

Syntax: The Analysis of Sentence Structure

William O'Grady

... the game is to say something new with old words.

—RALPH WALDO EMERSON, *Journals*, 1849

OBJECTIVES

In this chapter, you will learn:

- how we categorize words
- how words can be combined into phrases and sentences according to a systematic schema
- that words “choose” what they can combine with in the same phrase
- how to diagram the structure of sentences
- how questions are derived from statements
- how all languages are alike in the way sentences are constructed
- how languages can differ systematically in the way sentences are constructed



LaunchPad Solo
macmillan learning

For more helpful content and quizzes, go to the LaunchPad Solo
for *Contemporary Linguistics* at launchpadworks.com.

Not much can be said in English with a single word. If we are to use language to express complex thoughts and ideas, we must be able to combine and organize words into sentences. Not just any combination of words will do, however: the pattern in 1 is not permissible, even though the same words can be combined in a different way to form the acceptable sentence in 2.

- 1) *House painted student a the.
- 2) A student painted the house.

We say that an utterance is **grammatical** if native speakers judge it to be a possible sentence of their language. Hence, 2 is grammatical, but 1 is not.

This chapter will focus on **syntax**, the component of grammar that is concerned with the form of grammatical sentences. The starting point for work on syntax is the universally accepted idea that words belong to categories of different types (nouns, verbs, and so on) and that these categories can be combined in particular ways to form phrases—and ultimately, sentences. One widely accepted way to represent the internal structure of sentences makes use of tree diagrams, like the simplified one in Figure 5.1.

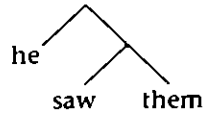


Figure 5.1 A simplified syntactic structure (tree diagram)

As we will see a little later in this chapter, structures like this are built by two interacting operations. A **Merge** operation combines words to create larger phrases and sentences, and a **Move** operation can carry an element to a new position within the structure. Before examining these operations, however, it is necessary to have a look at words and the categories to which they belong.

LANGUAGE MATTERS What's the Longest Sentence in English?

George Bernard Shaw wrote one that was 110 words long. William Faulkner's novel *Absalom, Absalom!* includes a 1,300-word sentence. James Joyce managed to produce a 4,391-word sentence (that goes on for forty pages) in *Ulysses*. But even that's not the longest known sentence—*The Rotter's Club* by Jonathon Coe contains a sentence that is 13,955 words long!

The bottom line is that there's no such thing as the world's longest sentence—any sentence can be made longer. That's because the operations that combine words can be used over and over again, without limit.

It's the right answer

I think it's the right answer

You know I think it's the right answer

Harry said you know I think it's the right answer.

....

a book

a book on the table

a book on the table near the bookcase

a book on the table near the bookcase in the office

Repeated application of the same rule to create an ever more complex structure is called **recursion**, and it's an essential part of our ability to build sentences.

1 Categories and Structure

A fundamental fact about words in all human languages is that they can be grouped together into a relatively small number of classes called **syntactic categories** or **parts of speech**. This classification reflects a variety of factors, including the types of meaning that words express, the types of affixes that they take, and the types of structures in which they can occur.

1.1 Categories of Words

Table 5.1 provides examples of the word-level categories that are most central to the study of syntax. The four most studied syntactic categories are **noun (N)**, **verb (V)**, **adjective (A)**, and **preposition (P)**. These elements, which are often called **lexical categories**, play a very important role in sentence formation, as we will soon see. A fifth and less studied lexical category consists of **adverbs (Adv)**, most of which are derived from adjectives.

Table 5.1 Syntactic categories

<i>Lexical categories (content words)</i>	<i>Examples</i>
Noun (N)	Harry, boy, wheat, policy, moisture, bravery
Verb (V)	arrive, discuss, melt, hear, remain, dislike
Adjective (A)	good, tall, old, intelligent, beautiful, fond
Preposition (P)	to, in, on, near, at, by
Adverb (Adv)	slowly, quietly, now, always, perhaps
<i>Nonlexical categories (functional categories)</i>	<i>Examples</i>
Determiner (Det)	the, a, this, these, no (as in <i>no books</i>)
Degree word (Deg)	too, so, very, more, quite
Auxiliary (Aux)	
Modal	will, would, can, could, may, must, should
Nonmodal	be, have, do
Conjunction (Con)	and, or, but

Languages may also contain **nonlexical** or **functional categories**, including **determiner (Det)**, **auxiliary verb (Aux)**, **conjunction (Con)**, and **degree word (Deg)**. Such elements generally have meanings that are harder to define and paraphrase than those of lexical categories. For example, the meaning of a determiner such as *the* or an auxiliary such as *would* is more difficult to describe than the meaning of a noun such as *hill* or *vehicle*.

A potential source of confusion in the area of word classification stems from the fact that some items can belong to more than one category.

3) *comb* used as a noun:

The woman found a comb.

comb used as a verb:

The boy should comb his hair.

4) *near* used as a preposition:

The child stood near the fence.

near used as a verb:

The runners neared the finish line.

near used as an adjective:

The end is nearer than you might think.

How then can we determine a word's category?

Meaning

One criterion involves meaning. For instance, nouns typically name entities (people and things), including individuals (*Harry, Sue*) and objects (*book, desk*). Verbs characteristically designate actions (*run, jump*), sensations (*feel, hurt*), and states (*be, remain*). Consistent with these tendencies, *comb* in 3 refers to an object when used as a noun but to an action when used as a verb.

The typical function of an adjective is to designate a property or attribute of the entities denoted by nouns. Thus, when we say *that tall building*, we are attributing the property *tall* to the building designated by the noun.

In a parallel way, adverbs typically denote properties and attributes of the actions, sensations, and states designated by verbs. In the following sentences, for example, the adverb *quickly* indicates the manner of Janet's leaving, while the adverb *early* specifies its time.

5) Janet left quickly.

Janet left early.

A word's category membership does not always bear such a straightforward relationship to its meaning, however. For example, nouns such as *difficulty, truth, and likelihood* do not name entities in the strict sense. Moreover, even though words for actions tend to be verbs, some nouns also express this type of meaning (e.g., *push* in *give someone a push* and *run* in *have a run*).

Matters are further complicated by the fact that in some cases, words with very similar meanings belong to different categories. For instance, the words *like* and *fond* are very similar in meaning (as in *Mice like/are fond of cheese*), yet *like* is a verb and *fond* is an adjective.

Inflection

Most linguists believe that meaning is only one of several criteria that enter into determining a word's category. As shown in Table 5.2, inflection can also be very useful for distinguishing among different categories of words.

Table 5.2 Lexical categories and their inflectional suffixes in English

<i>Category</i>	<i>Inflectional suffix</i>	<i>Examples</i>
Noun	plural -s possessive -'s	books, chairs, doctors John's, (the) man's
Verb	past tense -ed progressive -ing third person singular -s	arrived, melted, hopped arriving, melting, hopping arrives, melts, hops
Adjective	comparative -er superlative -est	taller, faster, smarter tallest, fastest, smartest

However, even inflection does not always provide the information needed to determine a word's category. In English, for example, not all adjectives can take the comparative and superlative suffixes (**intelligenter*, **beautifullest*) and some nouns cannot be pluralized (*moisture*, *bravery*, *knowledge*).

Distribution

A third and often more reliable criterion for determining a word's category involves the type of elements (especially functional categories) with which it can co-occur (its **distribution**). For example, nouns can typically appear with a determiner, verbs with an auxiliary, and adjectives with a degree word in the patterns illustrated in Table 5.3.

Table 5.3 Distributional properties of nouns, verbs, and adjectives

<i>Category</i>	<i>Distributional property</i>	<i>Examples</i>
Noun	occurrence with a determiner	a car, the wheat
Verb	occurrence with an auxiliary	has gone, will stay
Adjective	occurrence with a degree word	very rich, too big

In contrast, a noun cannot occur with an auxiliary, and a verb cannot occur with a determiner or degree word.

6) a noun with an auxiliary:

*will destruction

a verb with a determiner:

*the destroy

a verb with a degree word:

*very appreciate

Distributional tests for category membership are simple and highly reliable. They can be used with confidence when it is necessary to categorize unfamiliar words.

LANGUAGE MATTERS A Poem That Syntacticians Love

Thanks to distributional and inflectional clues, it's often possible to identify a word's category without knowing its meaning. The poem "Jabberwocky" by Lewis Carroll illustrates this point in a particularly brilliant way—it's interpretable precisely because readers are able to figure out that *gyre* is a verb (note the auxiliary verb to its left), that *borogoves* is a noun (it's preceded by a determiner and takes the plural ending), and so on.

