

Chapter 5

Syntax: the larger patterns of language

Learning Objectives

- Define the term *syntax*.
- Analyze the statement: “Syntax is basically subconscious knowledge.”
- List the names of the units that are larger than words and that make up sentences.
- List the names of different sentence types based on the types of clauses that construct each sentence type.
- Explain what types of sentences there are based on their meaning, function, or voice.
- Language is rule-governed. Discuss some of the general syntactic rules that a native speaker of a language knows.

- Word order is very important in some languages and less important in others. Explain why this is so.
- Report on who Noam Chomsky is and include what some of his contributions to linguistics are.
- Explain what phrase structure rules and phrase markers are.
- Discuss what is meant by saying that language has a hierarchical structure.
- Explain what is meant by the recursive property of language.
- Define transformational rules. List and explain the four basic types of transformations.
- Define grammaticality judgment.
- List the three types of ambiguity discussed in this chapter and provide at least one example of each.
- Explain how ambiguous and synonymous utterances can be seen in terms of deep and surface structure.

Syntax is a level of grammar that specifically refers to the arrangement of words and morphemes in the construction of sentences.

The word *syntax* is derived from the Greek elements *syn*, meaning *together*, and *tax*, which means *arranging*. **Syntax** is a level of grammar that specifically refers to the arrangement of words and morphemes (the lexicon) in the construction of structures such as phrases, clauses, and sentences. Syntax also deals with how these combined structures interface with external behaviors such as speech (sound), sign language (gestures), and writing to make the combined structures useful in communication. Lexical categories discussed in [Chapter 4](#) are syntactic categories for elements of language at the level of words.

Syntax can also be seen as the way in which the basically subconscious rules (tacit rules or knowledge) and categories that are part of each person's linguistic competence are used to construct sentences. Syntax deals with the interrelationship of the elements that make up sentences, and how different rules of arrangement are used to construct statements, questions, commands, and other types of utterances. In English, a native speaker

will know without having to consciously think about it that “*A to going minutes I be few in will store the” is not a grammatical sentence. However, “I will be going to the store in a few minutes” is a grammatical sentence and would be recognized as such by a native speaker of English. Words are not put together randomly to create phrases, clauses, and sentences.

When we say that syntactic rules are basically subconscious, we mean several things. First, people apply the rules of their language automatically and without noticing that they are doing anything special. Second, using the syntax of language is usually obligatory. Unless you make the grammar explicit (that is, you are consciously aware of it), you can't change it. Of course, under certain circumstances you might do just that. For example, if you are trying to imitate a dialect different from your own, you might study the grammar of your way of speaking and compare it to another.

Descriptive syntax or **descriptive grammar** refers to the mostly subconscious rules of a language that speakers use to combine smaller units into sentences. The term also refers to the study of these rules.

Prescriptive syntax or **prescriptive grammar** (as the term implies) refers to the concept that there is a correct and an incorrect way to speak, write, or sign.

When linguists and anthropologists study syntax, they are interested in describing the subconscious knowledge that people possess about the syntax of their language, not prescribing how they *should* construct sentences. What linguists and anthropologists are discovering is **descriptive syntax** or **descriptive grammar**. They listen to what people actually say and then attempt to discover the rules used. What a language teacher does in a grammar class by telling you that there is a correct or incorrect way to write or speak is **prescriptive syntax** or **prescriptive grammar**.

Transformational or **generative grammar** is a model of syntax that includes a finite set of rules that could hypothetically produce (generate) an infinite number of utterances.

There are many approaches to the study of syntax. Some are mostly descriptive and attempt to discover the rules of the syntax of a language and describe them; some are based on the analysis of the function of syntactic forms, while others are based on complex mathematical models. We will deal mostly with a model of syntax called **transformational** or **generative grammar** that incorporates descriptive and mathematical concepts but primarily is based on the assumption that many elements of syntax cannot be discovered just by studying linguistic performance. They are instead a reflection of how the human mind (the brain) works to form syntactic structures such as sentences.

Generative grammar (discussed further in the “Phrase structure rules” section later in this chapter) assumes that the general similarities found in all languages are a result of the prewired way in which the human brain must process linguistic input. What a person says, writes, or signs is the result of the complex processing of learned information in a partially innate way. This processing system is the result of human and prehuman evolution. It is, in part, what allows children to naturally, efficiently, and subconsciously learn language with relatively limited input from their social environment (see “poverty of the stimulus” in [Chapter 10](#)). In fact, in recent years some researchers have hypothesized that the abil-

ity of the human mind to automatically arrange certain types of information in a particular rule-governed way is also present in some of our closest living nonhuman relatives; that is, other primates, as well as some other animals, may have computational abilities.¹

In this chapter we are going to analyze syntactic structures in terms of traditionally defined units such as sentences. However, it should be pointed out that people often speak, sign, or write in utterances or texts that a language teacher might say are incomplete or not grammatical (see [Chapter 3](#) for a definition of utterance). This is obvious in such written forms as text messaging, tweets, and so on. When linguists analyze speech or sign language, they often also find this clipped or otherwise informal form of communication (see “Situational dialects or registers” in [Chapter 8](#)).

Syntactic construction

Types of syntactic structures

We will discuss syntax starting with larger syntactic structures, breaking them down into their smaller con-

stituents. So, we will start with sentences, proceed to clauses, move on to phrases, and end up with words.

A **sentence** is a string of words that is grammatically complete with at least two components, a subject and a predicate.

Constituents are the units being combined to create larger syntactic constructions.

The **subject** of a sentence is the topic of the sentence.

The **predicate** of a sentence in traditional grammar is a comment or assertion made about the topic. In most modern grammars, the predicate is seen as an element that assigns a property to another element in a sentence or helps to relate other elements to each

other, thereby completing the meaning of the predicate.

A **sentence** begins as a mental construction job. Sentences are not randomly combined morphemes but structures built on the basis of rules of combination. The units being combined are called **constituents**. In traditional approaches to grammar, a sentence is seen as having at least two main constituents; one is called a **subject**, and the other is called a **predicate**. In these traditional approaches the subject is the topic of the sentence and the predicate is a comment or assertion made about the topic.

In the sentence below, if analyzed by a traditional approach, *The student* is the subject of the sentence and *looked at a painting* is the predicate.

The student looked at a painting.

Arguments are necessary elements of a sentence used to complete the meaning of the predicate.

Adjuncts are optional elements of a sentence. They add information that is not essential to the meaning of the predicate.

In most modern syntactic analyses, the predicate is seen as an element that assigns a property to another element or elements in a sentence, or helps to relate those other elements to each other, thereby completing the meaning of the predicate. Using this rationale, the sentence above might be analyzed as: *looked (the student, at a painting)*. *Looked* is the predicate and the expressions in parentheses—*the student* and *at a painting*—complete the meaning of the predicate. The expressions *the student* and *at a painting* are called **arguments**. Arguments are necessary to complete the meaning of the predicate. *Looked* by itself is not a complete thought; nor is *the student looked*, or *looked at a painting*. The predicate in this case needs two arguments to complete it, a subject argument and prepositional phrase argument. Depending on the type of verb, one or more argument might be needed (obligatory). In addition to arguments that are obligatory, a predicate may be completed by elements, called **adjuncts**, which are optional. For instance, in the sentence

The art student looked at a very beautiful painting, the elements *art*, *very*, and *beautiful* are adjuncts. Adjuncts add information that is not essential to the meaning of the predicate.

Types of sentences and clauses

A **simple sentence** is a sentence with one subject and one predicate.

Using a traditional approach, sentences can be classified on the basis of how many subjects (topics) and how many predicates they contain and the types of clauses they possess. When a sentence consists of only one subject and one predicate, it is called a **simple sentence**. An example of a simple sentence is:

The dog ran away.

A **compound sentence** is made up of at least two simple sentences joined by a coordinating conjunc-

tion; in writing, punctuation can substitute for the conjunction.

Simple sentences can be combined to form **compound sentences**, such as: *The dog and the cat ran away*.

In this case, two sentences are combined using the coordinating conjunction *and*. The compound sentence tells us that *The dog ran away* and *The cat ran away*. Redundant elements are eliminated in forming this compound sentence. Compound sentences can be formed without a coordinating conjunction, as in the following sentence:

We studied all day for the test; now it is time to rest.

