complement word order. For Japanese, the X-bar schema looks like this:



Compare this to the English schema:



X-bar schema specifies a vast amount of syntactic knowledge in a concise way. If, as many linguistics believe, X-bar is universal (order aside) and hence part of children's innate endowment for language, it also helps explain how they so quickly learn the abstract hierarchical structures of phrases in their language (see Chapter 9). Upon hearing *Taro-ga inu-o mitsuketa* (Taro dog finds), the Japanese child automatically knows not only that NP complements precede the verb in his or her language, but also that all other complements do so as well. For example, NPs precede their prepositional heads, as in *niwa-de* (garden in). The English-speaking child will just as easily come to the opposite order based on sentences such as *John found the dog*.

What Heads the Sentence

Might, could, would—they are contemptible auxiliaries.

GEORGE ELIOT (MARY ANN EVANS), *Middlemarch*, 1872

We have suggested that the structure of all phrasal categories follows the X-bar schema. One category that we have not yet discussed in this regard is sentence (S). To preserve the powerful syntactic generalization that the X-bar schema offers, we want all the phrasal categories to have a three-tiered structure with specifiers, heads, and complements, but what would these be in the case of S? To answer this question, we first observe that sentences are always “tensed.” Tense provides a time-frame for the event or state described by the verb. In English, present and past tenses are marked on the verb:

John dances. (present)

John danced. (past)

Future tense is expressed with the modal *will* (*John will dance*). Modals also express notions such as possibility (*John may dance*); necessity (*John must dance*); and ability (*John can dance*). A modal such as *may* says it is possible that the event will occur at some future time, *must* that it is necessary that the event occur at some future time, and so on. The English modals are inherently “tensed,” as shown by their compatibility with various time expressions:

John may/must/can win the race today/tomorrow.

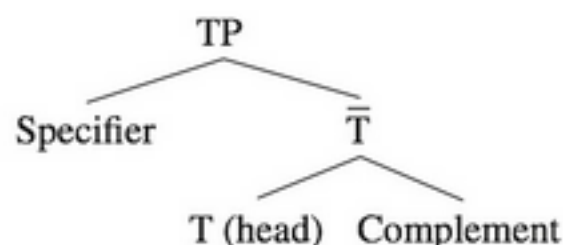
*John may/must/can win the race yesterday.

John could/would have tantrums when he was a child.

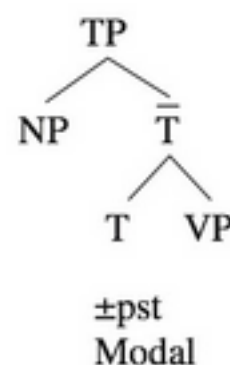
John could leave the country tomorrow.

Just as the VP is about the situation described by the verb—*eat ice cream* is about “eating”—so a sentence is about a situation or state of affairs that occurs at some point in time. Thus, the category Tense is a natural category to head S.

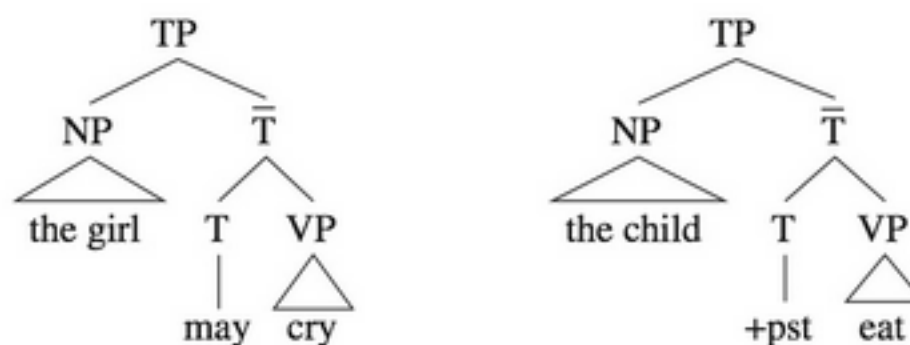
Using this insight, linguists refer to sentences as TPs (Tense Phrases) with the following structure conforming to the X-bar schema:



For sentences, or TPs, the specifier is the subject of the sentence and the complement of the T is a verb phrase. The head T contains the tense ($\pm$ pst) and modal verbs such as *can* or *would* and takes VP as its complement. The introduction of $\bar{T}$ gives the sentence its traditional subject–predicate form.



The NP left daughter of the TP functions as the subject of the sentence; the $\bar{T}$ right daughter is what is traditionally called the predicate. We are now able to represent the structures of such sentences as *The girl may cry* and *The child ate*:



In these structures, the T containing $+pst$ and *eat* is ultimately pronounced *ate*. When there is no modal under T, the present or past tense is realized on the verbal head of the VP.

Another way tense is expressed in English is by the tense-bearing word *do* that is inserted into negative sentences such as *John did not go* and questions such as *Where did John go?* In these sentences, *did* means “past tense.” Later in this chapter, we will see how *do*-insertion works.

While many of the details of X-bar syntax are beyond the scope of an introductory text, we will briefly show how the inclusion of an intermediate $\bar{X}$ tier allows the grammar to generate a wide range of sentences that could not be otherwise produced, and also further explains the recursive property of human languages.

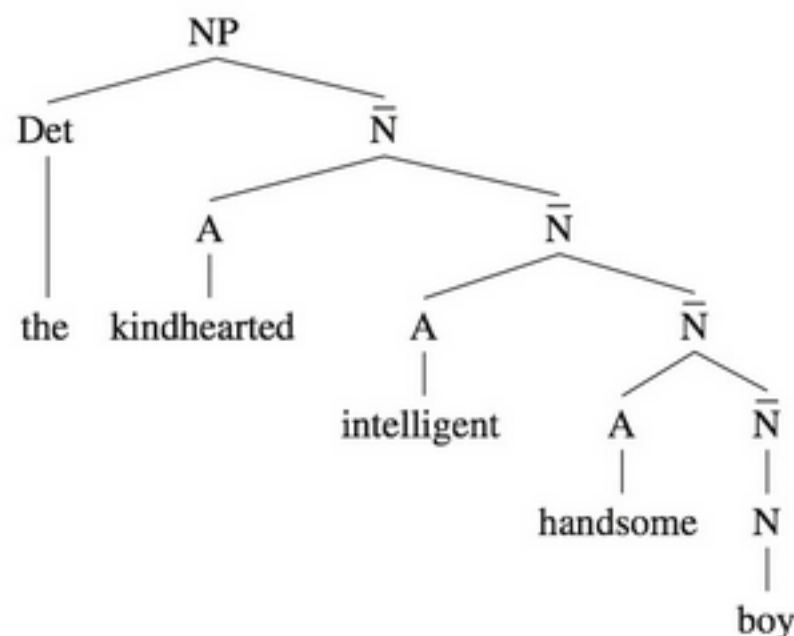
The Infinity of Language Revisited

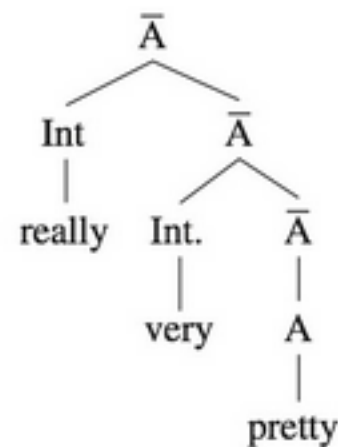
So, naturalists observe, a flea
Hath smaller fleas that on him prey;
And these have smaller still to bite 'em,
And so proceed ad infinitum.
JONATHAN SWIFT, “On Poetry, a Rhapsody,” 1733

We noted at the beginning of the chapter that languages have various means of creating longer and longer sentences. For example, an NP may contain any number of adjectives as in the *kind-hearted, intelligent, handsome boy*. One benefit of positing the abstract category $\bar{N}$ is that it allows us to account for the potentially limitless number of adjectives. Here we need a recursive rule—one that repeats itself—on $\bar{N}$:

9. $\bar{N} \rightarrow A \bar{N}$

This rule generates the NP structure in question:

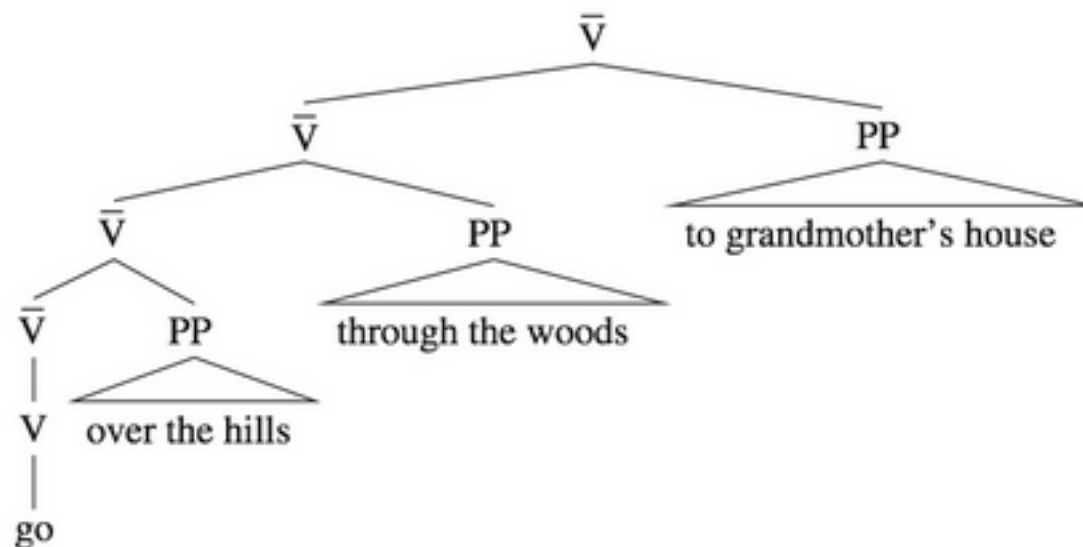


10. $\bar{A} \rightarrow \text{Int } \bar{A}$ 

A slightly different form of recursion involves PP recursion, as illustrated by *she went over the hills through the woods to grandmother's house*. . . . Sentences of this sort requires recursion on $\bar{V}$.