'Twas brillig, and the slithy toves
 Did gyre and gimble in the wabe;
 All mimsy were the borogoves,
 And the mome raths outgrabe.
 "Beware the Jabberwock, my son!
 The jaws that bite, the claws that catch!
 Beware the Jubjub bird, and shun
 The frumious Bandersnatch!"

1.2 Phrase Structure

Sentences are not formed by simply stringing words together like beads on a necklace. Rather, they have a hierarchical design in which words are grouped together into ever larger structural units called **phrases**—*the door*, *to the door*, *go to the door*, and so on.

The Blueprint

As a first approximation, it is often suggested that the internal structure of phrases follows the design shown in Figure 5.2, which is called the **X' Schema** (X' is pronounced 'X-bar').

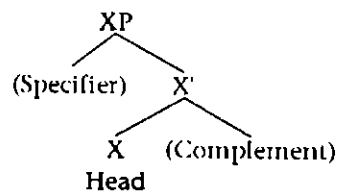


Figure 5.2 The X' Schema—a template for phrase structure

Heads

The **head** is the obligatory nucleus around which a phrase is built. For now we will focus on four categories that can function as the head of a phrase—nouns (N), verbs (V), adjectives (A), and prepositions (P).

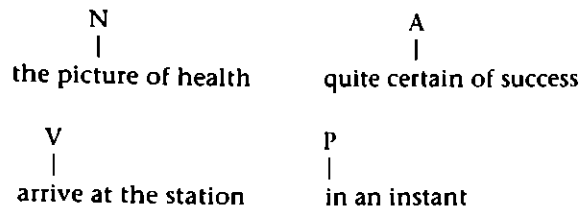


Figure 5.3 Some examples of heads

Specifiers

Specifiers have no single semantic function. Structurally, though, they are alike in that they occur at the edge of a phrase. As illustrated in Table 5.4, the specifier position in English is at the beginning of a phrase.

Table 5.4 Some specifiers

<i>Head</i>	<i>Specifier</i>	<i>Examples</i>
N	Determiner (Det) <i>the, a, some, this, those . . .</i>	<i>a picture, the map, those people, some guests</i>
V	Preverbal adverb (Adv) <i>never, perhaps, often, always, almost . . .</i>	<i>never quit, perhaps go, often failed, almost forgot</i>
A or P	Degree word (Deg) <i>very, quite, more, almost . . .</i>	<i>very smart, quite rich, almost in</i>

Note: *Almost* can be either an adverb or a degree word, depending on whether it is followed by a V or by an A or a P.

Complements

Complements, which are always phrases, provide information about entities and locations implied by the meaning of the head. For example, the meaning of *protect* implies something that is protected (*protect the environment*); the meaning of *in* implies a location (*in the house*); the meaning of *map* implies an area that is depicted, as in *a map of Oklahoma*; and so on.

As illustrated in Figure 5.4, the X' Schema ensures that when a phrase includes both a specifier and a complement in addition to the head, the specifier will occur higher than the complement. To simplify here, we don't show the internal structure of the complement phrases.

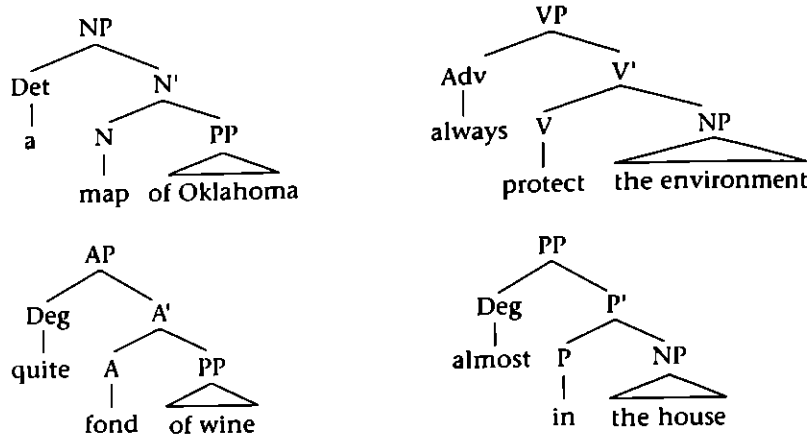


Figure 5.4 Phrases containing a head, a specifier, and a complement

However, it is common (and practical!) to represent tree structures in an abbreviated way, without the intermediate X' , when there is no specifier and/or complement, as shown in Figures 5.5 and 5.6.

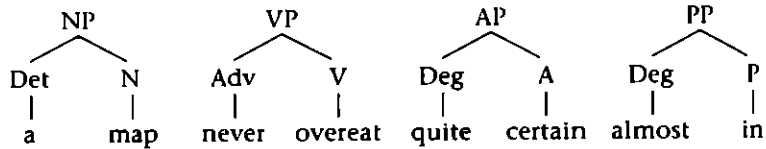


Figure 5.5 Phrases consisting of just a specifier and a head

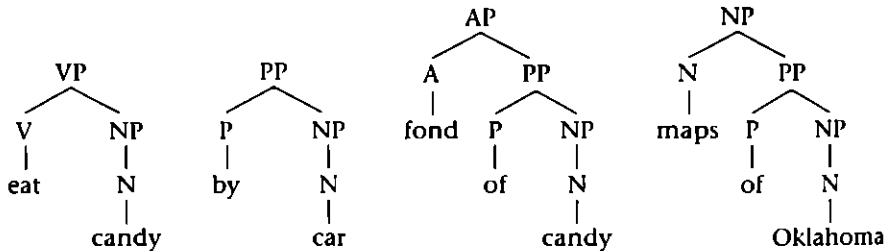


Figure 5.6 Phrases consisting of just a head and a complement

Only when the phrase contains both a specifier and a complement in addition to the head is it necessary to make use of the intermediate X' level.

In the interests of being able to consider the largest number of patterns possible, we will adopt two common additional assumptions. First, we will treat both names (*Mary, Bob, etc.*) and *pronouns* (*she, he, him, her, etc.*) as instances of the N category that do not normally take either specifiers or complements as shown in Figure 5.7.

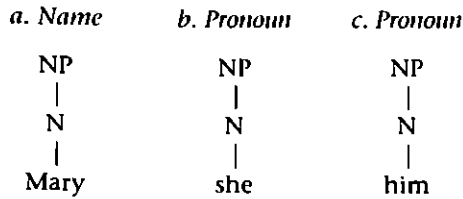


Figure 5.7 Names and pronouns

Second, we will assume that possessives (e.g., *the child's*, *Mary's*, *his*, etc.) are NPs that occur in the specifier position of a larger NP as shown in Figure 5.8.

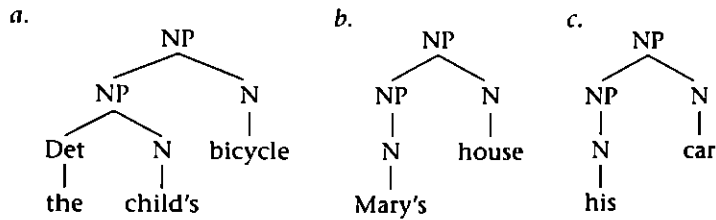


Figure 5.8 Possessives: NPs inside NPs

An appendix at the end of the chapter offers detailed instructions on how to draw tree structures; exercises 3 and 4 provide an opportunity to practice.

The Merge Operation

We can now formulate the following operation for sentence building.

7) Merge

Combine words in a manner compatible with the X' Schema.

As illustrated in Figure 5.9, the Merge operation is able to take a determiner such as *the* and combine it with the N *house* to form the NP *the house*. It is then able to take a preposition such as *in* and combine it with the NP *the house* to form the PP *in the house*.

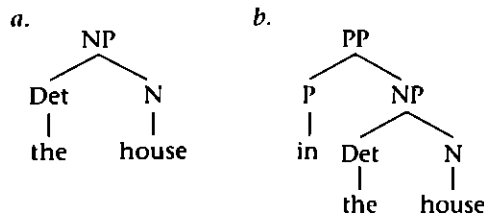


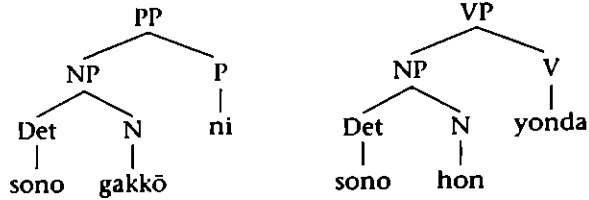
Figure 5.9 The Merge operation in action

LANGUAGE MATTERS The Mirror Image

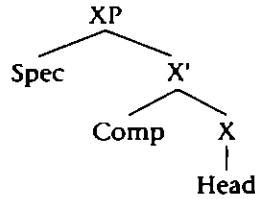
Many languages have a head-complement order that is the mirror image of the one found in English—the complement occurs before the head rather than after it. (In both types of language, the specifier appears before the head.) Japanese works that way: the V occurs at the end of the VP, the P at the end of the PP, and so on.