An **independent clause** is a simple sentence.

In this case, the semicolon takes the place of the conjunction. The two simple sentences in a compound sentence are said to be **independent clauses**.

A **dependent clause** has a subject and predicate but cannot stand alone as a simple sentence. It depends on an independent clause to make it complete.

A second type of clause used to construct sentences is the **dependent clause**. A dependent clause cannot stand alone as a simple sentence, but must be attached to an independent clause. A dependent clause often begins with a relative pronoun or a subordinating conjunction. Some examples of dependent clauses are the following:

although it is tempting
who would be traveling with us
if I come late

A **complex sentence** contains a simple sentence and one or more dependent clauses.

A sentence that contains a simple sentence and one or more dependent clauses is called a **complex sentence**. The following are complex sentences:

Although it is tempting, I will not be going to Las Vegas.
These are the people who would be traveling with us.
If I come late, start without me.

A **compound-complex sentence** has two or more independent clauses and at least one dependent clause.

Sentences that have two or more independent clauses and at least one dependent clause are called **compound-complex sentences**. For example:

When the teacher assigned the reading for the exam, many students were stunned, but they agreed to study.

When the teacher assigned the reading for the exam is a dependent clause; *many students were stunned* could stand alone as a simple sentence and is, therefore, an independent clause of the larger sentence. The same is true of *they agreed to study*. This independent clause is attached to the rest of the sentence by the coordinating conjunction *but*.

The terms *simple sentence*, *compound sentence*, *complex sentence*, and *compound-complex sentence* refer to the grammatical construction of a sentence. Sentences can also be classified on the basis of their meaning, purpose

(function), or voice. The following are some of the most common sentence types classified in these ways:

Declarative—Sentences that make a statement. *Christine just arrived.*

Interrogative—Sentences that ask a question. *Has Andrew just arrived?*

Imperative—Sentences that express a command or make a request. *Aaron, come here.*

Exclamatory—Sentences that show strong or sudden feeling. *Oh, if Jan were only here!*

Active or passive voice—Voice is the relationship of the grammatical subject of a verb to the action conveyed by that verb. In most English sentences, the grammatical subject precedes the verb. In an active sentence, the grammatical subject of the verb carries out an activity or purpose, as in the sentence *Mark hit the ball*. In the passive version of this sentence, the subject is receiving the action of the verb. So, in the above example, what was the direct object becomes the grammatical subject and what was the grammatical subject is moved to the position of the object. The result is *The ball was hit by Mark*. Note that the word *by* and an auxiliary verb *was* are added in this passive construction. Although the word *by* often indicates a passive construction, it does not have to be present in a passive sentence. The sentence *The computer was*

purchased yesterday is also passive. In this sentence the subject (*I, we, a person's name, etc.*) is missing altogether. A possible active version of the sentence would be *I purchased the computer yesterday*. In these two examples, the verbs *hit* and *purchased* are in the passive voice (see [Box 5-1](#)).

In addition to these types of sentences, various combinations of types can be formed. *Don't be hit by a ball* is a passive, imperative sentence.

Box 5-1 The passive voice

English teachers often tell students to avoid using the passive voice that is formed in the ways described in the text and by some uses of *to be* words, including *am, is, are, and were*. There is good reason for this. The passive voice can obscure who is doing what to whom or who is responsible for what. In some cases, *to be* words can be used to eliminate the person responsible for the action completely, as in: *The credit card payment will be made on the 15th day of the month*. This sentence does not state who is to make the payment. The reader or listener might assume that the person whose name is on the credit card is responsible for the payment. However, if this sentence were part of a legal document with several

parties, someone other than the cardholder might be responsible for the payment, such as the company that employs the card owner. The sentence *The card owner will make a credit card payment on the 15th day of the month* clarifies who is to make the payment.

The passive voice is also wordier, using more nouns and prepositional phrases. The following is a passive sentence:

Analysis and assessment of the quality of instruction by college presidents and deans is required so that suggestions for changes and improvements in instruction can be made.

The active version of the sentence is less wordy and less ambiguous:

College presidents and deans must analyze and assess the quality of instruction so that they can make suggestions for improving instruction.

The passive voice does have a place in writing and speech. It can be used to add variety to an utterance as long as it does not obscure meaning.

Phrases

A **phrase** is any constituent of a clause.

A **phrase** is any constituent of a clause. Phrases are commonly named for one of their main elements, so we speak of noun phrases, verb phrases, adjective phrases, adverb phrases, and prepositional phrases. A phrase may be a string of words or just one word. In the following sentence there are several phrases: *Jack went to the store*. *Jack* is a phrase, and so are *went to the store*, *to the store*, and *the store*. Notice that not only can a phrase be one word or a string of words, but that one phrase also can be embedded within another phrase.

The **head of a phrase** is the word that determines the syntactic or phrasal category of that phrase.

The **head of a phrase** is the word that determines the syntactic or phrasal category of that phrase—whether the phrase functions as a noun phrase, verb phrase, prepositional phrase, and so on. (Remember that in [Chapter 4](#) we discussed the head of a compound word, which is

the morpheme that determines the lexical category of the word; for example, whether the compound word is a noun or a verb.) The head of a noun phrase is a noun, the head of a verb phrase is a verb, and the head of a prepositional phrase is a preposition. If the phrase is made up of one word, then that word is the head of the phrase. If a phrase has two or more words in the lexical category that the phrase is named for, then the one that carries the central meaning of the phrase is the head of the phrase. In the noun phrase *the boat*, it is clear that *boat* is the head of the phrase. However, in the noun phrase *the title of the new movie*, there are two nouns, *title* and *movie*. Because the phrase is about the *title* of the movie and not about the *movie* itself, the head of the phrase is *title*.

The **dependent** or **dependents of a phrase** are all parts of a phrase that are not its head.

A **specifier** makes the meaning of the head more precise.

Complements provide further information about the head.

A **determiner** is a word used before a noun to indicate whether the noun refers to something that is specific or general.

All parts of a phrase that are not the head are called the phrase's **dependents**. In some approaches to syntax, these dependents are further broken down into **specifiers** and **complements**. In *the boat*, *the* is the specifier. In *the title of the new movie*, *the* is the specifier and *of the new movie* is the complement. The specifier makes the meaning of the head more precise. **Determiners** are specifiers for nouns, adverbs are specifiers for verbs, and degree words such as *very* and *more* are used as specifiers of adjectives and prepositions. Complements provide further information about the head. The phrase *the new movie* indicates the title is that of a movie as opposed to a book, a magazine, or play.

Some languages, including Spanish, French, Tiwi (an indigenous Australian language), and English, tend to place complements to the right of the head (head-first or right-branching languages). Other languages, such as Turkish, Korean, and Japanese, tend to put complements to the left of the head (head-last or left-branching languages), with Japanese doing this almost exclusively.

A **noun phrase** (often called a **nominal phrase**) does the work of a noun.

Noun phrases Among other functions, a **noun phrase** (often called a **nominal phrase**) can function in a sentence as the subject, direct object, and indirect object. A noun phrase could be a single noun or pronoun or a variety of longer forms:

1. Julian mailed a letter. (*Julian* is a noun phrase and the subject of the sentence; *a letter* is also a noun phrase and the direct object.)
2. Mary ate the hamburger. (*Mary* and *the hamburger* are noun phrases. *Mary* is the subject of the sentence; *the hamburger* is the direct object.)

3. Three people came late.

(*Three people* is the noun phrase and the subject of the sentence.)

4. The girl went into the house.

(*The girl* and *the house* are noun phrases; *the house* is an indirect object.)

5. He gave the card to me.

(*He* is a noun phrase, as is *me* and *the card*; *me* is an indirect object; *the card* is a direct object.)

Noun phrases can be abbreviated as NP. A noun might be preceded by an adjective or adjective phrase. The adjective phrase might include an adverb (*very fast horse*) or a subtype of adjective called a determiner (abbreviated as Det). In English, determiners fall into the following categories: definite and indefinite articles, demonstratives, possessives, and interrogatives. Determiners function to limit what the noun is referring to, such as to specify whether the referent is a specific thing or a general thing. Articles—*a(n)* and *the* (abbreviated as Art)—tell whether a noun refers to a definite (specific) thing, as in *the art*

student, or something that is not specified (a general thing), as in *a very beautiful painting*. In the phrase *the art student*, *the* is a definite article. *The art student* refers to a specific student to whom we might give a name. In the phrase *a very beautiful painting*, *a(n)* is an indefinite article because the phrase does not specify exactly what painting is being described. In English, articles are placed before the noun that they modify. They share this characteristic with demonstratives (*this* boy), possessives (*my* car), and interrogatives (*which* house).