11. $\bar{V} \rightarrow \bar{V} \text{ PP}$

giving rise to the following subtree

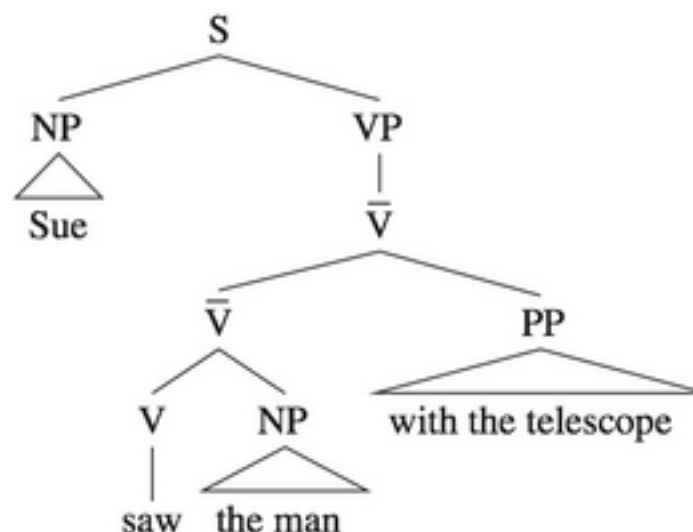


Note that the PP in (11), like the adjective in (9) and the intensifier in (10), are not complements, they are not sisters to the head of the phrase. Rather, they are sisters to $\bar{V}$, $\bar{N}$, and $\bar{A}$ respectively. A phrasal category that is sister to an $\bar{X}$ and daughter of a higher $\bar{X}$, as in the above structures, is called an **adjunct**. Like complements, adjuncts may be of any grammatical category.

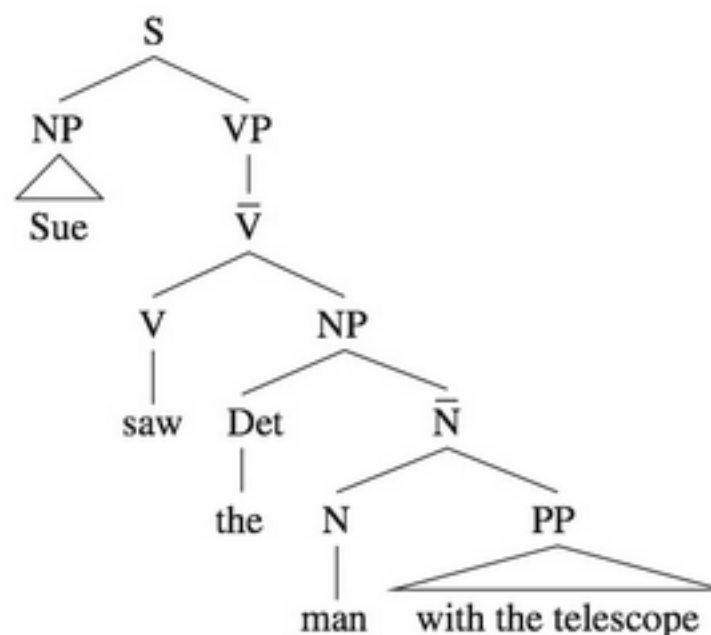
Distinguishing between complements and adjuncts is not always straightforward. Structurally, the distinctions are unambiguous: complements are sisters to X ; adjuncts are sisters to $\bar{X}$. But in analyzing sentences it is not always clear whether an addendum to a head is a complement or an adjunct. Here's one example illustrating the difference between a complement and an adjunct, an example that will bring us back to our discussion of the structural ambiguity of the sentence:

Sue saw the man with the telescope.

As discussed earlier, this sentence has more than one PS tree, each corresponding to a different meaning. Under the “instrumental” meaning (Sue used the telescope to see the man) the complement of *saw* is the simple NP *the man* and the PP is an *adjunct* introduced by rule 11. The sentence has the following constituent structure: (From now on we’ll adopt the convention of using a triangle when we are not concerned with the internal structure of the category.)



Under the second meaning of the sentence (the man is holding the telescope), *the man* and *with the telescope* form an NP constituent and the PP is a *complement* to the head noun, as illustrated in the following structure:



Thus, the different meanings arise from the fact that in the first case the PP *with the telescope* is sister to (hence modifies) the $\bar{V}$ *see the man*; but in the second case it is sister to (hence modifies) *man*. The two interpretations of this sentence are possible because the rules of syntax permit different structures for the same linear order of words.

Let us sum up our discussion thus far. We have seen that sentences have a tree-like organization. They are not simply “flat” strings of words, as shown by

various constituency tests, as well as structural ambiguities. Phrase structure trees specify (i) the grammatical categories of words and groups of words in a sentence, for example, N, V, VP, and so on, (ii) the position of categories with respect to each other, that is, word order, and (iii) the internal organization of words into hierarchically arranged phrases. The PS rules for a language thus define the (infinite set) of well-formed (grammatical) structures in that language.

Grammatical Dependencies

Method consists entirely in properly ordering and arranging the things to which we should pay attention.

RENÉ DESCARTES, *Oeuvres*, vol. X, c. 1637

In addition to the properties discussed above, the syntactic component of the grammar must describe various relationships and dependencies that hold across and within sentences. It is clear that certain sentence types are related, for example, the declarative-question pair below:

Homer will sleep.
Will Homer sleep?

Our grammar must reflect the speaker's knowledge of relationships of this sort.

Similarly, within a sentence two elements can be related even when they are separated by an arbitrary number of words. These "dependencies at a distance" provide further evidence for the hierarchical organization of sentences provided by the PS rules. Two such rules are subject-verb agreement, and question formation.

Subject-Verb Agreement

In many languages, including English, the verb must agree with the subject. The verb (in English) is marked with an -s when the subject is third-person singular and otherwise unmarked.

1. This **guy** **seems** kind of cute.
2. These **guys** **seem** kind of cute.

A simple rule that expresses the agreement relationship in terms of the linear adjacency of the noun (*guy/guys*) and verb (*seem/seems*) would work for the sentences in 1 and 2:

Linear Agreement Rule

The verb agrees in person and number with the word to its left.

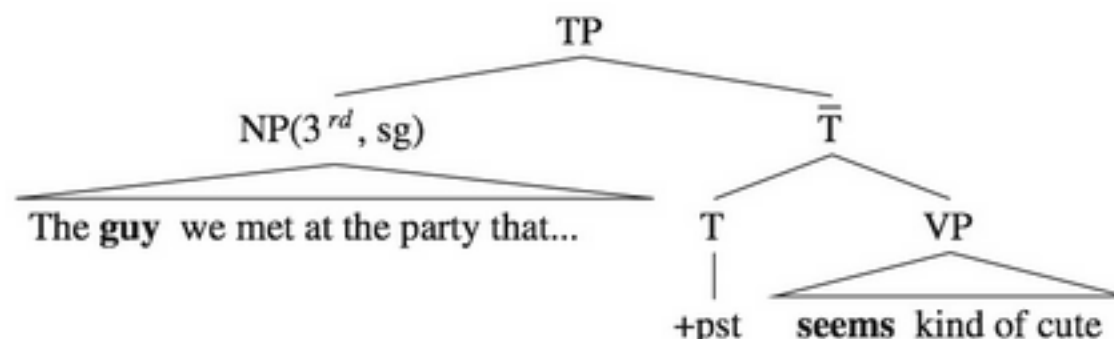
But what about the sentences in 3 and 4?

3. The **guy** we met at the party next door **seems** kind of cute.
4. The **guys** we met at the party next door **seem** kind of cute.

The verb *seem* must agree with the head of the subject NP, *guy* or *guys*, regardless of the number of words between the head noun and the verb. Moreover, there is no limit to how many words may intervene, or whether they are singular or plural, as the following sentence illustrates:

The **guy (guys)** we met at the party next door that lasted until 3 a.m. and was finally broken up by the cops who were called by the neighbors **seems (seem)** kind of cute.

The (much abbreviated) phrase structure tree below explains why this is so.



In the tree, the NP may in principle be indefinitely long and complex. However, speakers of English (and all other languages) know that agreement depends on sentence structure and not on the linear order of words: agreement is between the head of the subject NP and the main verb. As far as the rule of agreement is concerned, all other material can be ignored. (Although in actual performance, if the distance is too great, the speaker may forget what the subject was.) Thus, the rules of grammar that relate different elements in the sentence are **structure dependent** and therefore a more accurate agreement rule must be stated in terms of hierarchical structure:

Structure dependent agreement rule: The verb agrees in person and number with the *subject* of the sentence, where *subject* is defined as the NP immediately dominated by S (TP).

The fact that rules are structure dependent supports the tree-like arrangement of constituents in a sentence. If sentences were just flat strings of words, it would be impossible to state an agreement rule.

Structure dependency is a principle of Universal Grammar, and is thus found in all languages. In languages that have subject–verb agreement, the dependency is between the verb and the subject, and never some other NP such as the closest one, as shown in the following examples from Italian, German, Swahili, and English, respectively (the third-person singular agreement affix in the verb is in boldface and is governed by the boldfaced NP, not the underlined one, even though the latter is nearest the main verb):

La madre con tanti figli lavora molto.
Die Mutter mit den vielen Kindern arbeitet viel.
Mama anao watoto wengi **anajitahidi**.
The mother with many children works a lot.