[sono gakkō]-ni
that school at
'at that school'

[sono hon] yonda
that book read (+Pst)
'read that book'



The version of the X' Schema needed for these languages looks like this—with the head after its complement:



About half of the world's languages use this version of the X' Schema.

Further application of the Merge operation to additional words can lead to the formation of phrases and sentences of unlimited complexity.

1.3 Sentences

The largest unit of syntactic analysis is the sentence. Sentences typically consist of an NP (often called "the subject") and a VP that are linked together by an abstract category dubbed T (for tense). As illustrated in Figure 5.10, T serves as the head of the sentence, taking the VP as its complement and the subject NP as its specifier (+Pst = past, -Pst = nonpast). What we think of as a sentence or a sentential phrase, then, is really a TP.

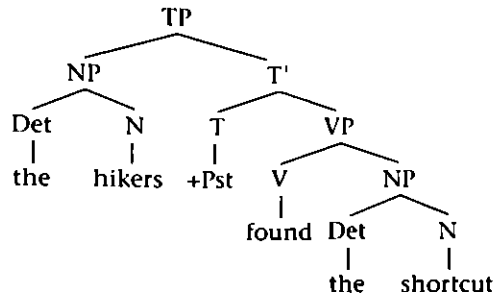


Figure 5.10 The structure of a typical sentence

The tense feature in T must be compatible with the form of the verb. So a sentence like the one above, whose head contains the feature +Pst, must contain a verb marked for the past tense (hence, *found* rather than *find*).

Although somewhat abstract, this analysis has the advantage of giving sentences the same internal structure as other phrases (with a specifier, a head, and a complement), making them consistent with the X' Schema. Moreover, because T, like all heads, is obligatory, we also account for the fact that all sentences have tense (i.e., they are all past or nonpast).

The TP structure also provides us with a natural place to locate **modal auxiliaries** such as *can*, *may*, *will*, and *must*, most of which are inherently nonpast, as shown by their incompatibility with time adverbs such as *yesterday*: **He can/will/must work yesterday*. (The modals *could* and *would* can be either past or nonpast: *He could swim when he was three/He could swim tomorrow*.) Although traditionally called auxiliary verbs, modals are treated as instances of the T category in contemporary syntactic analysis, as depicted in Figure 5.11. (Because modals have inherent tense, we will assume that it is not necessary to have the feature \pm Pst in the T position when they are used.)

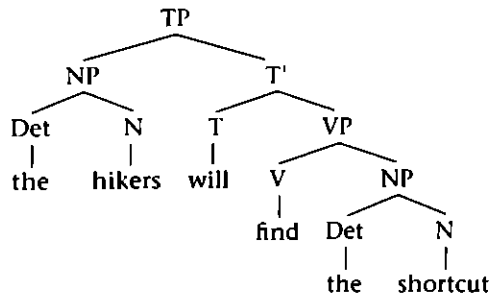


Figure 5.11 A sentence with an auxiliary verb in the T position

This structure neatly accounts not only for the fact that modals express an inherent tense but also for their positioning between the subject (the specifier) and the VP (the complement)—in the position reserved for the head of the sentence.

In fact, there are two types of auxiliary verbs, which differ in crucial ways. The modal auxiliaries are: *will, would, can, could, shall, should, may, must, might*. Because of their inherent tense, only modals are treated as instances of the T category. The **nonmodal auxiliaries** are *be, have, and do*. Unlike the modals, which are not inflected for tense or agreement, the nonmodal auxiliary verbs are marked for tense and agreement: *am-is-was; are-were; has-have-had; does-do-did*. When both types of auxiliaries appear in the same sentence, the modal always comes first, as in *They should have gone* or *They may be going*. We will consider the nonmodal auxiliaries further in Section 4.1.

The appendix at the end of the chapter outlines a procedure that will help you assign the right structure to sentences. Exercise 5 provides an opportunity to practice this procedure.

1.4 Tests for Phrase Structure

How can linguists be sure that they have grouped words together into phrases in the right way? The existence of the syntactic units, or **constituents**, found in tree structures can be independently verified with the help of special tests, although it must be noted that not every test works for every constituent. Consider, for instance, the tree structure that the X' Schema requires for the sentence *The children will stop at the corner* as shown in Figure 5.12.

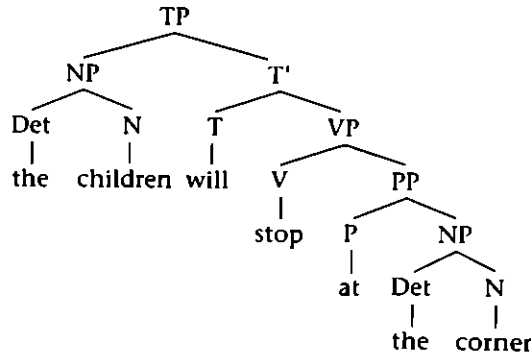


Figure 5.12 The tree structure for *The children will stop at the corner*

The Substitution Test

One piece of evidence for syntactic units comes from the fact that they can often be replaced by an element such as *they, she, he, it, do so*, and so on. (This is called a **substitution test**.)

As illustrated in 8, *the children* can be replaced by *they*, and *stop at the corner* can be replaced by *do so*—confirming that each is a syntactic unit, just as the tree structure shows by grouping the component parts together under a phrasal label such as NP, VP, and so on.

- 8) [_{NP} The children] will [_{VP} stop at the corner]. *They always do so.*
 (*they = the children; do so = stop at the corner*)

A substitution test also confirms that *at the corner* is a unit, as it can be replaced by a single word in a sentence such as 9.

- 9) The children stopped [_{PP} at the corner] and we stopped *there* too.
 (*there = at the corner*)

Elements that do not form a constituent cannot be replaced in this way. Thus, there is no word in English that we can use to replace *children stopped*, for example, or *at the*.

The Movement Test

A second indication that *at the corner* forms a constituent in Figure 5.12 is that it can be moved as a single unit to a different position within the sentence. (This is called a **movement test**.) In 10, for instance, *at the corner* can be moved from a position after the verb to the beginning of the sentence.

- 10) They stopped [_{PP} at the corner]. → [_{PP} At the corner], they stopped.

Of course, *at the*, which is not a syntactic unit, cannot be fronted in this manner (**At the, they stopped corner*). Note that the movement test often works better for PP than for other phrases.

The Coordination Test

Finally, we can conclude that a group of words forms a constituent if it can be joined to another group of words by a conjunction such as *and*, *or*, or *but*. (This is known as the **coordination test** since patterns built around a conjunction are called **coordinate structures**.) The sentence in 11 illustrates how coordination can be used to help establish that *stop at the corner* is a constituent.

- 11) The children will [_{VP} stop at the corner] *and* [_{VP} look both ways].

2 Complement Options

How can we be sure that individual words will occur with a complement of the right type in the syntactic structures that we have been building? Information about the complements permitted by a particular head is included in that head's entry in a speaker's lexicon. For instance, the lexicon for English includes an entry for *devour* that indicates that it requires an NP complement.

- 12) a. *devour* with an NP complement:
 The child devoured [_{NP} the sandwich].
 b. *devour* without an NP complement:
 *The child devoured.

The term **subcategorization** is used to refer to information about a word's complement options, such as the fact the verb *devour* belongs to a verb subcategory that requires an NP complement.

2.1 Complement Options for Verbs

Table 5.5 illustrates some of the more common complement options for verbs in English. The subscripted prepositions indicate subtypes of PP complements, where this is relevant. *Loc* stands for any preposition expressing a location (such as *near*, *on*, and *under*).

Table 5.5 Some examples of verb complements

Complement option	Sample heads	Example
∅	vanish, arrive, die	The rabbit vanished.
NP	devour, cut, prove	The professor proved [_{NP} <i>the theorem</i>].
AP	be, become	The man became [_{AP} <i>very angry</i>].
PP _{to}	dash, talk, refer	The dog dashed [_{PP} <i>to the door</i>].
NP NP	tell, hand, give	We handed [_{NP} <i>the man</i>] [_{NP} <i>a map</i>].
NP PP _{to}	hand, give, send	She gave [_{NP} <i>a diploma</i>] [_{PP} <i>to the student</i>].
NP PP _{for}	buy, cook, reserve	We bought [_{NP} <i>a hat</i>] [_{PP} <i>for Andy</i>].
NP PP _{loc}	put, place, stand	Chris put [_{NP} <i>the muffler</i>] [_{PP} <i>on the car</i>].
PP _{to} PP _{about}	talk, speak	I talked [_{PP} <i>to a doctor</i>] [_{PP} <i>about Sue</i>].
NP PP _{for} PP _{with}	open, fix	We opened [_{NP} <i>the door</i>] [_{PP} <i>for Andy</i>] [_{PP} <i>with a crowbar</i>].