Some possible noun phrases are as follows:

- a. Jim NP → N (N is the abbreviation for noun if it cannot be broken down further.)
- b. he NP → Pro (Pro = pronoun)
- c. the dog NP → Det N (Det = determiner, which in this case is an article.)
- d. six dogs NP → Num Noun (Num = numeral)
- e. the six dogs NP → Det Num Noun
- f. my dog NP → Det N (This determiner is a possessive.)
- g. what dog NP → Det N (This determiner is an interrogative.)
- h. that dog NP → Det N (This determiner is a demonstrative.)

The → in the formulas above means *can be rewritten as* or *can be expanded as* or *is made up of*. So in example (e), the formula reads that the noun phrase can be rewritten as a determiner plus a numeral and a noun.

A **verb phrase** tells you something about the subject. It includes a verb and can include an auxiliary verb, a direct or indirect object, and modifiers.

Verb phrases All English sentences (sentence is abbreviated as S) contain a noun phrase (NP) and a **verb phrase** (VP); that is, an English sentence is minimally as follows:

$$S \rightarrow NP + VP$$

Intransitive verbs can form a verb phrase by themselves. In the simple sentence, *Fish swim*,

swim is a verb phrase composed of just a verb (VP → V).

Verb phrases often include a noun phrase. Verbs that combine with a noun phrase are called transitive verbs. In the sentence

Mary ate the hamburger,

ate the hamburger is the verb phrase. It can be written as VP → V NP; *the hamburger* is the noun phrase within the verb phrase. All of the categories of verbs described in [Table 4-1](#) (Lexical Categories) can form verb phrases.

An **adjective phrase** is headed by an adjective, but might also include an adjective **modifier** (an element that adds a property to another lexical item). Adjective phrases modify nouns.

An **adverb phrase** is a modifier of a verb.

A **prepositional phrase** is a phrase headed by a preposition. It can function to modify a noun phrase or a verb phrase.

Other types of phrases In addition to noun phrases and verb phrases, other important phrasal categories are **adjective phrases** (AdjP), **adverb phrases** (AdvP), and **pre-**

positional phrases (PP). Adjective phrases are headed by an adjective, but might also include adjective **modifiers** (elements that add a property to another lexical item). Adjective phrases in turn modify nouns. Adverb phrases are headed by an adverb and might also include other adverbs and an adjective phrase or phrases. Adverb phrases modify verbs in the following ways:

1. frequency (They came *every* day.)
2. duration (The students have been coming *for the past five days*.)
3. time (Tim will be here *at 3 o'clock*.)
4. manner (You should do it *this way*.)
5. purpose (Christopher brought his report card home *to show it to his father*.)

Prepositional phrases are headed by a preposition and include a noun phrase. Both adjective and adverb phrases can use prepositions. The question becomes should the phrase with the preposition be called a prepositional phrase or an adverb or adjective phrase? Consider the following sentence:

The farmer from Iowa is going into the store.

There are two prepositional phrases in this sentence: *from Iowa* and *into the store*. The function of *from Iowa* is to

modify *the farmer*; it tells you where he is from. It is an adjective phrase, but because it is also a prepositional phrase some linguists and grammar teachers would call it an adjectival prepositional phrase. The phrase *into the store* is an adverb phrase (or an adverbial prepositional phrase). It modifies the verb by telling us where the farmer went.

Exercise 1 Syntactic construction

1. Label the subject and predicate of the following sentences.

Example: The black cat/ate all of the cat food.
 subject predicate

- a. I am going to the store.
 - b. The clown amused us.
 - c. Is this the place?
 - d. Come here.
 - e. We were amused by the clown.
2. Did you have any problems with d. and e. of Question 1? Explain.

3. Is there an analogous relationship between the concept of free and bound morphemes and independent and dependent clauses? Explain.

4. Determine which of the sentences listed below is simple, compound, complex, or compound-complex.

a. Who is at the door?

- b. We will be at the restaurant in twenty minutes. _____
- c. The children who came to the party are all from the same school. _____
- d. I have eaten two pies, yet my desire for sweets has not been satisfied. _____
- e. He walked as if someone was following him. _____
- f. We must find a teacher who understands our needs. _____
- g. The score was thirty-six to nothing; obviously there was little hope that the home team would win. _____
- h. All of the people enjoyed the concert and the dinner that followed it. _____

5. Rewrite the following noun phrases in terms of abbreviations and arrows.

Example: a bright color NP $\rightarrow$ Det Adj N

- a. the beautiful furniture

- b. a cow _____
- c. the most educated people

- d. six pens _____
- e. those pens _____

6. Mark all of the noun phrases in each of the following sentences and determine their function in the sentence.

Example: A few people came into the movie studio. NP—**subject** **Indirect object**

- a. Jill's house went on the market today.
- b. All guns are bad.
- c. It was a good thing that Shane came for dinner.
- d. Go home.
- e. Large cars require more gas.

7. Formulas such as NP $\rightarrow$ Det N express rules. This one simply says that a noun phrase can be a determiner plus a noun. However, this rule is somewhat too general; not all nouns can fol-

low a determiner. Can you determine which type or types of nouns do not fit this rule?

8. Mark all of the verb phrases in the following sentences.

Example: He photographed the flowers.
VP

- a. The dog ran after the car.
- b. Jack died.
- c. He has taken five tests.

9. Describe each of the verb phrases in Question 8 in terms of a formula.

Example: He photographed the flower. VP → V NP

10. Underline all of the adjective phrases in the following sentences.

- a. The blue ball rolled away.
- b. That is a really fat yellow cat.
- c. The candidate was quite upset at the reception he received.

11. Underline all of the adverb phrases in the following sentences.

- a. He arrived at noon.
- b. She usually gets up early.
- c. The farmer harvested the corn with a machine.
- d. We are going to take a vacation before the prices go up.
- e. The teachers all showed up to support the students.

12. In the previous question, identify the adjectival and adverbial prepositional phrases.

The constituent structure of sentences

Using an analogy, we can say that, similar to a sentence, a car is made up of constituents or parts. The largest part is the car itself, and can be compared to a sentence. The smallest parts of the car are individual pieces of metal, rubber, glass, and plastic, and for the purposes of this comparison could be compared to individual words (words of course are made of even smaller units—morphemes and phonemes). The small parts make up the whole but function differently than any of the parts alone. Between the whole and the individual parts are various assemblies of parts that go together. There is the brake assembly, which we could say is comparable to the subject noun phrase of a sentence; the front windshield assembly is comparable to the verb phrase; the steering wheel assembly is comparable to the object noun phrase; the engine assembly is comparable to a prepositional phrase, and so on. Each of these assemblies has a specific function, and the proper combination of individual parts makes the function possible. Furthermore, an individual

part from the brake assembly coupled with a part from the engine assembly would have no function and would not be a part (constituent) of the car, just as an individual part from a subject noun phrase and an individual part from the prepositional phrase would have no function and would not be a constituent of the sentence.

So, we can see a sentence as made up of small meaningful units (words). These units combine to make large units, and then these larger units combine into even larger ones until we have the entire sentence as the largest constituent of itself. Consider the following sentence:

The art student will look at a very beautiful painting.

Each word has meaning and a specific function in the sentence. So do the groupings:

- art student
- beautiful painting
- very beautiful painting
- at a very beautiful painting

Note that other groupings do not have a coherent meaning or function relative to the entire sentence. Groups like

- the art
- a very

- will look at a
- at a very

do not function as meaningful units for the sentence.

Why is *a very beautiful painting* a constituent of the sentence and *at a very* is not? The segment *a very beautiful painting* makes sense on its own as well as having meaning in the sentence. We can ask the question

What did the art student look at?

And we could answer

... a very beautiful painting.

There is no question that we can pose that will have the answer *at a very*. This is because *at a very* has no meaning in itself and is not a proper “assembly” relevant to the sentence under examination. Another way of saying this is that *a very beautiful painting* could be given a label as to its function in the sentence and *at a very* could not.