Question Formation Rules



THE ARGYLE SWEATER © 2012 Scott Hilburn. Dist.
By ANDREWS MCMEEL SYNDICATION. Reprinted with
permission. All rights reserved.

Yes–no questions

I put the words down and push them a bit.

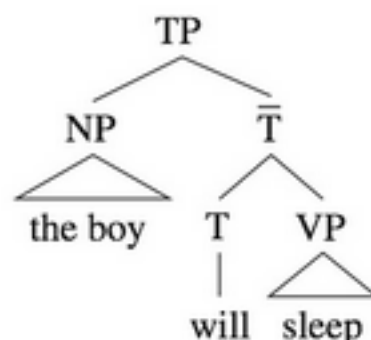
EVELYN WAUGH, quoted in *The New York Times*, April 11, 1966

Within any language certain sentence types relate systematically to other sentence types, such as the following pairs:

The boy will sleep.	Will the boy sleep?
The dog is barking.	Is the dog barking?
The man has eaten a fish.	Has the man eaten a fish?

Each pair of sentences is about the same situation. For example, the first sentence asserts that a “boy-sleeping” situation will happen. Such sentences are called **declarative** sentences. The corresponding question asks whether such a “boy-sleeping” situation will occur. Sentences of the second sort are called **yes–no questions**. The only actual difference in meaning between these sentences is that one asserts information while the other asks for confirmation of information. This meaning difference is indicated by the different word orders, illustrating that two sentences may have a structural difference that corresponds *in a systematic way* to a meaning difference. The grammar of the language must account for this fact.

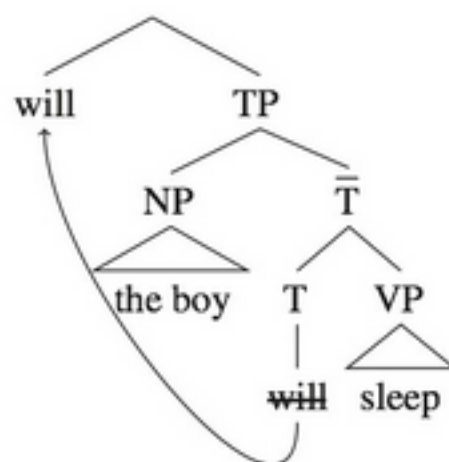
The standard way of describing these relationships is to say that the related sentences come from a common underlying structure. Yes–no questions are a case in point. A yes–no question begins life as a declarative sentence, a TP in the X-bar schema, for example:



The head of the TP, namely T (the modal *will* in this example), is central to the formation of yes–no questions as well as certain other types of sentences in English. In yes–no questions, the modal or auxiliary verb *have* or *be* appears in a different position; it precedes the subject.

The relationship between a declarative sentence and a yes–no question can be described by a rule that moves the material in T before the subject NP. This rule applies to the tree structure.

For the sentence *The boy will sleep* shown on the previous page to derive the structure below:



For descriptive purposes, we'll call this rule **Aux inversion**. Aux inversion is an example of what is traditionally referred to as a **transformational rule**. For now, we will leave unspecified the structural position that the auxiliary moves to in the tree above. We return to that below.

Thus, yes–no questions are thus generated in two steps:

1. PS-rules generate a basic structure.
2. Aux inversion applies to the basic structure to produce the derived structure.

By generating questions in two steps, we are claiming that a principled structural relationship exists between a question and its corresponding statement. Intuitively, we know that such sentences are related. The transformational rule is a formal way of representing this knowledge.

More generally, the basic structures of sentences are called **deep structures** or **d-structures**. Variants on the basic sentence structures are derived via transformational rules. The derived structures—the ones that follow the application of transformational rules—are called **surface structures** or **s-structures**. Loosely said, we *speak* and *hear* s-structures but mentally connect s-structures to d-structures. If no transformations apply, then d-structure and s-structure are the same. If transformations apply, then s-structure is the result after all transformations have taken effect.

In our discussion of the constituency test “move as a unit,” we saw other rules that dislocate elements of a sentence, for example, the active-passive pair in 1 and PP-preposing in 2:

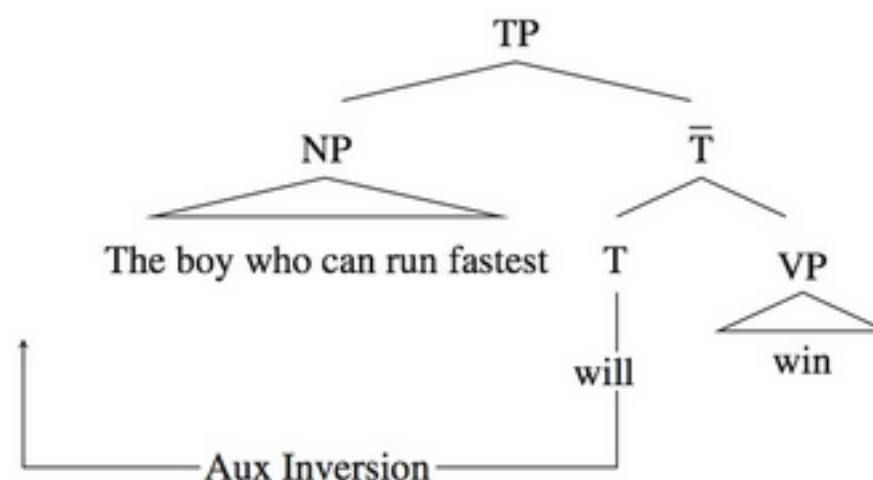
1. The child found a puppy → A puppy was found by the child.
2. The puppy played in the garden → In the garden the puppy played.

We saw earlier that the rule of subject-verb agreement is sensitive to structure and not to the linear position of elements in a sentence. We can now go further and state that all grammatical rules are structure dependent. For example, the PP-preposing rule in 2 cannot move just any string of words that begins with a preposition: It looks at the specific structure of the sentence containing the PP. This is made evident by the fact that *with a telescope, Sue saw the man* is not ambiguous. It has only the meaning “Sue used a telescope to see the man,” corresponding to the first phrase structure on page 106 where the PP is immediately dominated by the $\bar{V}$. In the structure corresponding to the other meaning, “the boy saw a man who had a telescope,” the PP is in the NP, as in the second tree on page 106. The PP-preposing transformation applies to the first structure but not the second.

Aux inversion provides yet another illustration of structure dependency.

1. The boy who **can** run fastest **will** win.
2. **Will** the boy who **can** run fastest win?
3. ***Can** the boy who run fastest **will** win?

The contrast in grammaticality of the sentences in 2 and 3 shows that to form a question Aux inversion applies to the modal within the $\bar{T}$ that is dominated by the root (highest) TP, and not simply to the *first* modal in the sentence, as illustrated in this highly abbreviated structure.



Let's now look at the structure of "Will the boy sleep?" in more detail. Thus far we have been assuming S (TP) is the root of the sentence. Strictly speaking, this is not correct. Remember our PS rule 7, repeated below.

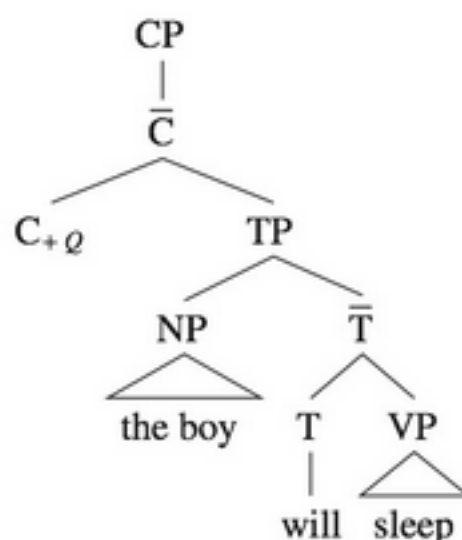
7. $CP \rightarrow C S$

From this rule, we see that CP (Complementizer Phrase) dominates S (TP). Though this rule was previously used only for embedded sentences such as *Marge noticed that Homer caught a pokémon*, yes–no questions (and many other structures) tell us that all sentences have CP as their root. Like all other categories, CP conforms to X-bar schema and hence we modify rule 7 accordingly:

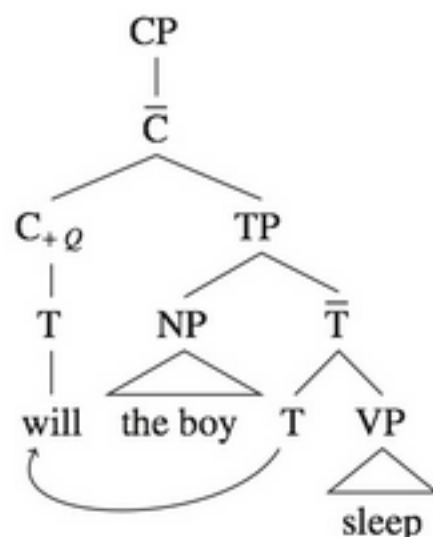
7(a).: $CP \rightarrow (\text{Spec}) \bar{C}$

7(b).: $\bar{C} \rightarrow C TP$

The sentence root is CP and TP is the complement to the head C. C contains the abstract element +Q for questions or –Q for declaratives. Putting aside the specifier of CP for the moment, the X-bar analysis of CP has the advantage that C provides a home for T when Aux inversion relocates it. The d-structure for questions is:



and the modal is moved to C:



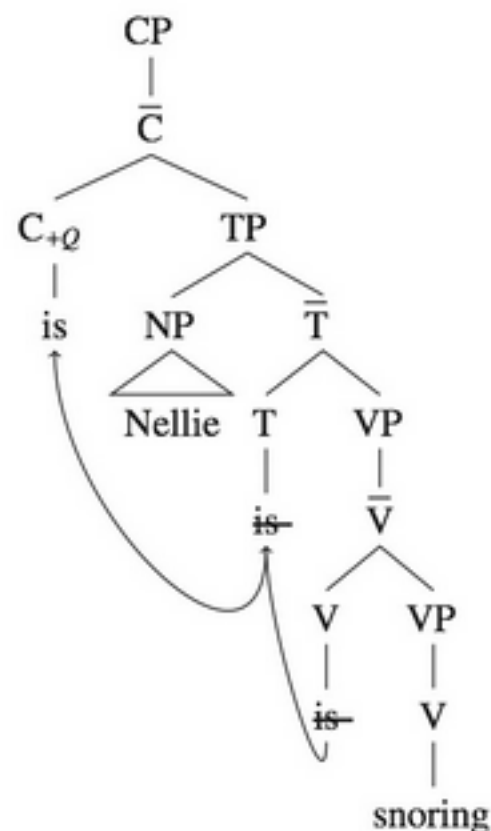
The auxiliaries *have* and *be* also undergo Aux inversion in yes–no question:

Spot has chased a squirrel. Has Spot chased a squirrel?
 Nellie is snoring. Is Nellie snoring?