The verbs in the first line of Table 5.5 (*vanish*, *arrive*, and *die*) don't take a complement, those in the second line take an NP complement, and so on.

When a verb's complement options include an NP, as in the case of *devour*, *give*, *buy*, and so on, it is said to be **transitive**, and its NP complement is often referred to as its **direct object**. Verbs like *vanish*, *arrive*, and *dash* that don't have an NP complement are called **intransitive**.

A word can belong to more than one subcategory. The verb *eat*, for example, can occur either with or without an NP complement and therefore belongs to both of the first two subcategories in Table 5.5.

13) After getting home, they ate (a snack).

Of course, not all verbs exhibit this flexibility. As we have already seen, *devour*—although similar in meaning to *eat*—requires an NP complement and therefore belongs only to the second subcategory in our table.

As the examples in Table 5.5 also show, some heads can take more than one complement. The verb *put* is a case in point, since it requires both an NP complement and a PP complement (or a locative adverb such as *there*).

- 14) a. *put* with an NP complement and a PP complement:
The librarian put [_{NP} the book] [_{PP} on the shelf].
- b. *put* without an NP complement:
*The librarian put [_{PP} on the shelf].
- c. *put* without a PP complement:
*The librarian put [_{NP} the book].

The VP *put the book on the shelf* has the structure in Figure 5.13, in which the VP consists of the head *put* and two complements—the NP *the book* and the PP *on the shelf*.

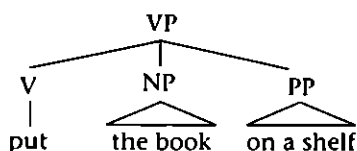


Figure 5.13 A verb with two complements

2.2 Complement Options for Other Categories

Various complement options are also available for Ns, As, and Ps. Tables 5.6, 5.7, and 5.8 provide examples of just some of the possibilities.

Table 5.6 Some examples of noun complements

Complement option	Sample heads	Example
∅	car, boy, electricity	the car
PP _{of}	memory, failure, death	the memory [_{PP} of a friend]
PP _{of} PP _{to}	presentation, description, donation	the presentation [_{PP} of a medal] [_{PP} to the winner]
PP _{with} PP _{about}	argument, discussion, conversation	an argument [_{PP} with Stella] [_{PP} about politics]

Table 5.7 Some examples of adjective complements

Complement option	Sample heads	Example
∅	tall, green, smart	very tall
PP _{about}	curious, glad, angry	curious [_{PP} about China]
PP _{to}	apparent, obvious	obvious [_{PP} to the student]
PP _{of}	fond, full, sick	fond [_{PP} of chocolate]

Table 5.8 Some examples of preposition complements

Complement option	Sample heads	Example
∅	near, away, down	(he got) down
NP	in, on, by, near	in [_{NP} the house]
PP	down, up, out	down [_{PP} into the cellar]

Here again, subcategorization ensures that particular heads can appear in tree structures only if there is an appropriate type of complement. Thus, the adjective *sick* takes an *of*-PP as its complement, while the adjective *satisfied* takes a *with*-PP.

- 15) a. sick [_{PP} of cafeteria food] (compare: *sick with cafeteria food)
 b. satisfied [_{PP} with cafeteria food] (compare: *satisfied of cafeteria food)

A good deal of what we know about our language consists of information about words and the type of complements with which they can appear. Much of this information must be stored in the lexicon, since it cannot be predicted from a word's meaning.

2.3 Complement Clauses

All human languages allow sentential phrases (or **clauses**, as they are often called) to function as complements. A simple example of this from English is given in 16.

- 16)
$$\begin{array}{c} \text{complement clause} \\ \downarrow \\ \underline{[\text{The fans hope } [\text{that the team won}]}]. \\ \uparrow \\ \text{matrix clause} \end{array}$$

The boldface bracketed phrase in 16 is called a **complement clause**; the larger underlined phrase in which it occurs is called the **matrix clause**. Words such as *that*, *whether*, and *if* are known as **complementizers** (Cs). Together with their TP complement, they form the CP (complementizer phrase) depicted in Figure 5.14.

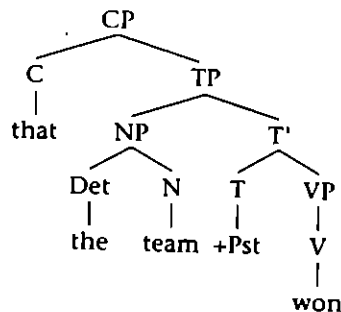


Figure 5.14 The internal structure of a CP

As we will see in Section 3.2, there is even a type of element that can occur in the specifier position under CP.

When a CP occurs in a sentence such as *16*, in which it serves as complement of the verb *hope*, the entire sentence has the structure shown in Figure 5.15.

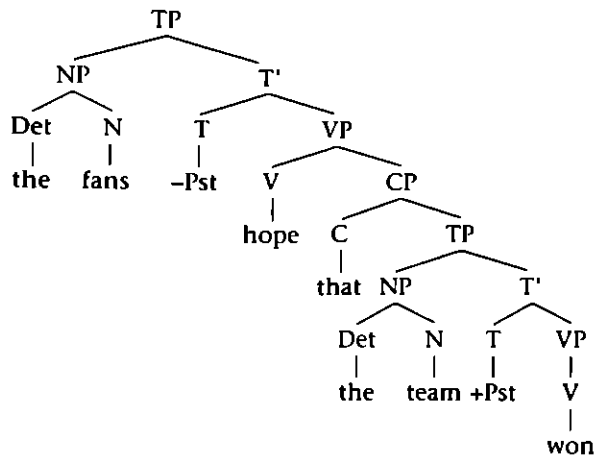


Figure 5.15 The structure of a sentence with an embedded CP

There is no limit on the number of embedded clauses that can occur in a sentence, as *17* shows.

17) Harry said [_{CP} that you know [_{CP} that I think [_{CP} that . . .

Table 5.9 provides examples of some verbs that are often found with a CP complement.

Table 5.9 Some verbs permitting CP complements

Complement(s)	Sample heads	Example
CP	believe, know, think, remember	They believe [_{CP} that Eric left].
NP CP	persuade, tell, convince, promise	They told [_{NP} Mary] [_{CP} that Eric had left].
PP _{to} CP	concede, admit	They admitted [_{PP} to Mary] [_{CP} that Eric had left].

3 Move

As we have seen, it is possible to build a very large number of different sentences by allowing the Merge operation to combine words and phrases in accordance with the

X' Schema and the subcategorization properties of individual words. Nonetheless, there are still many kinds of sentences that we cannot build. This section considers two such patterns and discusses the sentence-building operation needed to accommodate them.

3.1 Yes-No Questions


The sentences in 18 are examples of *yes-no* questions (so called because the expected response is usually “yes” or “no”).

- 18) a. **Should** those guys leave?
 b. **Can** we meet at the library?

A defining feature of *yes-no* questions is that the auxiliary verb occurs at the beginning of the sentence rather than in its more usual position after the subject, as illustrated in 19.

- 19) a. Those guys **should** leave.
 b. We **can** meet at the library.

How does the word order in 18 come about? The formation of question structures requires the use of an operation that we can call Move. Traditionally known as a **transformation** because it transforms an existing structure, Move transports the item in the T position to a new position in front of the subject.

- 20) Should those guys leave?


This analysis has at least two advantages. First, it allows us to avoid positing two types of modal auxiliary verbs in English: one that occurs between the subject and the VP and one that occurs in front of the subject. Thanks to Move, all modal auxiliaries belong in the same place—in the T position, from which they can then be moved in front of the subject in order to signal a question.

Second, the use of Move automatically captures the fact that the sentence *Should those guys leave?* is the question structure corresponding to *Those guys should leave*. According to the analysis presented here, both sentences initially have the same basic composition. They differ only in that the Move operation has applied to the T category in the question structure.

A Landing Site for T

In what position does the modal auxiliary land when it is moved in front of the subject? One promising idea assumes that TPs occur within a larger CP shell, in which the C position carries information about whether the sentence is a statement or a question. For the sake of illustration, we use the symbol +Q to indicate a question; sentences with the feature -Q in their C position will be interpreted as statements.