The manner in which constituents of a sentence are arranged is related to the meaning of the sentence. However, constituents can be arranged ungrammatically and the sentence might still have the same meaning as it would have had if the constituents had been arranged grammatically. A native speaker of English will recognize

that the following sentence is ungrammatical, but might still understand what it means:

Will look the student at paintings very beautiful.

Also, a sentence might be well-formed (grammatical) but not have any meaning, as exemplified by the “Jabberwocky” example on page 128.

Labeling the constituents of a sentence

The words The smallest constituents of a sentence are the morphemes that make it up. Morphemes make up words. Each word can be labeled as to its part of speech or the subtype of its part of speech (lexical category) (see [Table 4-1](#), Lexical Categories). For an example, examine the sentence we have been considering:

The	art	student	will	look	at	a	very	beautiful	painting.
Det	Adj	N	Aux	Verb	Prep	Det	Adv	Adj	N

A **lexicon** for a specific language is a list of all the morphemes that are used in that language to form words.

A **co-occurrence restriction** is a limitation on the use of a morpheme.

The lexicon A **lexicon** for a specific language is a list of all of the morphemes that are used in that language to form words. A dictionary often lists only the words. Each morpheme in a lexicon is accompanied by a set of specifications. These specifications include information on the meaning, pronunciation, and various other grammatical features of each morpheme. The lexicon specifies whether each lexical entry (each morpheme in the lexicon) is a prefix, suffix, or root. If it is a root, then the lexical category is also included. Any **co-occurrence restrictions** are also mentioned by labeling each root as to the subtype of the category, such as whether a verb is transitive or intransitive. A co-occurrence restriction is a limitation on the use of a morpheme. For instance, a transitive verb is limited to sentences that have a direct object.

**The boy threw.*

This sentence is ungrammatical because *threw* must co-occur with an object, as in *The boy threw the ball*.

Labeling phrases

We have already discussed the labeling of phrases. A phrase can be labeled as a noun phrase, verb phrase, adjective phrase, adverb phrase, prepositional phrase, and others. Consider the sentence under question in [Figure 5-1](#).

A **tree diagram** is an illustration in the form of an upside-down tree shape that shows the constituents of an utterance, with the most general at the top and more specific constituents at the bottom of the tree.

A **phrase marker** or a **phrase structure tree** is a tree diagram that specifies the function of each constituent of an utterance.

The phrase constituent structure, along with labels of each word, can also be represented in a **tree diagram**. A tree diagram that specifies the function of each constituent

ent is called a **phrase marker** or a **phrase structure tree**. Examine the phrase marker produced in [Figure 5-2](#).

A **node** is a point in a tree diagram where branching occurs.

Each point at which branching occurs is called a **node**. Notice that tree diagrams are upside down. What should be the root is at the top, and what would be the top is represented by the most specific constituents (the individual words).

[Figure 5-1](#) and the tree diagram in [Figure 5-2](#) represent the hierarchical structure of language. That is, one constituent is often a constituent of a higher level or is a dominant constituent; and all constituents are part of the highest-level or most dominant constituent, the sentence itself.

We said earlier in the chapter that an English sentence must contain a NP and a VP. You will see in [Figure 5-2](#) that there are three primary constituents or heads of the sentence being diagramed, not two; a noun phrase, a verb phrase, and an auxiliary. Although in traditional grammars auxiliaries are seen simply as part of the verb, most contemporary syntacticians see auxiliaries as one of the

heads of a sentence. Auxiliaries are seen as a separate mental construct. The auxiliary says something about the time frame (that is, tense) of the action or state of the subject noun phrase; for instance, the auxiliary *will* indicates future tense. The auxiliary *might* indicates a possibility of an action or state, exemplified by the word *may* (see [Table 4-1](#) under verbs for a further explanation of auxiliaries and their many functions).

The art student / will look at a very beautiful painting.
noun phrase verb phrase

/at a very beautiful painting.
prepositional phrase

/a very beautiful painting.
noun phrase

FIGURE 5-1 The constituent structure of a sentence

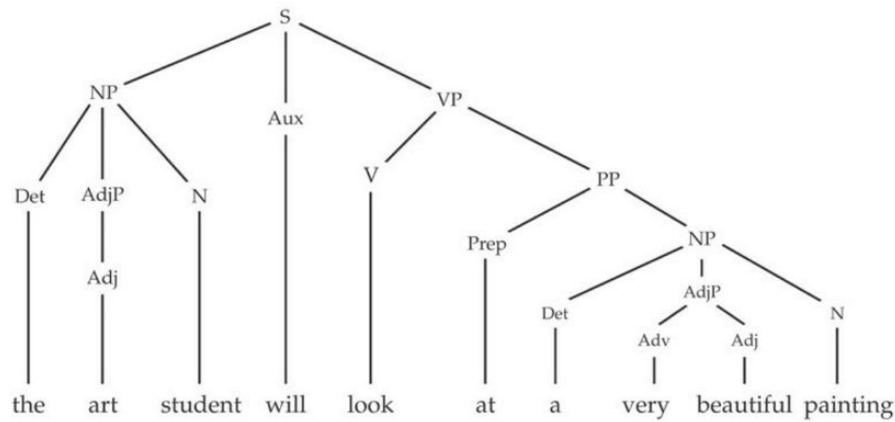


FIGURE 5-2 Phrase marker

Notice that the auxiliary verb *will* indicates the future tense. However, tense is indicated by an inflectional bound morpheme for some past tense verbs, such as in the sentence: *The art student looked at some very beautiful paintings*. In this case, the auxiliary is considered part of the form of the main verb, but Aux is often still put in the phrase marker even though there is no separate word. The inflectional morpheme *-ed* converts the verb to the past tense. This also applies to irregular verbal forms such as *went*. Even though there is no physical inflectional morpheme indicating past and no physical auxiliary verb, the concept *past* is included in the word to the same extent as if there was an *-ed* added to the verb or an auxiliary separate from the verb. Some linguists use the

symbol *I* for inflection, instead of *Aux* to represent tense in a phrase marker, but we will continue to use *Aux*. See page 119 for an example of a phrase marker with the past tense marker *-ed*. Auxiliaries can indicate things other than tense, such as aspect, modality, or voice.

Exercise 2 Constituents and phrase markers

1. The smallest constituents of a sentence are individual words. Label the word type of each word in the following sentences:
 - a. Fluent speakers have an enormous subconscious knowledge of their language.
 - b. The rabbit quickly jumped into the big hole.
 - c. We may visit our good friend.
 - d. The boy will do the homework.
2. Draw a phrase marker for each of the sentences above. (See phrase marker on page 119 to see how you should handle the past tense.)

Phrase structure rules

Phrase structure rules specify how constituents of an utterance are arranged and what constituents can occur as parts of other constituents (the hierarchical structure of a sentence).

We have already said that statements such as $S \rightarrow NP \text{ Aux VP}$ are rules. Because these rules have to do with how constituents are arranged and what constituents can occur as parts of other constituents (the hierarchical structure of a sentence), they are called **phrase structure rules**. A sentence can be described by listing a series of phrase structure rules starting with the most general (the top of a phrase marker) and ending with the most specific rules (the bottom of the phrase marker). The sentence

The art student will look at a very beautiful painting

can be represented in phrase structure rules as follows (see the phrase markers in [Figure 5-2](#)):

Sentence A: *The art student will look at a very beautiful painting.*

$S \rightarrow NP \text{ Aux VP}$

$NP \rightarrow \text{Det AdjP N}$

$\text{AdjP} \rightarrow (\text{Adv}) \text{Adj}$

$\text{Aux} \rightarrow \text{Tense}$

$\text{VP} \rightarrow \text{V PP}$

$\text{PP} \rightarrow \text{Prep NP}$

Other sentences may fit these same set of phrase structure rules or may be described by other rules. For instance, the following phrase structure rules describe this sentence: In the phrase structure rules here, *Noun* is abbreviated as *N* in some places and spelled out in others. It is spelled out as *Noun* when the noun can be broken down further. In the example above, the word *paintings* can be broken down into $\text{Noun} \rightarrow \text{N pl}$. The pl stands for *plural*. In general, if a component of a phrase can be broken down further, a longer representation of the component is used. When the component cannot be broken down further, we use a shorter representation.