But the d-structure position of these auxiliaries is not under T. We know this because they can also occur with modals (which occupy the T position) as in:

Nellie may be snoring.
 Spot must have found a squirrel.

Moreover, like other verbs in English (and unlike modals) *have* and *be* inflect for tense (and agreement): *am, is, are, was, were, have, has, had*. These observations lead us to conclude that *have/be* originate under V. When there is no modal in the sentence, *have* or *be* can undergo a movement that is not available to other verbs: they can “raise” from the position under V to T, and then undergo a second movement to C to form a question, as follows:

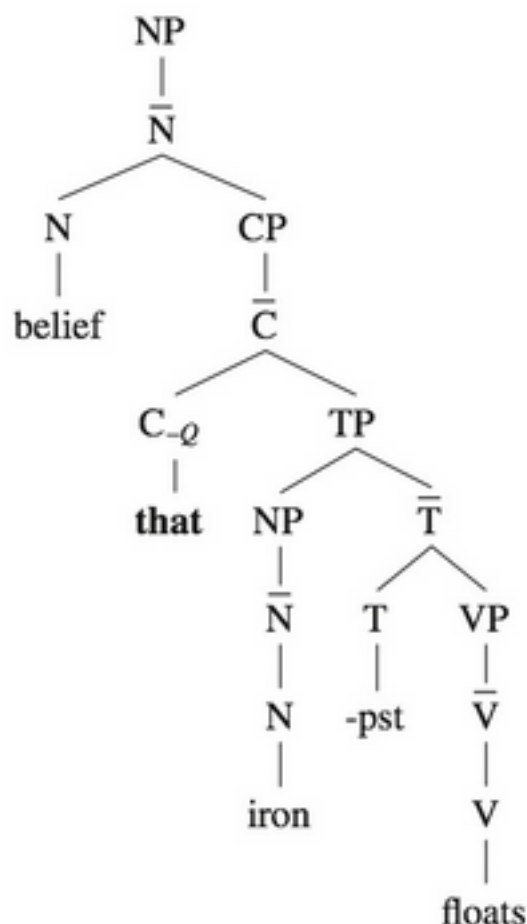


Additional PS rules would be needed to account for sentences with both *have* and *be* such as *Spot has been chasing squirrels* and even such unusual sentences as *The squirrels have been being chased by Spot*.

In addition to questions, the need for the complementizer phrase (CP) is provided by phrasal categories that take sentences (TPs) in their complements (underlined):

belief that iron floats (CP complement to head N)
 wonders if iron floats (CP complement to head V)
 happy that iron floats (CP complement to head A)
 about whether iron will sink (CP complement to head P)

The words *that*, *if*, and *whether* are complementizers and the CP has a place for them under its head C, for example:



Wh Questions

Whom are you? said he, for he had been to night school.

GEORGE ADE, "The Steel Box," in *Bang! Bang!*, 1928

We have shown that syntactic rules are structure dependent and do not pay attention to the length or content of the words in a sentence. Nowhere is this better illustrated than in **wh questions** such as the following.

1. (a) What will Max chase _____?
- (b) Where should Pete put his dog bone _____?
- (c) Which toys does Pete like _____?

Wh questions contain *wh* phrases of various syntactic categories, for example, *what* is an NP, *which* is a determiner, and *where* is a PP. They are inserted into a PS tree under the appropriate category node, like all other words. In English and many other languages, *wh* phrase generally have to move from their d-structure position, indicated by the _____ in the sentences in (1), to the beginning of the sentence ("echo questions" like *you ate what!?* behave differently and we'll ignore them here).

Several clues tell us that the *wh* phrases in (1) have undergone movement. For example, the verb *chase* in sentence (a) is transitive, yet there is no direct object

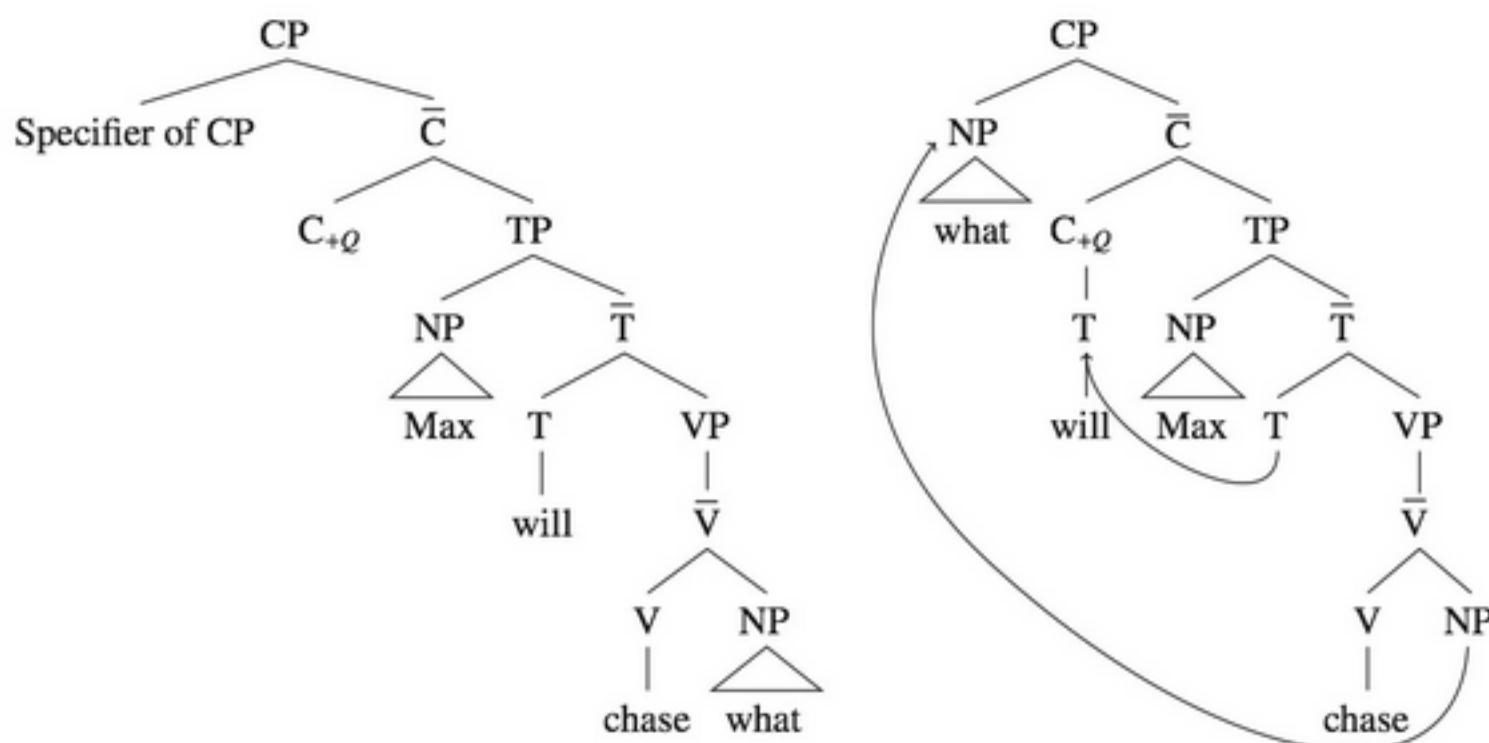
following it. There is a “gap” where the direct object should be. The verb *put* in sentence (b) is subcategorized for a direct object and a prepositional phrase, yet there is no PP following *his bone*. Finally, in sentence (c) *like* is followed by a gap and also has the third-person singular -s morpheme though it is preceded by a plural noun.

We can explain the grammaticality of the sentences in (1) despite these “abnormalities” by assuming that in each case the *wh* phrase originates in the position of the gap, as in (2), and is then moved to the beginning of the sentence by transformational rule.

2. (a) Max will chase *what*?
- (b) Pete should put his dog bone *where*?
- (c) Pete likes *which* toys?

The sentences in (1) are grammatical because the requirement that *chase* and *like* have a direct object is satisfied by the *what* and *which* toys, while the PP requirement of *put* is satisfied by *where*. The subcategorization requirements of the verbs are met prior to movement of the *wh* phrase. In any *wh* question, there is a dependency between the *wh* phrase at the beginning of the sentence and a gap somewhere else in the sentence.