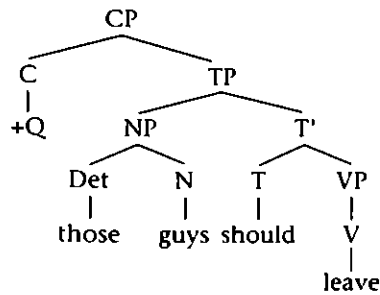


Figure 5.16 A TP inside a CP shell, with the C carrying the +Q feature

In some languages, the Q feature is spelled out as a separate morpheme (see the example from Yoruba in the box on page 187). In languages like English, where there is no such morpheme, the feature must attract another element to its position. The modal auxiliary in the T position is that element. As illustrated in Figure 5.17, T is drawn to the C position, where it attaches right next to the +Q feature.

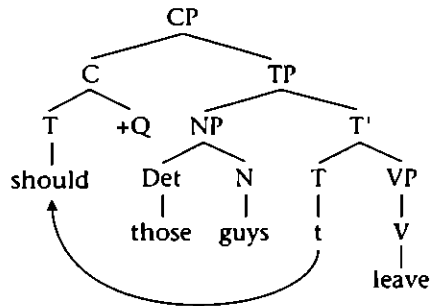


Figure 5.17 The T moves to the C position

A Move operation can do no more than change an element's position. It does not change the categories of any words and it cannot eliminate any part of the structure created by the Merge operation. Thus, *should* retains its T label even though it is moved into the C position (it changes its address, not its name). Moreover, the position that T formerly occupied remains in the tree structure. Called a **trace** and marked by the symbol *t*, it records the fact that the moved element comes from the head position within TP.

The Move operation used for *yes-no* questions is often informally called **Inversion**; it can be formulated as follows.

21) *Inversion*

Move T to the C position.

Interesting evidence that T does in fact end up in the C position comes from patterns such as 22, which contain an embedded CP.

22) I wonder [_{CP} whether those guys should leave].

Here, as Figure 5.18 shows, the C position in the embedded clause is occupied by the complementizer *whether*.

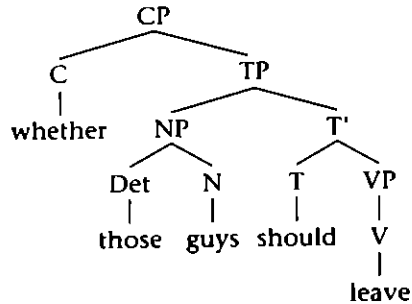
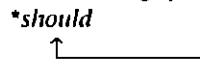


Figure 5.18 The C position in the embedded CP is filled by *whether*

Assuming that no more than one word can occur in a head position, we predict that Inversion should not be able to apply in the embedded clause since there is nowhere for the moved auxiliary verb to land. The ungrammaticality of 23 shows that this is correct.

23) Attempted inversion when there is a complementizer—the landing site is full:

*I wonder [_{CP} **whether** those guys *t* leave].



Crucially, the acceptability of Inversion improves quite dramatically when there is no complementizer in the C position. (In fact, such sentences are perfectly acceptable in Appalachian English. For other English speakers, they may sound most natural when the embedded clause is interpreted as an indirect question.)

24) Inversion in an embedded CP that does not have a complementizer:

I wonder [_{CP} **should** those guys *t* leave].



To summarize before continuing, we have introduced two changes into our system of syntactic analysis. First, we assume that TPs occur inside CP shells even when there is no visible complementizer. Second, we assume that the Inversion transformation moves T (and its contents) to the C position in order to indicate a question. In addition to giving the correct word order for the question structure, this analysis helps explain why the result of applying Inversion sounds so unnatural when the C position is already filled by another element, as in 23.

LANGUAGE MATTERS Another Way to Ask a Yes-No Question

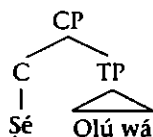
Although Inversion is a widely used question-marking strategy around the world, many languages go about things in an entirely different way. Instead of moving something to the C position, they place a special question morpheme there to begin with. Yoruba (a Benue-Congo language spoken in Nigeria) works that way.

Yoruba

Şé Olú wá?

+Q Olu come

'Did Olu come?'



Information from: Oluseye Adesola, *Yoruba: A Grammar Sketch*, Version 1.0, <http://www.africananaphora.rutgers.edu/images/stories/downloads/casefiles/YorubaGS.pdf>.

3.2 *Wh* Questions

Consider now the question construction exemplified in 25. These sentences are called *wh* questions because of the presence of a question word beginning with *wh*.

- 25) a. [_{NP} Which languages] can Jerry speak?
 b. [_{NP} What] will they talk about?

Depending on the *wh* word and its place in the sentence, *wh* words can belong to various syntactic categories, as Table 5.10 shows.

Table 5.10 The syntactic category of *wh* words

Wh word	Syntactic category	Examples
<i>who</i>	N	<i>Who did you contact?</i>
<i>what</i>	N, when it occurs by itself Det, when it occurs with a noun	<i>What did you see?</i> <i>What movie do you want to see?</i>
<i>which</i>	N, when it occurs by itself Det, when it occurs with a noun	<i>Which do you prefer?</i> <i>Which car do you prefer?</i>
<i>where</i>	Adv	<i>Where are you going?</i>
<i>when</i>	Adv	<i>When did you move to Texas?</i>
<i>why</i>	Adv	<i>Why did you leave the room?</i>
<i>how</i>	Adv, when it asks about a verb Deg, when it occurs with an adjective	<i>How did they escape?</i> <i>How rich are they?</i>

There is reason to believe that the *wh* elements at the beginning of sentences such as those in 25 have been moved there from the positions indicated in 26.

- 26) a. Jerry can speak [_{NP} which languages]
 b. They will talk about [_{NP} what]

As illustrated here, *which languages* corresponds to the complement of *speak* (compare: *Jerry can speak two languages*) and *what* corresponds to the complement of *about* (compare: *They will talk about politics*).

How, then, do the *wh* phrases end up at the beginning of the sentence? The answer is that they are attracted there by the +Q feature, which triggers the application of a Move operation known as **Wh Movement**.

- 27) a. [Which languages] can Jerry speak *t*?
 ↑ Wh Movement
 b. [What] will they talk about *t*?
 ↑ Wh Movement

A Landing Site for *Wh* Words

Because *wh* phrases end up in front of the C position (filled in 27 by a moved modal), we can infer that they end up in the specifier of CP—the only available position in that region of the sentence. We can make this idea precise by formulating the **Wh Movement** operation as follows.

28) *Wh Movement*

Move a *wh* phrase to the specifier position under CP.

The sentence *Which languages can Jerry speak?* can now be analyzed in steps, the first of which involves formation of the structure in Figure 5.19a, which includes an open specifier position under CP. *Wh Movement* and *Inversion* then apply, as depicted in Figures 5.19b and 5.19c.

LANGUAGE MATTERS Pied Piping

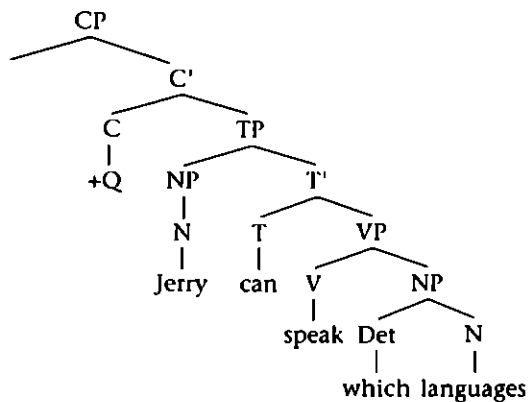
In more formal varieties of English, there is a second possibility—the entire PP containing the *wh* word can undergo *Wh Movement*.

Movement of the PP *about what*:

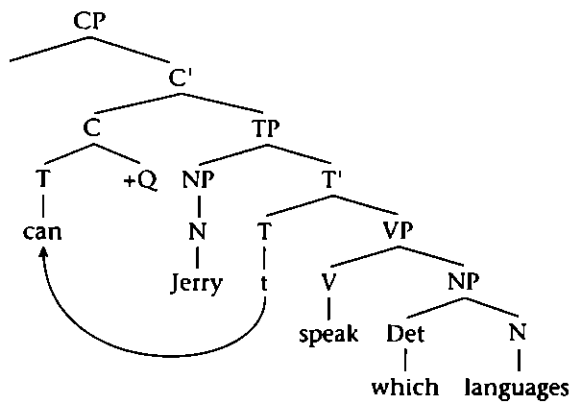
[_{PP} About what] will they *t* talk *t*?
 ↑ Inversion
 ↑ Wh Movement

This phenomenon is known as *pied-piping*, a whimsical reference to the folk tale *The Pied Piper of Hamelin*, in which (in the words of Robert Browning) “the Piper advanced and the children followed.”