Sentence B: *He will look at some beautiful paintings.*

$S \rightarrow NP \text{ Aux VP}$

$NP \rightarrow \left\{ \begin{array}{l} \text{Pro} \\ \text{AdjP Noun} \end{array} \right\}$

$\text{AdjP} \rightarrow \text{Adj Adj}$

$\text{Noun} \rightarrow \text{N pl}$

$\text{Aux} \rightarrow \text{Tense}$

$\text{VP} \rightarrow \text{V PP}$

$\text{PP} \rightarrow \text{Prep NP}$

Sentence B differs from sentence A in the first noun phrase. In sentence B, *the art student* is replaced by the pronoun *he*. In the phrase structure rules for sentence B, braces are used for the NP. Braces mean *either-or*. A noun phrase in sentence B can be a pronoun or an adjective phrase and a noun; in this case, two adjectives plus a noun. An individual noun phrase cannot include both a pronoun and two adjectives and a noun. The parentheses around a constituent mean that it is optional. For example, the line $\text{AdjP} \rightarrow (\text{Adv}) \text{Adj}$ in the phrase structure rules for sentence A means that of the two adjective phrases in the sentence, one includes an adverb and the other does not.

We can now combine the rules for sentences A and B.

The combined rules will describe both sentences. These are the combined rules: We could write the rules for a third, fourth, fifth sentence, and so on and each time incorporate the individual rules of each sentence into a more general set of rules. If, at the end of this procedure, we combined the rules of every type of English sentence into a general set of rules, we would have a complete grammar of the syntactic component of English. Such a grammar would be called a generative grammar. A generative grammar is a finite set of rules that could hypothetically produce an infinite number of utterances. It would enable us to generate all sentences an English speaker could produce. And it would never produce an ungrammatical sentence. The production of infinite utterances is made possible by the recursive property of language. Recursion allows one type of syntactic structure to be included inside another structure of the same type (such as a noun phrase) to create infinitely long sentences or an infinite number of different sentences (see [Box 5-2](#)).

$$S \rightarrow NP \text{ Aux VP}$$

$$NP \rightarrow \left\{ \begin{array}{l} \text{Pro} \\ (\text{Det}) \text{ AdjP Noun} \end{array} \right\}$$

$$\text{AdjP} \rightarrow (\text{Adv}) (\text{Adj}) \text{ Adj}$$

$$\text{Noun} \rightarrow \text{N (pl)}$$

$$\text{Aux} \rightarrow \text{Tense}$$

$$\text{VP} \rightarrow \text{V PP}$$

$$\text{PP} \rightarrow \text{Prep NP}$$

Box 5-2 Recursion in language

The behavioralists believed that a child acquires language by learning a limited number of representative sentences and then producing new sentences based on the pattern of the representative sentences. But this does not account for productivity in everyday language, let alone poetry and other creative speech and writing. Noam Chomsky and his colleagues have demonstrated one process that generates productivity; it is called *recursion*. Recursive rules allow, for instance, a noun phrase to be made up of other noun phrases with connecting elements. So we can say in this utterance:

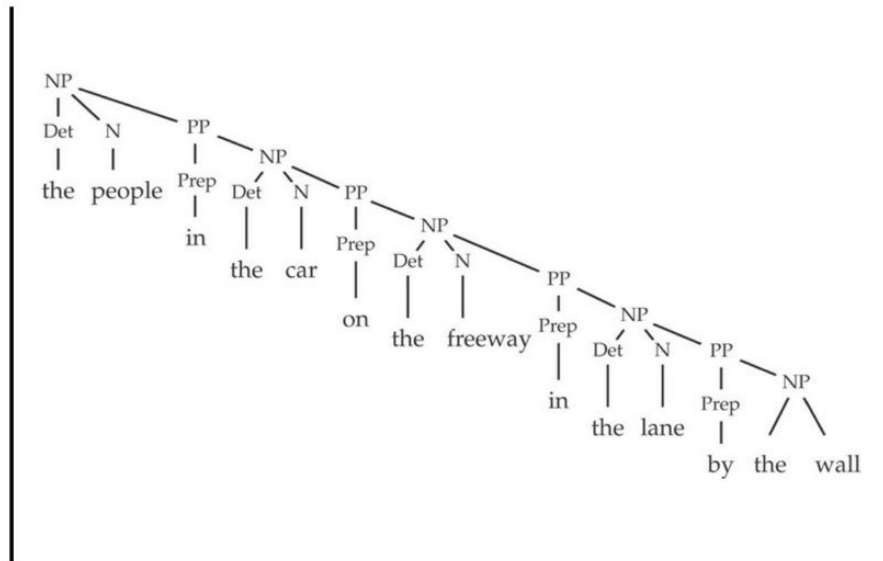
the boy that chased the dog that chased the cat that chased the squirrel that chased the bird that chased the bug.

In this way we can create an “infinitely” long sentence.

Phrase structure rules are recursive rules. For instance, the phrase structure rule $\text{PP} \rightarrow \text{Prep NP}$ is recursive in that it can generate all the prepositional phrases in the sentence:

The people in the car on the freeway in the lane by the wall drove home.

The rule $\text{PP} \rightarrow \text{Prep NP}$ can be applied over and over again in the same sentence, leading to a more and more complex sentence, as illustrated in the diagram below that provides a phrase marker only for the application of the rule $\text{PP} \rightarrow \text{Prep NP}$, and does not diagram the entire sentence (the VP is not included).



There are several different approaches to generative grammar. Depending on the approach, phrase structure rules might be written quite differently. In addition, we have only written rules for simple declarative sentences ($S \rightarrow NP \text{ Aux VP}$). The sentence below is a yes/no interrogative sentence:

Will you come here?

The general rule for this type of sentences would be $S \rightarrow \text{Aux NP VP}$.

We have not covered phrase structure rules or phrase markers for compound, complex, or compound-complex sentences or for other classes of sentences. We leave that

and the various notational systems used to represent sentences as well as other syntactic structures, and the concepts behind these notational systems, to a course specifically on the topic of syntax.

Noam Chomsky and generative grammar and beyond

Noam Chomsky, born in 1928, has been perhaps the best-known linguist in the world for more than 50 years, and his influence goes well beyond linguistics. In 1957, he began to revolutionize the study of language with his book *Syntactic Structures*. His ideas were highly influenced by his interest in logic and mathematics. In the 1950s, Chomsky broke with the dominant school of thought in linguistics, the structural approach. Leonard Bloomfield (1887–1949) and others championed the structural approach in linguistics. Bloomfield was one of the best-known American linguists of the first part of the twentieth century; the type of structuralism he developed was descriptive and is often referred to as Bloomfieldian linguistics. His approach started with describing and classifying sounds and then morphemes in terms of their function. From principles developed from the study of phonology and morphology, more abstract

units would be studied, with syntax the ending point of the analysis. Bloomfield studied meaning (semantics), but most other Bloomfieldians thought that semantics was too abstract to be studied in any verifiable (empirical) way, and therefore ignored it.



Noam Chomsky. dpa picture alliance/Alamy Stock Photo

With *Syntactic Structures*, Chomsky began to change the Bloomfieldian consensus. In Chomsky's view, language learning is motivated by an internal capacity to acquire language. This capacity evolved, as hominins evolved, into a universal innate human ability to learn and analyze linguistic information. (Some anthropologists use the term

Universal grammar (UG) is the system involving phonemic differences, word order, and phrase recognition that is the basis for the theory of the innateness of language acquisition.

hominids instead of *hominins*.) This universal grammar provides the general rules that allow us, at least as children, to learn any language, even with minimum input from the environment. **Universal grammar (UG)** is a general blueprint that permits the child to proceed from the general rules of all languages to the rules specific to his or her own language. Chomsky believes that language learning is guided by an innate language acquisition device that is a result of human evolution. We will discuss this idea further in [Chapter 10](#) along with alternative ideas.

Surface structure refers to an actual utterance that can be broken down by conventional methods of syntactic analysis.

Deep structure refers to a highly abstract level of language that represents the basic meaning of a sentence.