Wh questions such as those in (1) are generated in several steps: phrase structure principles provide the basic declarative word orders in (2) (or more precisely the d-structure) with the *wh* expression in complement position, as required by the X-bar schema and the selectional properties of the verbs *chase*, *put* and *like*. Transformational operations then apply. Taking (2a) as illustrative, the rule *wh* movement relocates the *wh* expression from its d-structure position to a structural position at the beginning of the sentence, which we now identify as the Specifier of CP. Aux inversion moves the modal to the C, as in the derivation of yes–no questions. Following are the d-structure and s-structure trees for the sentence “What will Max chase?” is:

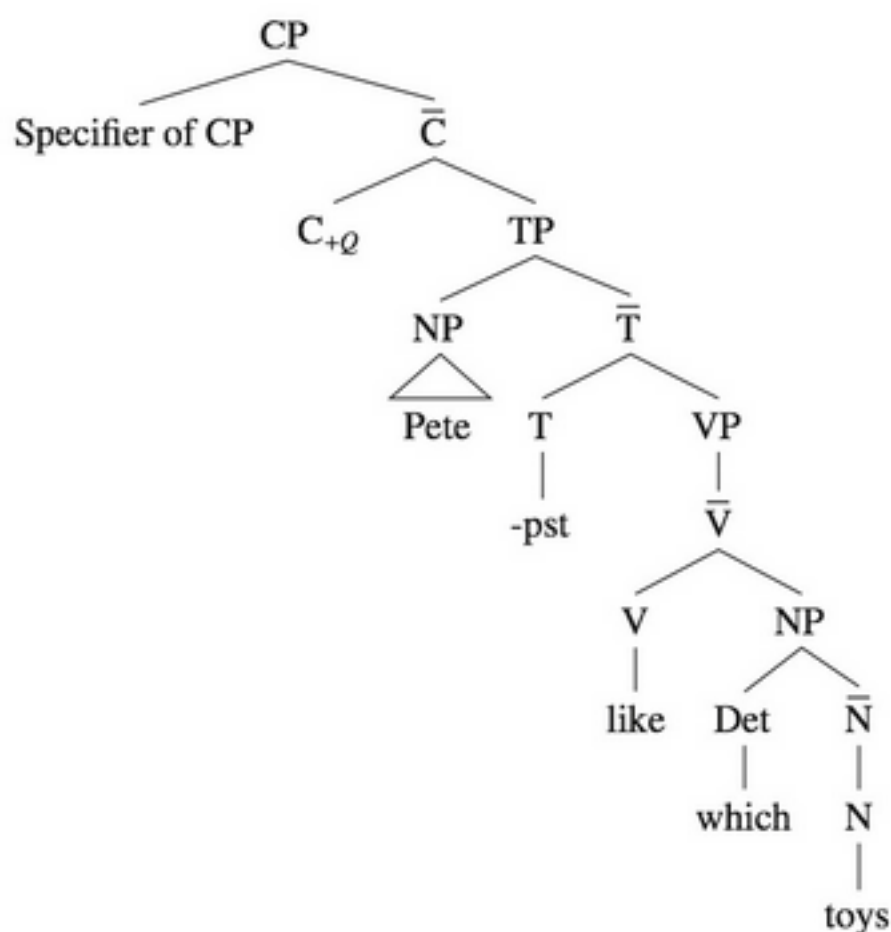


At this point, it is worth observing that like proper names and pronouns, *wh* expressions such as *what* and *who* are full NPs, not Ns. Unlike nouns, *who* and *what* cannot appear with a determiner, an adjective, or any other NP element.

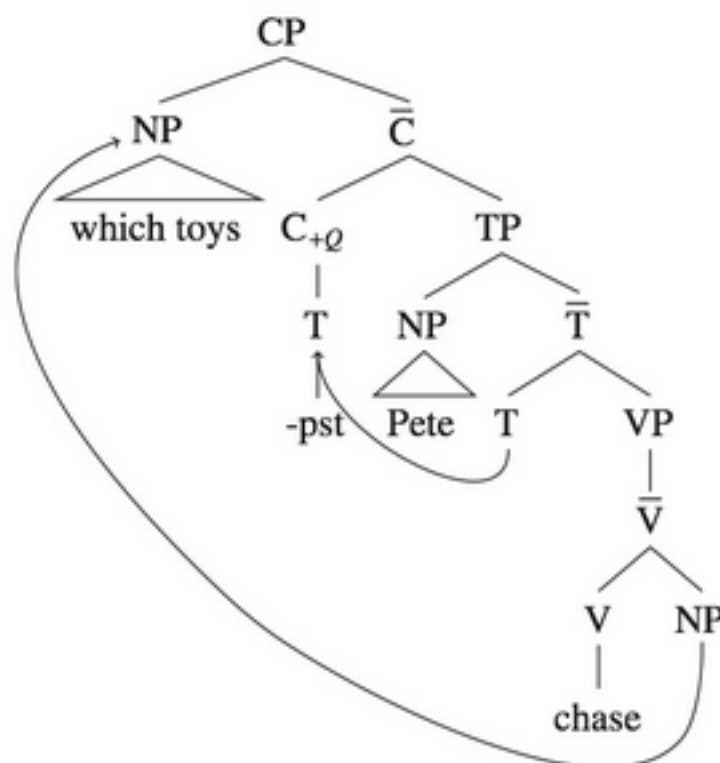
*The what did you see?

*The fast who won the race?

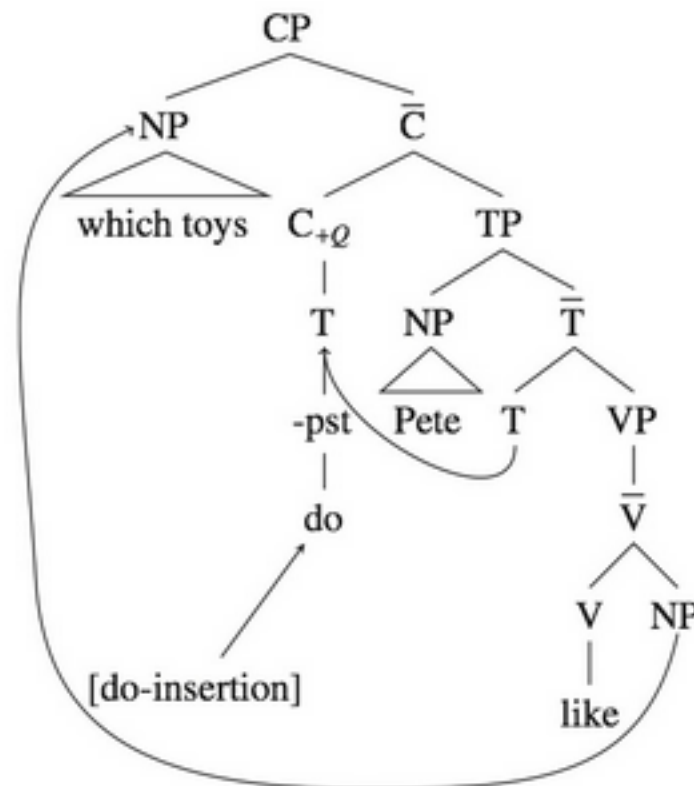
However, *what* can be a determiner, like *which*. This is reflected in the structure for the sentence *Which toys does Pete like?* Following is the d-structure of this sentence:



After *wh* movement and Aux Inversion have done their work we have this near s-structure:



Additionally, when T lacks an actual word and carries only the present tense, as in this sentence, it still undergoes movement because Aux Inversion is *structure dependent* and doesn't pay attention to the particular words (or lack thereof) under a category. With T separated from the main verb by an NP, something is needed to carry the tense. That something is the “dummy” word *do*, and it is put in place by a transformational rule of *do-insertion*, yielding the final s-structure:



Do combines with [-pst] to yield the present tense *does*. Rules that convert inflectional features such as *past tense* or *third-person present tense* into their proper phonological forms are called **spell-out rules**. They apply to the syntactic output of s-structures.

A notable property of *wh* questions is that the *wh* phrase is relocated to a position outside its original d-structure clause. Indeed, there is no limit to the distance that a *wh* phrase can move, as illustrated by the following sentences. The dashes indicate the d-structure position from which the *wh* phrases has been moved.

- Who did Helen say the senator wanted to hire _____?
- Who did Helen say the senator wanted the congressional representative to try to hire _____?
- Who did Helen say the senator wanted the congressional representative to try to convince the Speaker of the House to get the Vice President to hire _____?

“Long-distance” dependencies such as those created by *wh* movement are a fundamental part of human language. They provide still further evidence that sentences are not simply strings of words but are supported by a rich scaffolding of phrase structure trees. These trees express the underlying structure of a sentence as well as its relation to other sentences in the language, and as always reflect a person's knowledge of syntax.

UG Principles and Parameters



Whenever the literary German dives into a sentence, that is the last you are going to see of him till he emerges on the other side of the Atlantic with his Verb in his mouth.

MARK TWAIN, *A Connecticut Yankee in King Arthur's Court*, 1889

In this chapter, we have largely focused on English syntax, but many of the grammatical structures we have described for English also hold in other languages. This is because Universal Grammar (UG) provides the basic design for all human languages, and individual languages are simply variations on this basic blueprint. Imagine a new housing development. All of the houses have the same floor plan, but the occupants have some choices to make. They can have carpet or hardwood floors, curtains or blinds; they can choose their kitchen cabinets and the countertops, the bathroom tiles, and so on. This is more or less how the syntax operates. Languages conform to a basic design, and then there are choice points or points of variation.

All languages have structures that conform to the X-bar schema. Phrases consist of specifiers, heads, and complements; barred categories express recursive properties; sentences are headed by T, which is specified for information such as tense and modality; and so on.