- a. The structure produced by the Merge operation, with *which languages* functioning as complement of *speak*



- b. Inversion: T moves to the C position



- c. *Wh* Movement: the *wh* phrase moves to the specifier position in CP

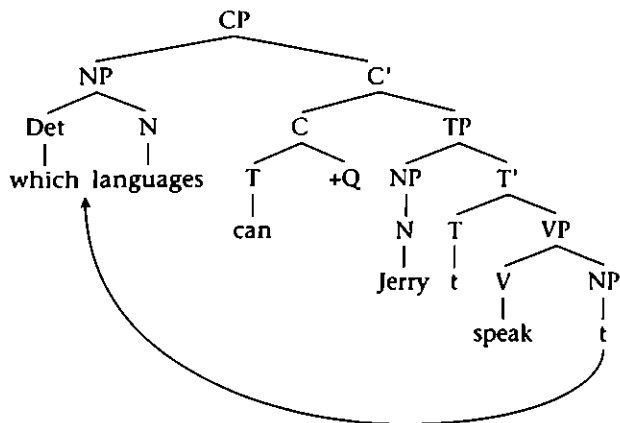
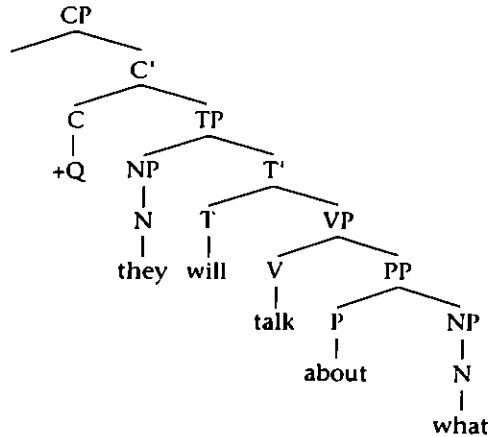


Figure 5.19 Steps for forming the sentence *Which languages can Jerry speak?*

Like Inversion, *Wh* Movement cannot eliminate any part of the previously formed structure. The position initially occupied by the *wh* phrase is therefore not lost. That is because the Move operation leaves behind an empty category (known as a trace) that marks the earlier position of the moved element. In the case at hand, the trace indicates that the NP *which languages* originates as the complement of the verb *speak*.

Figure 5.20 provides a second example, involving sentence 27b.

a. The structure produced by the Merge operation



b. Inversion and *Wh* Movement (compressed here into a single step to save space)

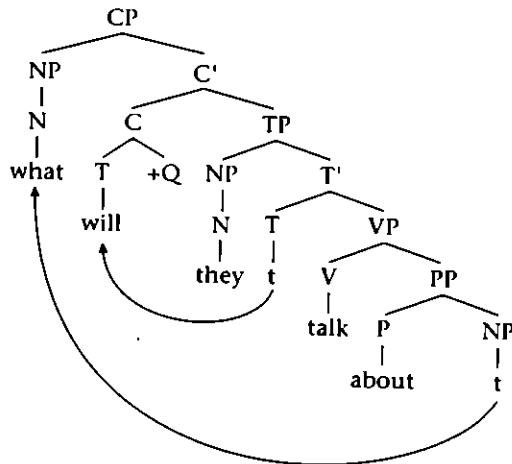


Figure 5.20 Formation of the sentence *What will they talk about?*

In the examples considered so far, the *wh* word originates as the complement of a verb or preposition. In sentences such as the following, however, the *wh* word asks about the subject (the person who will walk the dog).

29) Who will walk the dog?

The *wh* word in these patterns originates in the subject position. For the sake of generality, we assume that it subsequently moves to the specifier position in CP, even though the actual order of the words in the sentence does not change as a result of this movement (see Figure 5.21). (We will assume that there is no Inversion in this type of question structure.)

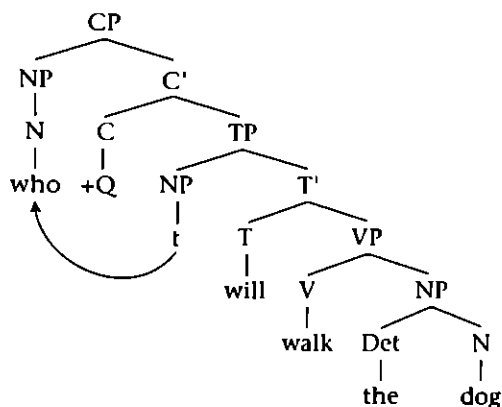


Figure 5.21 Movement of a subject *wh* phrase

3.3 Deep Structure and Surface Structure

The preceding examples show that two distinct types of mechanisms are involved in structure building. The first is the Merge operation, which creates tree structures by combining categories in a manner consistent with their subcategorization properties and the X' Schema. The second is the Move operation, which can modify these tree structures by moving an element from one position to another.

In the system sketched here, all instances of the Merge operation take place before any instances of the Move operation. This yields two distinct levels of syntactic structure, as shown in Figure 5.22. The first, called **deep structure** (or **D-structure**), is formed by the Merge operation.

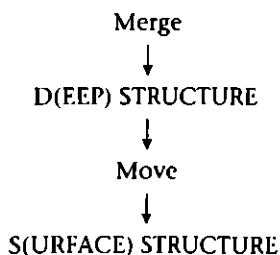


Figure 5.22 Structure-building operations

The second level of syntactic structure corresponds to the final syntactic form of the sentence. Called **surface structure** (or **S-structure**), it results from applying whatever Move operations are appropriate for the sentence in question.

A system that explicitly builds syntactic structure in this way is often called a **(transformational) generative grammar**.

4 Universal Grammar and Parametric Variation

An intriguing aspect of work on syntax is the emphasis on **Universal Grammar** (UG), the system of categories, operations, and principles shared by all languages. The key idea is that despite the many superficial differences among languages, there are certain commonalities with respect to how syntax works, including categories such as noun and verb, structure-building operations such as Merge and Move, and general constraints such as those imposed by the X' Schema.

This does not mean that languages must be alike in every way, though. Universal Grammar leaves room for variation, allowing individual languages to differ with respect to certain **parameters**. (You can think of a parameter as the set of options that UG permits for a particular phenomenon.) We have already seen one example of this sort of variation with regard to the X' Schema, which allows the head to precede its complement (as in English) or to follow it (as in Japanese). This illustrates the **head-complement parameter** stated in 30:

30) *The Head-Complement Parameter*

- Option a: The head precedes its complement.
- Option b: The head follows its complement.

Another parameter involves the placement of *wh* words. As we have seen, *wh* words move to the specifier position under CP in simple *wh* questions in English. In Chinese, though, they stay in their original position.

31) Ni mai le **shenme**?

you buy PAST **what**
'What did you buy?'

This suggests the existence of a **Wh Movement parameter** with the two options summarized in 32.

32) *The Wh Movement Parameter*

- Option a: *Wh* words move to the specifier position in CP.
- Option b: *Wh* words don't move.

The next section presents yet another example of parametric variation, this one involving verb movement.

LANGUAGE MATTERS When Wh Words Don't Move in English

English requires *Wh* Movement when the question is a simple request for information but not in certain other types of questions.

1. *Incredulity questions, expressing disbelief or surprise.* Such questions usually have a high-rising intonation and heavy stress on the *wh* word.

Speaker A: The President appointed his brother to the cabinet.

Speaker B: The President appointed WHO to the cabinet?!!

2. *Pure echo questions, which request a repetition due to partial unintelligibility.* They also manifest a rising intonation and stress on the *wh* word, although perhaps less extreme than in the case of incredulity questions.

Speaker A: The President appointed [mumble, mumble] to the cabinet.

Speaker B: The President appointed WHO to the cabinet?

3. *Quizmaster questions, sometimes used by courtroom attorneys and quiz program announcers.* They have a flat or falling intonation.

Now, Mr. Smith, you said you were where the night the Stanley Cup was stolen?

For \$15,000, the Lewis and Clark expedition began in what year?

Information from: J.-Marc Authier, "Nonquantificational *Wh* and Weakest Crossover," *Linguistic Inquiry* 24, 1 (1993): 161–168.

4.1 Verb Raising

Consider the contrast between the following two English sentences.

- 33) a. Paul always works.
b. *Paul works always.

The ungrammaticality of the second sentence is expected since the preverbal adverb *always* functions as specifier of the verb and therefore should occur before it, as in 33a. Surprisingly, however, the equivalent adverb must follow the verb in French, even though specifiers in French normally precede the head, just as they do in English.

- 34) a. If the adverb precedes the verb, the sentence is ungrammatical:
*Paul toujours travaille. (= English 33a)
Paul always work
'Paul always works.'
- b. If the adverb follows the verb, the sentence is grammatical:
Paul travaille toujours. (= English 33b)
Paul work always
'Paul always works.'