Bloomfieldian linguistics emphasized linguistic performance, what the speaker actually says, and what some linguists called the **surface structure**. At first, Chomsky emphasized linguistic competence, what the speaker subconsciously knows about his or her language, and what might be called the **deep structure**. Early Chomskian linguistics dealt with how the deep structure is transformed into the surface structure and how an infinite number of utterances can be generated from a finite number of rules and lexical items. For this reason, the Chomskian approach has been called a mentalist approach in which the subconscious knowledge of the native speakers of a language is emphasized. This is in contrast to the Bloomfieldian school, which emphasizes what is called the mechanistic approach; a rigid set of learned rules is used to form grammatical utterances. The mechanistic approach cannot explain many things that are explained by the mentalist approach. One of these things is productivity. The Bloomfieldian idea that language is learned

by mimicry does not account for the fact that young children can produce utterances that they have never produced or heard before from a finite number of words. The mentalists' approach postulates that the human mind is like a software program designed to generate new sentences on demand.

In the 1980s, Chomsky introduced the principles and parameters theory. One problem with the concept of universal grammar is that languages vary so much in their surface structure. Or do they? The principles and parameters theory postulates that even though there is variation in languages, these differences have specific principles and parameters (limitations). For instance, modifiers, such as adjectives and adverbs, can come before or after the thing that they modify, but not several words away from what is being modified. So the language acquisition device is "programmed" to allow certain possibilities and not others. Children have to learn from the speaking environment which possibility fits their language.

In 1995, Chomsky further modified his concepts with the formulation of what is called minimalism (or the minimalist program). In minimalism, Chomsky maintains the concept that there is one single grammatical system for all languages, but eliminates the concepts of deep and surface structure as well as other features of

earlier conceptualizations of syntax. Explanations of the details of this are beyond the scope of this text. There have always been competing theories of syntax; minimalism is quite controversial. See Chomsky's 1995 book to learn more; and Pieter A.M. Seuren's book offers critical analysis (see "Suggested reading" at the end of this chapter).

Although there are competing theories of syntax, Chomsky's ideas have been and continue to be very influential. One point of disagreement is that not all linguists, psychologists, or biologists believe that there is a language acquisition device in a physical sense (see [Chapter 10](#)). However, the discovery of the FoxP2 gene, which, if "defective," causes problems with the acquisition of language, points toward a genetic potential for the acquisition of language (see [Chapter 1](#)).

Exercise 3 Phrase structure rules

1. What are phrase structure rules?

2. Write the phrase structure rules for all of the sentences in [Exercise 1](#), Question 8.

Example: He photographed the flowers.

$S \rightarrow NP \text{ Aux VP}$

$NP \rightarrow \{\text{Pro Det Noun}\}$

$\text{Noun} \rightarrow N \text{ pl}$

$\text{Aux} \rightarrow \text{Tense}$

$VP \rightarrow V NP$

3. Write the phrase structure rules for the following sentences, and then draw the phrase markers (tree diagrams) for each.

a. You may sit on the fence:.

b. Six boys are playing quietly.

c. The bad dog will bite the man.

4. Now combine the rules of individual sentences a., b., and c. into one set of general rules.

Transformational rules

Consider the following sets of sentences:

A1 The boy passed out the candy.

A2 The boy passed the candy out.

B1 Linguists often use large words.

B2 Large words are often used by linguists.

Each of the sets contains synonymous sentences. Yet the forms of sentences 1 and 2 of each set are different. The phrase markers used to represent these synonymous sentences would be different.

Transformational rules (T-rules) relate the spoken form of sentences (surface structure) to their underlying meaning (deep structure).

Transformational rules (T-rules) relate the spoken form of sentences to their underlying meaning. More

technically, transformational rules relate the surface structure of sentences to their deep structure. There are many hypotheses about the importance or even the existence of deep and surface structure, as well as the importance and types of transformations. This section is included to provide background into how these ideas were originally used. For detailed discussions of the various concepts of syntax, see Andrew Carnie's book, listed in "Suggested reading."

Surface structure refers to an actual utterance that can be understood and broken down by traditional syntactic analysis. The deep structure was proposed in early generative hypotheses as a highly abstract level that represents the basic meaning of an utterance. Different surface structures may have the same deep structure, or different deep structures may have the same surface structure.

We intuitively judge sentences A1 and A2 to have the same deep structure. Yet the sentences take slightly different forms.

A **phrasal verb** is a verb phrase consisting of a verbal base and a verbal particle. It can have an idiomatic or special meaning.

The **verbal base** is the main part of the verb.

Verbal particles are prepositions that co-occur with some verbs and can appear to the left or right of the direct object noun phrase.

Passed out is an example of certain types of verbal expressions, called **phrasal verbs**, that include a **verbal base**, the main part of the verb, and a **verbal particle** (Prt), in this case a preposition. In the sentence

The boy passed out the candy,

passed out is the verb, *passed* is the verbal base, and *out* is the verbal particle. Verbal particles are part of the verb, yet they can be separated from the verbal base. Sentences A1 and A2 show that the preposition can occur on either side of the direct object noun phrase. However, in other sentences the preposition can be restricted to one side of the direct object noun phrase. In the following example, the verbal particle is restricted to the left side of the direct

object:

C1 Please go over your homework tonight.

C2 *Please go your homework over tonight.

Of the two sentences listed in A, A1 might be judged to be more basic, since the verbal base and verbal particle are together. If we take this as our assumption, then sentence A2 is a transformed version of A1. The rule that relates sentences A1 and A2 to each other is the particle movement transformation. The rule takes this form:

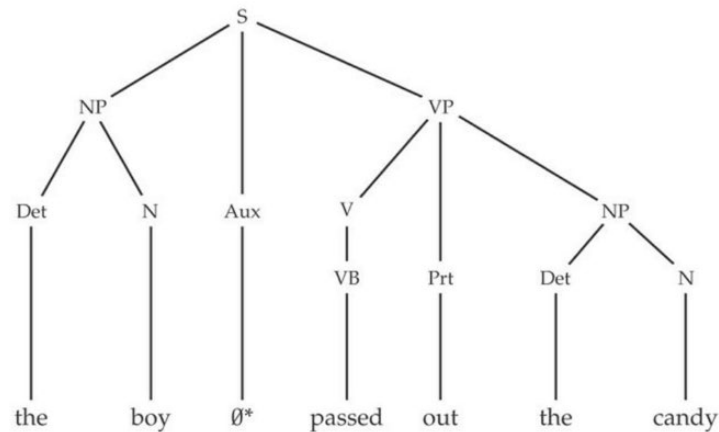
$$X_1 + \text{VB} + \text{Prt} + \text{NP} + X_2 \rightarrow X_1 + \text{VB} + \text{NP} + \text{Prt} + X_2.$$

In this notational system, X_1 is any element to the left of the verbal base (VB) and X_2 is any element to the right of the direct object noun phrase.

We can show the particle movement transformation with phrase markers.

Basic phrase marker

The diagram below shows a basic phrase marker.

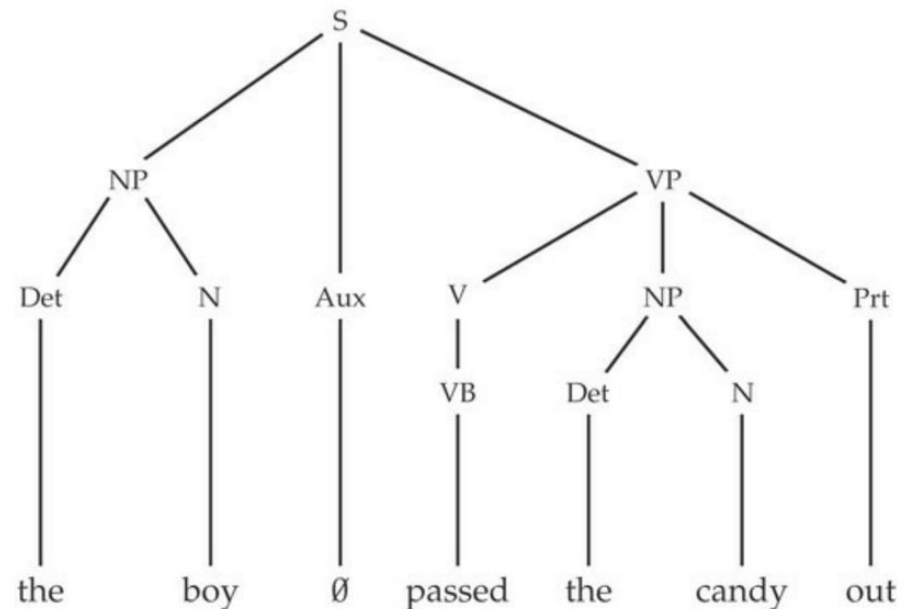


* The Ø means that the auxiliary is not represented by a word such as *will*. Instead the inflectional bound morpheme *-ed* takes the place of a word and indicates the concept *past tense*, just as the word *will* indicates future tense.