However, languages may have different orders within the phrases and sentences. The word order differences between English and Japanese, discussed earlier, illustrate this interaction of general and language-specific properties. UG specifies the structure of a phrase. It must have a head and may take a complement of some type and have adjuncts. However, each language defines for itself the relative order of these constituents: English is head-initial, Japanese is head-final. We call the points of variation **parameters**.

All languages appear to have transformational rules for reordering elements to achieve certain purposes such as creating questions or emphasizing certain

constituents. In Dutch, for example, in which the modal moves, if there is one, as in (1), and otherwise the main verb moves, as in (2):

1. Zal Femke fietsen? (from “Femke zal fietsen.”)
will Femke bicycle ride
(Will Femke ride her bicycle?)
2. Leest Meindert veel boeken? (from “Meindert leest veel boeken.”)
reads Meindert many books
(Does Meindert read many books?)

Main verbs in Standard American English do not move. Instead, *do* spells out the stranded tense and agreement features. All languages have expressions for requesting information about *who*, *when*, *where*, *what*, and *how*. Even if the question words in other languages do not necessarily begin with “wh,” we will refer to such questions as *wh* questions. In some languages, such as Japanese and Swahili, the *wh* phrase does not move. It remains in its original d-structure position. In Japanese the sentence is marked with a question suffix *-no*:

Taro-ga	nani-o	mitsuketa-no?
Taro	what	found

Recall that Japanese word order is SOV, so the *wh* phrase *nani* (“what”) is an object and occurs before the verb.

In Swahili, the *wh* phrase—*nani* by pure coincidence—also stays in its base position:

Ulipatia	nani	kitabu?
you gave	who	a book

However, in all languages with *wh* movement (i.e., movement of the question phrase), the question element moves to the same sentence-initial position. The “landing site” of the moved phrase is determined by UG. Among the *wh* movement languages, there is some variation. In the Romance languages, such as Italian, the *wh* phrase moves as in English, but when the *wh* phrase questions the object of a preposition, the preposition must move together with the *wh* phrase. In English the preposition can be “stranded” (i.e., left behind in its original position):

A chi hai dato il libro?
To whom (did) you give the book?
*Chi hai dato il libro a?
Who(m) did you give the book to?

In some dialects of German, long-distance *wh* movement leaves a trail of *wh* phrases:

Mit	Wem	glaubst	Du	Mit	wem	Hans	spricht?
With	whom	think	you	with	whom	Hans	talks

(Whom do you think Hans talks to?)

Wen	willst	Du	wen	Hans	anruft?
Whom	want	you	whom	Hans	call

(Whom do you want Hans to call?)

In Czech a quantity question phrase can be moved, leaving behind the NP it modifies:

Jak	velké	Václav	koupil	auto?
How	Big	Václav	bought	car

(How big a car did Václav buy?)

Despite these variations and despite the fact the *wh* phrase can move a very long distance, there are certain instances in which it cannot apply and these constraints are universal and structure dependent. For example, consider the following three “sentences.” (Remember that the position from which the *wh* phrase has moved is indicated with _____.)

1. (a) Spock asked Kirk if Scotty had fixed the warp drive?
 (b) Who did Spock ask _____ whether Scotty had fixed the warp drive?
 (c) *Who did Spock ask Kirk whether _____ had fixed the warp drive?

The only difference between the grammatical (1b) and the ungrammatical (1c) is that in (1b) the *wh* phrase originates in the higher clause, whereas in (1c) the *wh* phrase comes from inside the *whether* clause. This illustrates that the constraint against movement depends on structure and not on the length of the sentence. Some sentences can be very short and still not allow *wh* movement:

2. (a) George admired Martha’s mother.
 (b) Who did George admire?
 (c) Whose mother did George admire?
 (d) *Whose did George admire mother?

The sentences in (2) show that a *wh* phrase cannot be extracted from inside a possessive NP. In (2b) it is okay to question the whole direct object. In (2c) it is even okay to question a piece of the possessive NP, providing the entire *wh* phrase is moved, but (2d) shows that moving the *wh* word alone out of the possessive NP is illicit.

The principle of structure dependency, the X-bar principles governing the organization of phrases, and the constraints on movement just illustrated, are part of UG. These aspects of grammar need not be learned. They are part of the innate blueprint for language that the child brings to the task of acquiring a language. What children must learn are the language-specific aspects of grammar. Where there are parameters of variation, children must determine the correct choices for their language. The Japanese child must determine that the verb comes after the object in the VP, and the English-speaking child that the verb comes before it. The Dutch-speaking child acquires a rule that moves the

verb to make a question, while the English-speaking child has a more restrictive rule regarding such movement. Italian, English, and Czech children learn that to form a question the *wh* phrase moves, whereas Japanese and Swahili children determine that there is no movement. As far as we can tell, children fix these parameters very quickly. We will have more to say about how children set UG parameters in Chapter 9.

Sign Language Syntax

All languages have rules of syntax similar in kind, if not in detail, to those that we have seen for English, and sign languages are no exception. Signed languages have phrase structure (PS) rules that build hierarchical structures out of linguistic constituents and specify the word order of a given signed language. ASL is an SVO language. The signer of ASL knows that the first two sentences below are grammatical sentences of ASL, but the third is not. [The capitalized words represent signs.]

CAT CHASE DOG
 “The cat chased the dog.”
 DOG CHASE CAT
 “The dog chased the cat.”
 *CHASE CAT DOG

Unlike in English, however, adjectives can follow the head noun in ASL, as in Spanish, for example, and other spoken languages.

The PS rules also determine the grammatical functions of a sentence such as subject and object, so that a signer of ASL knows that while the first two sentences are both grammatical, they differ with respect to who is chasing whom. Finally, the PS rules of signed languages exhibit language-specific variation, just as those of spoken languages do. The grammatical sentences given above for ASL would not be grammatical for signers of Italian Sign Language (LIS or “Lingua dei Segni Italiana”), because LIS is an SOV language.

In ASL, as in English and other spoken languages, the basic word order can be modified by movement rules. For example, a direct object or other constituent such as a temporal adverb can be moved to the beginning of the sentence in a process called *topicalization*. This is done to bring attention to this constituent:

BOOK, JOHN READ YESTERDAY
 YESTERDAY, JOHN READ BOOK

It is also possible for movement to apply iteratively, giving a double topicalization structure, as in:

YESTERDAY, BOOK, JOHN READ

Topicalization in ASL is accompanied by raising the eyebrows and tilting the head upward, marking the special word order, much as intonation does in English. The use of such non-manual markers is a salient feature of signed languages and something that distinguishes them from spoken languages. Spoken

language may be accompanied by facial expressions and other non-manual gestures. But however expressive or informative such gestures are, they do not form part of the grammatical system of a spoken language as they do in signed languages.

Wh questions in ASL may also be formed via movement. In contrast to English, the movement is optional. In ASL, *wh* phrases may remain in the d-structure position as in Japanese and Swahili. The ASL equivalents of *Who did Bill see yesterday?* and *Bill saw who yesterday?* are both grammatical. As in English and other spoken languages, *wh* movement in signed languages is constrained in various ways (see Appendix D). For example, in ASL it is not possible to question one member of a coordinate structure:

*WHO JOHN KISS MARY AND _____ YESTERDAY?

*“Who did John kiss Mary and yesterday?”

Similar constraints operate in topicalization. For example, a constituent cannot be moved out of the clause beginning with another *wh* phrase:

*MOTHER, I NOT-KNOW WHAT LIKE

*“(As for) Mother, I don’t know what _____ likes.”

Wh questions in ASL are accompanied by an obligatory facial expression with a tilted head and furrowed brows. These nonmanual markers are analogous to the special intonation that indicates interrogatives in many spoken languages.

Signed languages also have complex structural means to express notions such as tense, modality, and negation. In ASL, as in English, there are several forms of negation, including NO, NOT, NONE, and NEVER, and they may follow different rules. The sign NOT, for example, can come at the end of an ASL sentence, quite unlike the behavior of the English word *not*. The structural rules for negation in ASL also require that the signer shake his or her head while producing a negative sentence, and even allow a signer to “shorten” or “reduce” the negation of a sentence to just a head shake, without producing the actual sign for NOT or NEVER. This is similar to how a speaker of English can shorten *not* to *n’t*.

ASL and other sign languages show an interaction of universal and language-specific properties. The rules of sign languages are structure-dependent, and movement rules are constrained in various ways. Other properties such as the nonmanual markers and the use of space are an integral part of the grammar of sign languages but not of spoken languages. The fact that sign languages appear to be subject to the same principles and parameters of UG that spoken languages are subject to shows us that the human brain is designed to acquire and use language, not simply speech.

Summary

Speakers of a language recognize the grammatical sentences of their language and know how the words in a sentence must be ordered and grouped to convey a certain meaning. All speakers are capable of producing and understanding an unlimited number of new sentences that have never before been spoken or heard. They also recognize ambiguities, know when different sentences mean the same thing, and correctly interpret the grammatical relations in a sentence,

such as **subject** and **direct object**. This kind of knowledge comes from their knowledge of the rules of syntax.

Sentences have structure that can be represented by **phrase structure trees** containing **syntactic categories**. Phrase structure trees reflect the speaker's mental representation of sentences. Ambiguous sentences may have more than one phrase structure tree.

Phrase structure trees reveal the linear order of words and the constituency of each syntactic category. There are different kinds of syntactic categories: **Phrasal categories**, such as NP and VP, are composed of other syntactic categories; **lexical categories**, such as Noun and Verb, and **functional categories**, such as Det and T, often correspond to individual words. The hierarchical structure of the phrasal categories is universal and is specified by **X-bar schema**. NPs, VPs, and so on are headed by nouns, verbs, and the like. The sentence (S or TP) is headed by T, which carries such information as tense and modality.