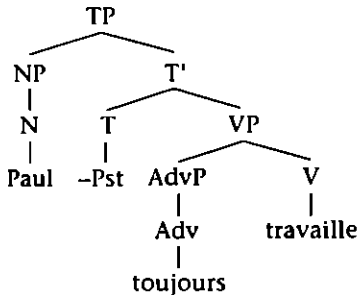
Why should this be? One possibility is that the tense feature in the T category attracts the verb to the T position in French, just as the Q feature can attract T to the C position in some languages. As a result, French has the **Verb Raising** rule outlined in 35.

35) *Verb Raising*

Move V to the T position.

This Move operation brings about the change depicted in Figure 5.23, adjoining the verb to the tense feature with which it is associated.

a. The verb in its original position



b. The verb raises to the T position

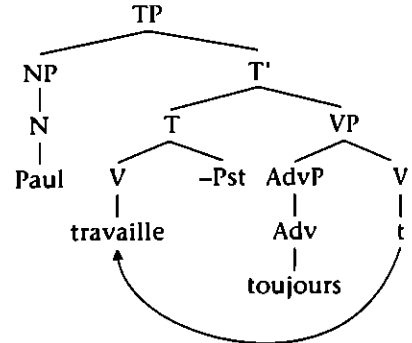


Figure 5.23 Verb Raising in French

An important piece of independent evidence for the existence of Verb Raising in French comes from Inversion. As we have already seen (Section 3.1), this operation moves the T category to the C position. In English, only auxiliary verbs occur in the T position, which explains why only they can undergo Inversion.

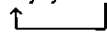
36) a. Inversion of an auxiliary verb in English:

Will you *t* stay for supper?



b. Inversion of a nonauxiliary verb in English:

*Stay you *t* for supper?



In French, however, ordinary verbs can occur in the T position, thanks to Verb Raising. This predicts that Inversion in French should be able to apply to these verbs as well as to auxiliaries. This is correct. Like English, French can form a question by moving an auxiliary leftward (when the subject is a pronoun) as illustrated in 37.

37) Inversion of an auxiliary:

As-tu *t* essayé?



'Have you tried?'

However, unlike English, French also allows inversion of nonauxiliary Vs.

38) Inversion of a nonauxiliary verb:

Vois-tu t le livre?

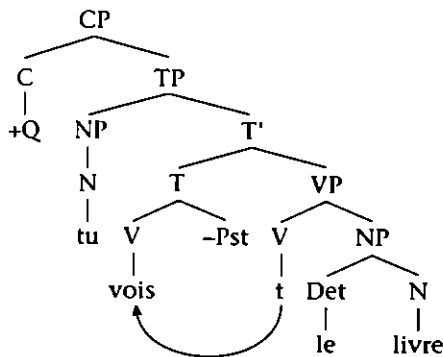


see you the book

'Do you see the book?'

Figure 5.24 depicts the interaction between Verb Raising and Inversion: the V first raises to the T position; the T complex then moves to the C position.

a. Verb Raising: the V raises to T



b. Inversion: the T complex moves to C

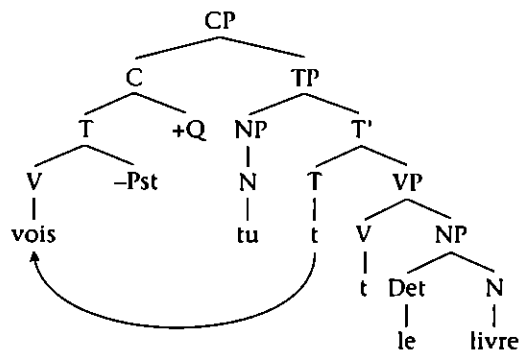


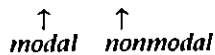
Figure 5.24 A *yes-no* question in French

Verb Raising in English

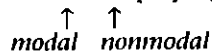
At this point, it might seem that there is a simple Verb Raising parameter with two options—raising (as in French) and no raising (as in English). This neatly accounts for the facts that we have considered so far, but matters are not so simple. As we'll see next, Verb Raising can apply in English, but only to *have* and *be*.

To begin, consider the sentences in 39, which contain two auxiliaries—one modal and one nonmodal.

39) a. The students should have finished the project.



b. The children could be playing in the yard.



As we have already seen, modal auxiliaries occur under T, but what about nonmodal auxiliaries? As depicted in Figure 5.25, they are considered to be a special type of V that takes a VP complement.

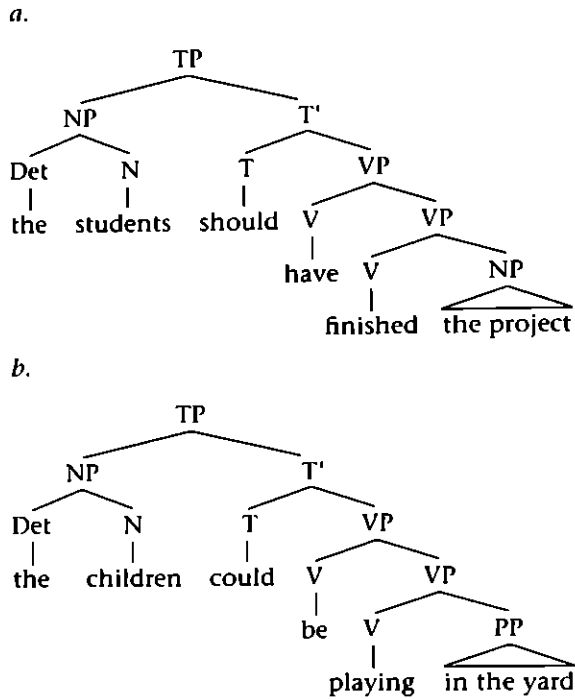


Figure 5.25 Patterns containing a modal auxiliary and a nonmodal auxiliary

As expected, only the modal auxiliary can undergo Inversion in these structures.

40) a. The modal auxiliary verb moves to the C position (grammatical):

$[_{CP} \text{Should } [_{TP} \text{the students } t \text{ have finished the project}]]?$



b. The nonmodal auxiliary moves to the C position (ungrammatical):

$*[_{CP} \text{Have } [_{TP} \text{the students should } t \text{ finished the project}]]?$



Crucially, however, a nonmodal auxiliary can undergo Inversion when there is no modal.

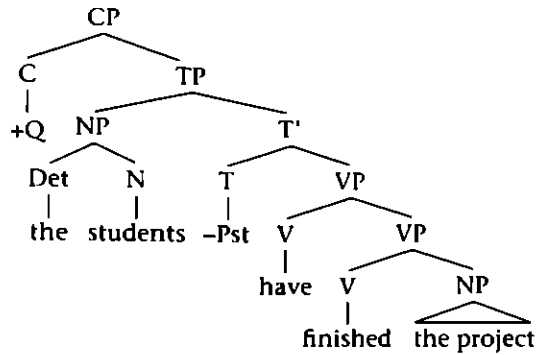
41) $[_{CP} \text{Have } [_{TP} \text{the students } t \text{ finished the project}]]?$



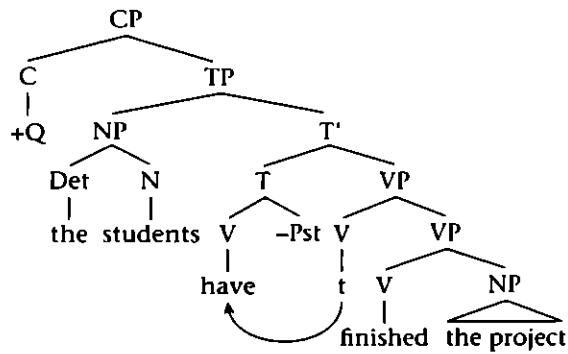
(from: The students have finished the project.)

Since Inversion involves movement from T to C, the auxiliary in 41 must have moved to the T position, and from there to the C position, as depicted in Figure 5.26.

a. Before Verb Raising



b. Verb Raising



c. Inversion (raising of the T complex to C)

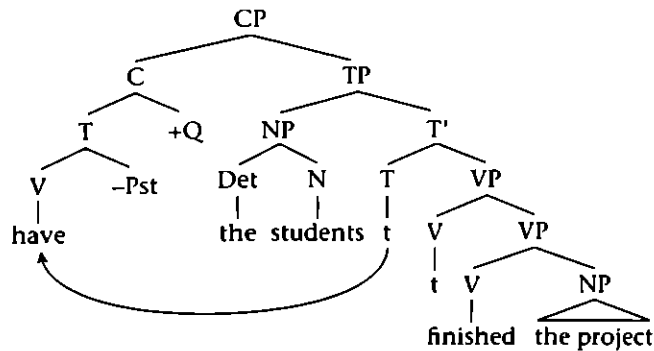


Figure 5.26 The V moves to the T position (Verb Raising); the T complex then raises to C (Inversion)

In sum, then, it appears that the two options permitted by the Verb Raising parameter should be stated as follows.