Derived phrase marker

A **derived phrase marker** is a phrase marker after transformational rules have been applied.

The phrase marker after the transformational rule has been applied is called the **derived phrase marker**.



Movement transformations The particle movement transformation not only explains the relationship between sentences A1 and A2, but also explains all similar sets. Some other examples of sentences related in this way are the following:

Anthony took off his shoes.

Anthony took his shoes off.

The students passed in their tests.

The students passed their tests in.

The particle movement transformation is only one of many movement transformations in which an element of

the deep structure is moved from its position in the deep structure to another location.

Topicalization is another kind of movement transformation. The topicalization transformation creates a derived sentence with a different focus or emphasis than the basic sentence.

Topicalization is another kind of movement transformation. The topicalization transformation creates a derived sentence with a different focus or emphasis than the basic sentence. The basic sentence

I love Christine

can be transformed to

Christine, I love.

It would be written in abstract as:

$$X_1 + V + X_2 \rightarrow X_2 + X_1 + V$$

Other types of transformations

We have seen that transformational rules relate to sets of sentences that have the same element in different places within the sentences. Transformations explain three other processes in addition to movement rules:

- Deletion
- Insertion
- Substitution.

Deletion transformations A sentence that undergoes transformation must have the same meaning as the sentence from which it was derived. Transformations never change meaning. In the imperative sentence

You come here

the pronoun may be deleted. The derived sentence

Come here

has the same meaning as the basic sentence. This transformation is the imperative transformation.

Redundant elements in the deep structure of the basic sentence may also be deleted. For instance, in the sentence

If Stephen says he will study for the test, he will study for the test

all but the auxiliary of the second verb phrase can be deleted. The result is:

If Stephen says he will study for the test, he will.

This transformation is called the verb phrase deletion rule.

Insertion transformations Words inserted into a basic sentence may not add meaning to the basic sentence. In the sentences

A1 He knew she was here

A2 He knew that she was here

that is inserted in the second sentence. But *that* has no meaning. In this case, the addition of *that* is optional. However, consider the following:

B1 *He won the race is history.

B2 That he won the race is history.

Even though the meaning might be clear, B1 is not a grammatical English sentence. The *that* in B2 is a word inserted to introduce the noun phrase *he won*. Although *he won* is a grammatical sentence, here it is a part of the larger sentence. In sentence B2, *he won* is a NP that is the subject of the sentence. A sentence that is part of another

sentence is called an *embedded sentence*. *So he won* is an embedded sentence, acting as a noun phrase in the larger sentence. The insertion of *that* to form a surface structure sentence is, not surprisingly, called the *that* insertion transformation.

Substitution transformations The only substitution transformations are those that substitute a pronoun for some other part of speech or syntactic category. For instance,

Anthony thought that Anthony was the best

becomes

Anthony thought that he was the best.

This substitution of a pronoun is a pronominalization transformation. Like all transformations, it does not change meaning.

Optional and obligatory transformations

As with phonological rules, transformations can be optional (stylistic) or obligatory. The particle movement transformation, topicalization rule, imperative transformation, verb phrase deletion rule, and pronominalization are all optional rules in English. They may or may not be applied.

In American Sign Language (ASL), the topicalization transformation is obligatory; the object is always signed first. In English, the *that* insertion rule as it applies to embedded sentences is obligatory. It must be applied to the deep structure to render a grammatical surface structure.

Another rule that is obligatory in English is yes/no question formation. Transformational grammarians assume that the deep structure of a yes/no question is similar to that of a declarative sentence, but with an abstract element labeled Q at the beginning of the sentence, as shown below. (The symbol # marks the beginning or end of a sentence.)

Q Aaron will eat his lunch

The form shown here is a deep structure that must undergo transformations to become a question at the surface structure. The transformation simply involves moving the first auxiliary verb to the left of the subject NP ($NP + Aux + V + X \Rightarrow Aux + NP + V + X$). The result would be the following:

Will Aaron eat his lunch?

In Japanese, a yes/no question is formed not by a movement transformation as it is in English, but by an insertion transformation. The suffix *-ka* is inserted on the end of a verb to form a question from a statement, but the order of the words is not changed:

Kyou (watashi wa) gakkou ni ikimashita.—(I) went to school today.

Kyou gakkou ni ikimashita *ka*?—Did you go to school today?

Sequences of transformations

So far, we have discussed deep structures that have undergone only one transformation to derive a surface

structure. A deep structure may undergo many transformations. Consider the following two sentences:

Did the dog chase the cat?
The dog was blind.

A child may say these sentences separately, but most adults would say

Did the blind dog chase the cat?

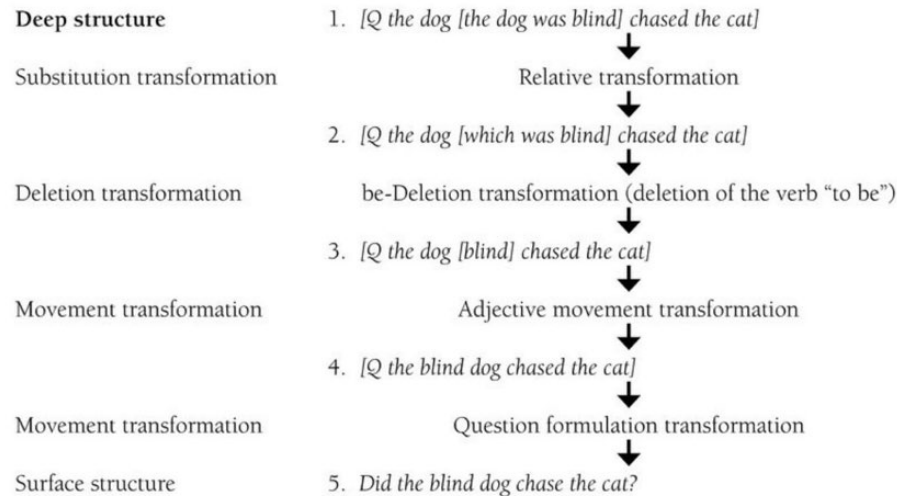


FIGURE 5-3 Transformation from deep to surface structure

Many transformationalists hypothesize that the deep structure of the adult utterance is similar to the child's use of two sentences. However, one sentence would be

embedded in the next as:

[Q the dog [the dog was blind] chased the cat]

Then it could be proposed that a series of transformations would apply to change this deep structure into a surface structure, as shown in [Figure 5-3](#).

Exercise 4 Transformations

1. Examine each of the following sets of sentences and determine whether a movement, deletion, insertion, or substitution transformation has occurred. The symbol $\Rightarrow$ means a transformation has been applied.

- # a girl was on a swing # $\Rightarrow$ There was a girl on a swing.
- # NEG Jessica will have gone to bed # $\Rightarrow$ Jessica will not have gone to bed. (Negation is considered by some linguists to be the first thought that occurs in the formation of what will become a negative sentence. NEG symbolizes this negative thought.)

- c. # Alma went to school early and Juan went to school early, too # $\Rightarrow$ Alma went to school early and Juan did so, too.
- d. # you will leave my house # $\Rightarrow$ Leave my house!
- e. # the fish, who is fat, swims slowly # $\Rightarrow$ The fat fish swims slowly.

2. The names of transformations are relatively descriptive. First, can you guess the names of the transformations in Question 1, parts a. through c.? What are the names of the transformations in Question 1, parts d. and e.? (Check the text.) Second, describe what each transformation does.

- a. _____

- b. _____

c. _____

3. a. _____

- b. _____

4. Label a. through e. as being either optional or obligatory.

- a. _____

- b. _____
- c. _____
- d. _____
- e. _____

Grammaticality judgments and ambiguity

As with all levels of language—phonetic, phonological, morphological, and semantic—the syntactic level is rule-governed. The rules that govern each of these levels or systems are often subconsciously known.