The particular order of elements within the phrase is subject to language-particular variation and can be expressed through the **phrase structure rules** of each language, which conform to the X-bar schema. Here is a composite of all the phrase structure rules given in this chapter renumbered to keep phrasal types together.

1. $S (= TP) \rightarrow NP \bar{T}$
2. $\bar{T} \rightarrow T VP$
3. $NP \rightarrow (Det) \bar{N}$
4. $\bar{N} \rightarrow N (XP)$
5. $\bar{N} \rightarrow A \bar{N}$
6. $VP \rightarrow (Spec) \bar{V}$
7. $\bar{V} \rightarrow V (XP)$
8. $\bar{V} \rightarrow \bar{V} PP$
9. $PP \rightarrow P NP$
10. $CP \rightarrow (Spec) \bar{C}$
11. $\bar{C} \rightarrow C TP (= S)$
12. $AP \rightarrow A PP$
13. $\bar{A} \rightarrow Int \bar{A}$

A grammar is a formally stated, explicit description of the mental grammar or the speaker's linguistic competence. The **lexicon** represents the knowledge that a speaker has about the vocabulary of his or her language. This knowledge includes the syntactic categories of words as well as the **subcategorization** or **C-selection** properties of particular lexical items that specify the complements they can take, for example, whether a verb is **transitive** or **intransitive**. The lexicon also contains semantic information, including the kinds of NPs that can function as semantically coherent subjects and objects: **S-selection**. Selectional restrictions must be satisfied in the **d-structure** representation of the sentence.

Transformational rules such as Aux Inversion, Wh Movement, and do-insertion account for relationships between sentences such as declarative and interrogative pairs, including *wh* questions. The output of the transformational rules is the **s-structure** of a sentence, the structure that most closely determines

how the sentence is to be pronounced (or signed). Inflectional information, such as tense, may be represented as abstract features in the phrase structure tree. After the rules of the syntax have applied, these features are sometimes spelled out as affixes such as *-ed* or as function words such as *do*.

The basic design of language is universal. Universal Grammar specifies that syntactic rules are **structure-dependent** and that movement rules may not move phrases out of certain structures, among many other constraints, including a need to not violate the X-bar schema. These constraints exist in all languages—spoken and signed—and need not be learned. UG also contains parameters of variation, including the order of heads and complements, and the variations on movement rules. A child acquiring a language must fix the parameters of UG for that language.

References for Further Reading

- Baker, M. C. 2001. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.
- Carney, A. 2012. *Syntax: A generative introduction, 3rd ed.* Cambridge, MA: Blackwell.
- Chomsky, N. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- _____. 1972. *Language and mind, rev. ed.* New York: Harcourt Brace Jovanovich.
- _____. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Jackendoff, R. S. 1994. *Patterns in the mind: Language and human nature*. New York: Basic Books.
- Pinker, S. 1999. *Words and rules: The ingredients of language*. New York: HarperCollins.
- Radford, A. 2009. *Analysing English sentences: A minimalist approach*. Cambridge, UK: Cambridge University Press.
- _____. 2004. *English syntax: An introduction*. Cambridge, UK: Cambridge University Press.

Exercises

1. Besides distinguishing grammatical from ungrammatical sentences, the rules of syntax account for other kinds of linguistic knowledge, such as:
 - a. when a sentence is structurally ambiguous. (Cf. *The boy saw the man with a telescope*.)
 - b. when two sentences with different structures mean the same thing. (Cf. *The father wept silently*. and *The father silently wept*.)
 - c. systematic relationships of form and meaning between two sentences, like declarative sentences and their corresponding interrogative forms. (Cf. *The boy can sleep*. and *Can the boy sleep?*)

Draw on your linguistic knowledge of English to come up with an example illustrating each of these cases. (Use examples that are different from the ones in the chapter.) Explain why your example illustrates the point. If you know a language other than English, provide examples in that language, if possible.

2. Consider the following sentences:

- a. I hate war.
- b. You know that I hate war.
- c. He knows that you know that I hate war.
 - i. Write another sentence that includes sentence (c).
 - ii. What does this set of sentences reveal about the nature of language?
 - iii. How is this characteristic of human language related to the difference between linguistic competence and performance? (Hint: Review these concepts in Chapter 1.)

3. Paraphrase each of the following sentences in two ways to show that you understand the ambiguity involved:

Example: Smoking grass can be nauseating.

- i. Putting grass in a pipe and smoking it can make you sick.
- ii. Fumes from smoldering grass can make you sick.
- a. Dick finally decided on the boat.
- b. The professors appointment was shocking.
- c. The design has big squares and circles.
- d. That sheepdog is too hairy to eat.
- e. Could this be the invisible mans hair tonic?
- f. The governor is a dirty street fighter.
- g. I cannot recommend him too highly.
- h. Terry loves his wife and so do I.
- i. They said she would go yesterday.
- j. No smoking section available.
- k. We will dry clean your clothes in 24 hours.
- l. I bought cologne for my boyfriend containing 25% alcohol.
- m. The new magazine has between one and two billion readers.

4. Here are two examples where structural ambiguities lead to humorous results.

For sale: an antique desk suitable for lady with thick legs and large drawers.

We will oil your sewing machine and adjust tension in your home for \$10.00.

Using square brackets to delineate constituents, explained the ambiguity and the resulting humor of these two sentences by doing a constituent analysis.

5. Following the X-bar schema, draw two phrase structure trees to represent the two meanings of the sentence *The magician touched the child with the wand*. Be sure you indicate which meaning goes with which tree. (Hint: this is similar to an example in the text)

6. Indicate the grammatical category of each of the words in the following sentences.
 - a. The girls love sushi.
 - b. That boy has won many races.
 - c. Mary will finish her homework in the library.
 - d. A strong wind uprooted the tall trees.
 - e. My dog is exceptionally smart.
7. Consider the phrase structure rule $NP \rightarrow (Det) \bar{N}$ as given in this chapter to account for NPs such as *dogs* in *dogs bark*. Why would this rule not work for expressions such as *who*, *he*, *Mary*.
8. Draw the NP subtrees for the italicized NPs in the following sentences:
 - a. *Every mother* hopes for good health.
 - b. *A big black dog* is barking.
 - c. *Angry men in dark glasses* roamed the streets.
 - d. Challenge exercise: *Melissa's garden* is beautiful.
9. In all languages, sentences can occur within sentences. For example, in Exercise 2, sentence (b) contains sentence (a), and sentence (c) contains sentence (b). Put another way, sentence (a) is embedded in sentence (b), and sentence (b) is embedded in sentence (c). Sometimes embedded sentences appear slightly changed from their normal forms, but you should be able to recognize and underline the embedded sentences in the following examples. Underline in the non-English sentences, when given, not in the translations (the first one is done as an example):
 - a. Yesterday I noticed my accountant repairing the toilet.
 - b. Becky said that Jake would play the piano.
 - c. I deplore the fact that bats have wings.
 - d. That Guinevere loves Lorian is known to all my friends.
 - e. Who promised the teacher that Maxine wouldn't be absent?
 - f. It's ridiculous that he washes his own Rolls-Royce.
 - g. The woman likes for the waiter to bring water when she sits down.
 - h. The person who answers this question will win \$100.
 - i. The idea of Romeo marrying a 13-year-old is upsetting.
 - j. I gave my hat to the nurse who helped me cut my hair.
 - k. For your children to spend all your royalty payments on recreational drugs is a shame.
 - l. Give this fork to the person I'm getting the pie for.
 - m. *khâw chyâ waă khruu maa. (Thai)*
 He believe that teacher come
 He believes that the teacher is coming.
 - n. *Je me demande quand il partira. (French)*
 I me ask when he will leave
 I wonder when he'll leave.
 - o. *Jan zei dat Piet dit boek niet heeft gelezen. (Dutch)*
 Jan said that Piet this book not has read
 Jan said that Piet has not read this book.

10. Adhering to the X-bar schema, draw phrase structure trees for the following sentences.
 - a. The puppy found the child.
 - b. A surly passenger insulted the attendant.
 - c. The house on the hill collapsed in the earthquake.
 - d. The ice melted.
 - e. The hot sun melted the ice.
 - f. The old tree swayed in the wind.
 - g. The wondrous, beautiful, blue guitar sold for a song.
11. Create three phrase structure trees of 6, 7, 8, 9, and 10 words. Use your mental lexicon to fill in the bottoms of the trees.
12. Using one or more of the constituency tests (i.e., stand alone, move as a unit, replacement by a pronoun) discussed in the chapter, determine which of the boldfaced portions in the sentences are constituents (and which are not). Provide the grammatical category of the constituents.
 - a. Martha found **a lovely pillow** for the couch.
 - b. The **light in this room** is terrible.
 - c. I wonder **whether Bonnie has finished packing her books**.
 - d. Melissa slept **in her class**.
 - e. **Pete and Max** are fighting over the bone.
 - f. I gave a bone to Pete **and to Max** yesterday.
 - g. I gave a bone to **Pete and** to Max yesterday.
13. The two sentences below contain a **verbal particle** *up*:

He *ran up* the bill.
He *ran* the bill *up*.

 - a. The verbal particle *up* and the verb *run* depend on each other for the unique idiosyncratic meaning of the phrasal verb *run up*. (*Running up a bill* involves neither running nor the location up.) Does *up the bill* form a constituent? Give at least one argument that favors your answer.
 - b. List five other verb+ particle combinations in English.
14. In terms of C-selection restrictions, explain why the following are ungrammatical:
 - a. *The man located.
 - b. *Jesus wept the apostles.
 - c. *Robert is hopeful of his children.
 - d. *Robert is fond that his children love animals.
 - e. *The children laughed the man.
15. The complement of V may be a single NP direct object as for *find*. English also has **ditransitive verbs**, ones whose complement may be two NPs, such as *give*:

The emperor gave the vassal a castle.
Think of three other ditransitive verbs in English and give example sentences.