42) *The Verb Raising Parameter*

Option a: Any type of verb raises to T (French).

Option b: Only auxiliary verbs raise to T (English).

5 Some Additional Structures

Now that we have in place a basic system for forming sentences, it is possible to extend it to encompass various other syntactic phenomena, four of which will be considered here.

5.1 Modifiers

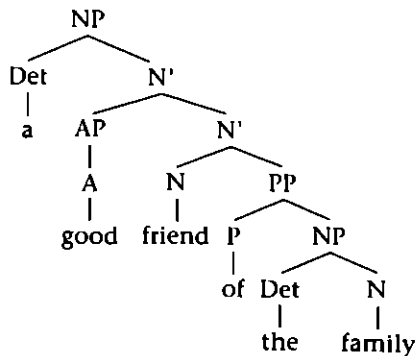
The term **modifier** is used for words and phrases that denote properties of heads. For example, adjective phrases (APs) commonly serve as modifiers of Ns, while adverb phrases (AdvPs) modify Vs.

- 43) a. a [_{AP} good] friend of the family
 b. read the instructions [_{AdvP} very carefully]

The adjective *good* denotes a property of the friend, while the AdvP *very carefully* describes the manner in which the reading occurred.

How do modifiers fit into phrase structure? For the purposes of this introduction to syntax, we will assume (as most syntacticians do) that they occur in an intermediate position, lower than specifiers but higher than complements, as illustrated in Figure 5.27.

a. An adjectival phrase



b. An adverbial phrase

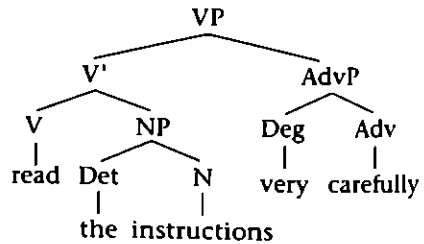


Figure 5.27 The place of modifiers in phrase structure

In Figure 5.27a, the modifier (on the left side of the head) occurs lower than the specifier (the determiner *a*), but higher than the complement (the PP *of the family*), creating two intermediate levels in the phrase. In Figure 5.27b, the modifier (the adverbial phrase) occurs on the right side of the head, once again above the complement.

5.2 Relative Clauses

Sometimes even CPs can serve as modifiers. In the following sentence, for instance, a special type of CP called a **relative clause** provides information about the N before it.

- 44) a. The friend [_{CP} *who* Leslie visited _] lives in Colorado.
 b. the choice [_{CP} *which* most people prefer _]

Relative clause structures resemble *wh* questions in two respects. First, they can begin with a *wh* word such as *who* or *which* (a so-called **relative pronoun**). Second, there is an empty position within the sentence from which the *wh* phrase has apparently been moved. (In 44a and b, that position occurs right after the transitive verb.)

The first step in the formation of the relative clause in 44a involves the D-structure in Figure 5.28; the **+Rel feature** in the C position indicates that the CP is a relative clause. (The CP here includes an open specifier position for subsequent use as a landing site for *Wh* Movement.)

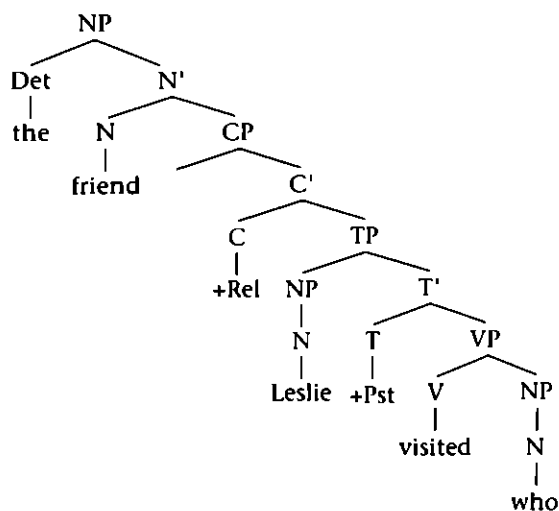


Figure 5.28 The D-structure for a relative clause

Here, the *wh* word *who* occurs as complement of the verb *visit* since it corresponds to the direct object (the person who was visited). The next step involves the application of *Wh* Movement (triggered by the **+Rel** feature in the C position) to give the structure in Figure 5.29, with the *wh* word ending up in the specifier position within CP.

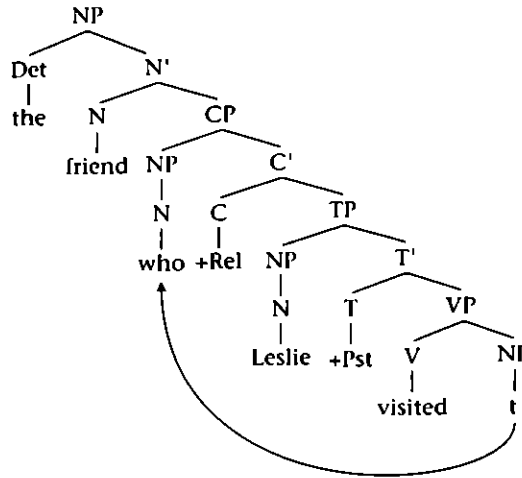


Figure 5.29 The S-structure for a relative clause: the *wh* word moves to the specifier position in CP

In the preceding example, the *wh* word originates in the direct object position. In 45, in contrast, it originates in the subject position.

45) Sue met some people [_{CP} who live in Arizona].

Here *who* corresponds to certain people who live in Arizona. The D-structure for this sentence therefore has the *wh* word in the subject position. Like other *wh* words, it subsequently moves to the specifier position within CP even though the actual order of the words in the sentence does not change as a result of this movement.

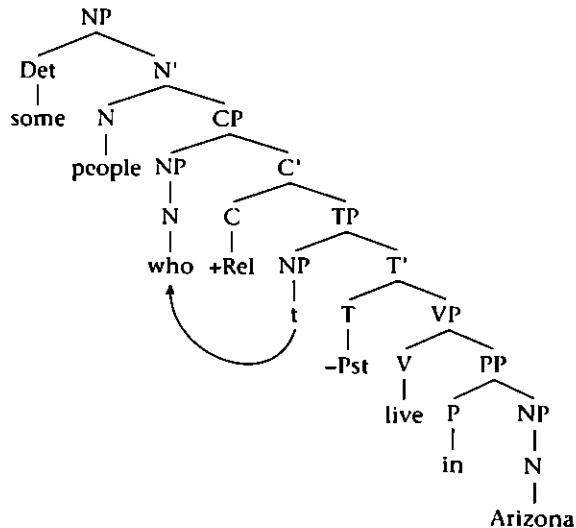


Figure 5.30 A subject relative clause

5.3 Passives

Another important syntactic phenomenon involves the relationship between the two sentence types exemplified in 46. (The first type is called **active** because the subject denotes the **agent** or instigator of the action denoted by the verb, while the second is called **passive**.)

- 46) a. Active sentence type:
A thief stole the painting.
b. Passive sentence type:
The painting was stolen (by a thief).

We will focus here on two key properties of passive constructions. First, passive constructions involve a major reduction in the importance of the agent. Whereas the agent serves as subject of an active clause, it is not expressed at all in the vast majority of passive sentences in English.

- 47) The painting was stolen.

Second, some other NP—usually the direct object of the corresponding active sentence—functions as subject in the passive sentence. This too can be seen in example 47, where the NP *the painting* serves as direct object in the active sentence and as subject in the passive sentence.

The D-structure for a passive sentence such as *The painting was stolen* is depicted in Figure 5.31. Note that the auxiliary *be* is treated as a V that takes a VP complement. We include an empty subject position under TP. (When the agent is expressed as part of a PP [e.g., *by the thief*], the PP is attached to the lower VP, to the right of the verb's NP complement.)

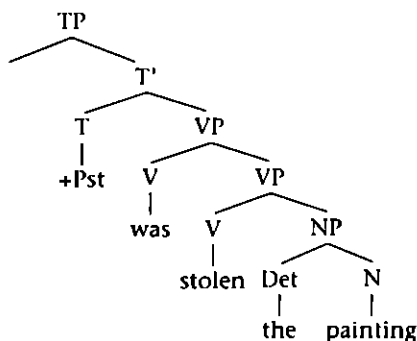


Figure 5.31 The D-structure for the passive sentence *The painting was stolen*

This D-structure is admittedly abstract—it does not sound like any sentence that we actually utter. However, it neatly captures the two key properties of passive constructions. First, the agent is not expressed, leaving the subject position open. Second, the verb has a direct object that can take over as subject.

The D-structure in Figure 5.31 is converted into an S-structure with the help of the Move operation in 48, known as **NP Movement**.