16. Tamil is a language spoken in India by upward of 70 million people. Following are word-for-word translations of PPs from Tamil to English:

Tamil to English Meaning

the bed on	“on the bed”
the village from	“from the village”

- a. Based on these data, is Tamil a head initial or a head final language?
 - b. What would the PS tree for a Tamil PP look like?
17. Here are three more word-for-word glosses in Tamil:

a story tell	“tell a story”
the boy a cow saw	“the boy saw a cow”
woman this slept	“this woman slept”

Do these additional data support or argue against your analysis in Exercise 16? What would the pertinent VP and NP trees look like in Tamil, based on these data?

18. Provide the d-structure for each of the following *wh* questions. Then state the grammatical function of the *wh* phrase (e.g., subject, object, etc.)

Example: Who did Mary see?

d-structure: Mary saw who

Who is the direct object

- a. Who left the party early?
 - b. Where did Mary leave her wallet?
 - c. What did you eat for dinner last night?
 - d. Who did Mary write to?
 - e. Which book did Sue read?
 - f. Whose jacket are you wearing?
 - g. Who did Al forget that Betty invited to the party?
 - h. Which bear did Goldilocks say that Mama Bear gave porridge to?
 - i. What did Goldilocks say that Mama Bear gave to Baby Bear?
19. As illustrated in the last two examples of the previous question, a *wh* phrase can move a very long distance. But as we saw in the chapter, there are certain instances in which it cannot apply and these constraints are universal and structure dependent. Consider the following sentences and state (in your own words) the constraint that blocks *wh* movement in the ungrammatical examples.
- a. Paul and John play beautiful music.
 - b. Who plays beautiful music?
 - c. Paul and who (else) play beautiful music?
 - d. *Who do Paul and play beautiful music.
 - e. The children love macaroni and cheese?
 - f. What do the children love?
 - g. What do the children love macaroni with?
 - h. *What do the children love macaroni and?

20. There are many systematic, structure-dependent relationships among sentences similar to the one discussed in the chapter between declarative and interrogative sentences. Here are some example sentences based on existential *there* sentences:

- a. (i) A boy is on the roof
(ii) There is a boy on the roof.
- b. (i) A boy danced on the roof.
(ii) *There danced a boy on the roof.
- c. (i) A boy was dancing on the roof.
(ii) There was a boy dancing on the roof.
- d. (i) A wallet was left in the restaurant.
(ii) There was a wallet left in the restaurant.
- e. (i) The boy is on the roof.
(ii) *There is the boy on the roof.
- f. (i) A man seems to be in the garden.
(ii) There seems to be a man in the garden.

Formulate in your own words a transformational rule relating the pairs in (i) and (ii). Be careful that your rule generates the grammatical pairs but not the ungrammatical ones.

Challenge exercise:

Now consider the following pair of sentences:

- g. (i) Students must be in the dorm by midnight.
(ii) There must be students in the dorm by midnight.

Can you describe the difference in meaning between sentence (i) and (ii)?

Does this meaning difference undermine the transformational analysis of *there* sentences?

21. State at least three differences between English and the following languages, using just the sentence(s) given. Ignore lexical differences (i.e., the different vocabulary). Here is an example:

Thai:	Dèg	khon	níi	kamlang	kin.
	boy	classifier	this	progressive	eat

“This boy is eating.”

Măa	tua	nán	kin	khâaw.
dog	classifier	that	eat	rice

“That dog ate rice.”

Three differences are (1) Thai has “classifiers.” They have no English equivalent. (2) The words (determiners, actually) *this* and *that* follow the noun in Thai, but precede the noun in English. (3) The “progressive” is expressed by a single separate word in Thai. The verb does not change form. In English, the progressive is indicated by the presence of the verb *to be* and the adding of *-ing* to the verb.

a. French

Cet	homme	intelligent	comprendra	la question.
this	man	intelligent	will understand	the question

'This intelligent man will understand the question.'

Ces	hommes	intelligents	comprendront	les questions.
these	men	intelligent	will understand	the questions

"These intelligent men will understand the question."

b. Japanese

Watashi	ga	sakana	o	tabete	iru.
I	<i>subject</i>	fish	<i>object</i>	eat (ing)	am
	<i>marker</i>		<i>marker</i>		

"I am eating fish."

c. Swahili

Mtoto		alivunja		kikombe.	
m-	toto	a-	li-	vunja	ki-
class	child	he	<i>past</i>	break	<i>class</i>
<i>marker</i>					<i>marker</i>
					kombe
					cup

"The child broke the cup."

Watoto		wanavunja		vikombe.	
wa-	toto	wa-	na-	vunja	vi-
<i>class</i>	child	they	<i>present</i>	break	<i>class</i>
<i>marker</i>					<i>marker</i>
					kombe
					cup

"The children break the cups."

d. Korean

Ki	sonyɔn-iee		wiyu-lil		masi-ass-ta.
ki	sonyɔn-	iee	wiyu-	lil	masi-
the	boy	<i>subject</i>	milk	<i>object</i>	drink
		<i>marker</i>		<i>marker</i>	<i>past</i>
					ta
					<i>assertion</i>

"The boy drank milk."

ki	nin		muɔs-il		mɔk-ass-ninya
ki	nin	muɔs-	il	mɔk-	ass-
he	<i>subject</i>	what	<i>object</i>	eat	<i>past</i>
	<i>marker</i>		<i>marker</i>		<i>Question</i>
					ninya

"What did he eat?"

e. Tagalog

Nakita	ni	Pedro-ng		puno	na	ang	bus.
nakita	ni	Pedro	-ng	puno	na	ang	bus
saw	<i>article</i>	Pedro	that	full	already	<i>topic</i>	bus
						<i>marker</i>	

"Pedro saw that the bus was already full."

- 22.** Transformations may delete elements, as in the following "elliptical" sentences:

- (i.) Mary will study hard for the exam and John will too.
- (ii.) John wrote a letter to someone, but I don't know who.
- (iii.) John loves carrots and Mary broccoli.

- a. Identify the omitted constituent in each of the examples above. (Hint: Do this by providing the d-structure for each sentence.)
- b. Provide three more examples of each kind of “ellipsis” illustrated above.

Research question: Consult with a speaker of another language and determine whether this language has the same kinds of ellipsis as English does, or other kinds. (If you know another language you can use your own intuitions to answer this question.)

23. **Challenge exercise:** Compare the following French and English sentences:

French	English
Jean boit toujours du vin. Jean drinks always some wine (*Jean toujours boit du vin.)	John always drinks some wine. *John drinks always some wine.
Marie lit jamais le journal. Marie reads never the newspaper (*Marie jamais lit le journal.)	Mary never reads the newspaper. *Mary reads never the newspaper.
Pierre lave souvent ses chiens. Pierre washes often his dogs. (*Pierre souvent lave ses chiens.)	Peter often washes his dogs. *Peter washes often his dogs.

- a. Based on the above data, what would you hypothesize concerning the relative positions of adverbs of frequency (e.g., *toujours*, *jamais*, *souvent*, *always*, *never*, *often*) and the verbs they modify in French and English?
 - b. Now suppose that UG specifies that in *all languages* the adverbs of frequency must precede $\bar{V}$ reflecting a phrase structure rule $\bar{V} \rightarrow \text{Adv } \bar{V}$. Describe in words—don’t worry about the details—a transformational rule needed to derive the correct surface word order for French? (Hint: review the discussion of the auxiliaries *have* and *be* raising and the phrase structure tree that illustrates that process.)
 - c. In terms of that transformational rule how do French and English differ?
24. Dutch and German are Germanic languages related to English, and as in English, *wh* questions are formed by moving a *wh* phrase to sentence-initial position.

In what way are the rules of question formation in Dutch and German different from those in English? Base your answer on the following data:

German	Dutch
i. Was hat Karl gekauft? what has Karl bought “What has Karl bought?”	Wat heeft Wim gekocht? what has Wim bought “What has Wim bought?”

	German	Dutch
ii.	Was kauft Karl? What buys Karl “What does Karl buy?”	Wat koopt Wim? what buys Wim “What does Wim buy?”
iii.	Kauft Karl das Buch? buys Karl the book “Does Karl buy the book?”	Koopt Wim het boek? buys Wim the book “Does Wim buy the book?”

25. **Challenge exercise:** We noted that it is often not straightforward to distinguish adjuncts from complements. “One-replacement” provides a test: only nouns with adjuncts can be substituted for by *one*, as in *a patient with a broken arm* and *one with a broken leg* (adjunct), but nouns with true complements do not allow one-replacement, so that **a patient of the doctor* and *one of the chiropractor* is not well-formed. Here are four examples of complements and four of adjuncts. Apply the *one*-replacement test to determine which is which:
- the man with the golden arm
 - a voter for proposition eighteen
 - my cousin’s arrival at his home
 - the construction of a retaining wall
 - the boat in the river
 - the ocean white with foam
 - the desecration of the temple
 - the betrayal of Julius Caesar
26. **Challenge research exercise:** X-bar theory demands binary branching and that a head may have one and only one complement. Ditransitive verbs such as *write* and *give*. (they are numerous) pose problems insofar as fitting into the strict (dare we say “Procrustean”) requirements of X-bar. This research project asks you to examine the work that has been done to accommodate the facts of ditransitive verbs with X-bar theory.