Fluent speakers of a language possess enormous subconscious knowledge, known as linguistic competence, of the rules of their own language. There is more knowledge of language in the mind of a fluent speaker than in all the grammar texts combined. On the other hand, we know nothing about the rules of a language we do not understand. A foreign language, although governed by rules just as our language is, may sound like gibberish to us. We do not know where one word ends and another

starts, let alone anything more complex. In other words, we have an enormous competence for languages in which we are fluent and none for languages we have not learned. Even universal features of language will not be recognized in a language that we do not understand.

What subconscious knowledge do we have about the syntax of our own language and languages we have learned? A fluent speaker of a language knows whether or not an utterance is complete; that is, whether or not it is missing an obligatory component or not. A fluent speaker will also know about proper word order, the proper relationship of words, and will often recognize ambiguous utterances.

A **grammatical (well-formed)** sentence is one in which the sequence of words conforms to the syntactic knowledge (rules) of native speakers of a language.

If you are a fluent speaker of a language, your subconscious knowledge allows you to produce **grammatical** or **well-formed** sentences. A sentence is grammatical if the sequence of words and the relationship between words conform to the syntactic knowledge (rules) of fluent

speakers of a language and if the sentence contains all of its required components.

An **ungrammatical (ill-formed)** sentence is one in which the sequence of words does not conform to the syntactic knowledge (rules) of fluent speakers of a language.

A fluent speaker also will immediately recognize that certain sentences are **ungrammatical** or **ill-formed**. A sentence is ungrammatical if the sequence of words and the relationship between words do not conform to the syntactic knowledge (rules) of native speakers of a language or if the sentence does not contain all of its required components. You cannot just randomly arrange lexical items to create a sentence.

Grammaticality judgments about completeness

*That house not pretty.

Any fluent speaker of English would recognize this sen-

tence as being incomplete. Furthermore, a fluent speaker could easily say why it is incomplete. There is no verb. The corrected sentence might read

That house is not pretty.

All English sentences include a verb. The necessity for a verb is one of the more consciously known rules of English syntax.

Although all languages have nouns and verbs, not every language requires a verb in every sentence.

[gwa? a kari kaa tutu?uli]

is a Yaqui (Native American language) sentence that loosely translates to *that house not pretty*. There is no verb in the Yaqui sentence, yet it is a grammatical sentence in Yaqui. *That house* is the subject of this simple sentence, and *not pretty* is an adjective phrase that acts as a comment (predicate) about the subject of the sentence.

Some words must occur with another word. For instance, the verb *holds* in some contexts needs another word to complete its meaning. The following is not a grammatical sentence:

*The boy holds.

The word *holds* is a transitive verb. Transitive verbs require a direct object. Thus, the following sentence would be complete:

The boy holds a ball.

Intransitive verbs do not take direct objects.

Grammaticality judgments about word order

Just as English speakers would recognize

*That house not pretty

as an incomplete sentence, they would recognize

*You had early get up wanted

as not being the correct English word order. Yet, that is the word-to-word translation of the German sentence

Du hattest früh aufstehen wollen.

Word-order rules are relative to each language. In English, looking at just the verb, we can see that in a complex construction like

had wanted to get up early

the word order can be analyzed as follows: helping or auxiliary verb (*had*), verb (*wanted*), verb complement (*to get up*), adverb (*early*). In German, the word order is: helping or auxiliary verb (*hattest*), adverb (*früh*), main verb (*aufstehen*), modal verb (*wollen*).

In a simpler sentence, such as

He sees a man

Er sieht einen Mann,

English and German display the same word order. Both languages in this case display the word order subject-verb-object (S-V-O). French, Thai, Swahili, and many other languages display this general S-V-O word order. Others may usually have an S-O-V word order, as in Bengali, Turkish, Persian, Japanese, and Navajo. Still other languages may display V-S-O (Tagalog, Irish, and Welsh), V-O-S (Fijian and Malagasy), O-V-S (Carib, a language from Brazil), or O-S-V (Xavante of Brazil) word orders. Also, a few languages, such as some indigenous languages of Australia, allow the speaker flexibility in the choice of word order.

Linear word order is the specific sequence that different types of words follow.

S-V-O refers to a sentence's **linear word order**, the specific sequence that different types of words (lexical categories) follow. Linear word order is often specific for the type of sentence. For instance, the S-V-O word order of English describes declarative sentences, but not interrogative (question) sentences. In the interrogative sentence

Did he see a man?

the helping (auxiliary) verb is at the beginning of the sentence.

Case indicates the function of nouns, pronouns, and adjectives within a sentence and the relationship of these words to verbs and other words within the sentence.

Case indicates the function of nouns, pronouns, and adjectives within a sentence and the relationship of those words to verbs and other words within the sentence. In English and other analytic languages, linear word order alone usually indicates the grammatical function of a word (its case). In the sentence

Dogs chase cats

dogs is the subject of the sentence (nominative case) by virtue of its placement before the verb. If the sentence was

Cats chase dogs

cats are doing the chasing and the word *cats* is the subject of the sentence. In addition to placement in the sentence, the form of the word or an inflectional morpheme might indicate its case. In English, *I*, *me*, and *my* have the same meaning; that is, "the person speaking or writing." The form *I* indicates that the person is the subject of the utterance (nominative case); the word *me* indicates that the person is the object of the verb or preposition (accusative case), and the word *my* indicates the possessive or genitive case. With a word such as *Jack's*, as in

This is Jack's coat

the insertion of the inflectional bound morpheme *-s* indicates the possessive case. The word *Jack's* means “belonging to Jack.” For a plural noun such as *boys* the possessive is indicated by *-s'*; *boys' toys*.

Old English (449–1100 CE), Latin, and many modern languages have many more inflectional bound morphemes to indicate case. Modern English has just two, *-s* and *-s'*. In fact, in languages that have many case endings, linear word order is not always important to indicate case. The inflectional morpheme alone tells what the function of the word is in the sentence. For example, the Latin word *domus* means *house*. The *-us* marks the word as being the nominative case (singular), meaning that it is the subject of the verb. If the word ends in the bound morpheme *-i* (*domi*), it is in the singular genitive case (*of the house*). An *-o* bound morpheme marks the word as being singular and in the dative case (indirect object). So *domo* could mean *to the house*. The morpheme *-um*, as in *domum*, marks the singular accusative case (direct object); so *domum* would mean *house* in the sentence *He bought the house*. (Each case also has a distinct ending for the plural.) The word *domus* would be the subject of the sentence regardless of its position in a sentence. The word *domo* would be the indirect object whether or not it was at the beginning, middle, or end of the sentence. And the word *domum* would be the direct object no matter the

word order of the sentence.

Grammaticality judgments about word combinations

Some lexical categories of words can occur together and others cannot. For instance, “the looked” is not a possible combination of words in English. The articles *a*, *an*, and *the* do not occur before verbs. Articles occur before many types of nouns or a gerund. A gerund is a verbal form ending in *-ing* and it acts as a noun. An example would be “the *running* of the Kentucky Derby.”

Adverbs are not used to modify nouns. So, a fluent speaker of English would find the following sentence ungrammatical: **The quickly person is home*. Other classes of words only co-occur with other specific categories of words. For instance, *will* or *may* only occur with certain types of verbs.

Whereas transitive verbs require a direct object, intransitive verbs such as *to fall* never take a direct object. (See the discussion on co-occurrence restrictions in the “Labeling the constituents of a sentence” section earlier in this chapter). Thus, the following sentence is grammatical:

Jack fell.

However, this following sentence is ungrammatical:

*Jack fell the stairs.

Intransitive verbs either end a sentence or combine with (are modified by) a prepositional phrase, an adverb, or an adverb phrase. So the following sentences would be grammatical:

Jack fell down the stairs.

Jack fell quickly.

Jack fell very quickly.

Some verbs may or may not take a direct object. The verb *to drink* is an example of a verb for which a direct object is optional. Both of the following sentences are grammatical.

Annie drank.

Annie drank milk.

Grammaticality judgments: several nonfactors

Grammaticality of an utterance is not based on whether

or not you have heard that utterance before. Language is productive ([Chapter 1](#)), so that most of the sentences you create and hear or see (if written or signed) you have not experienced before. Also, the grammaticality of an utterance does not depend on whether you understand the words in the utterance or not. You might not understand the sentence *Polystyrene microbeads can be coated with a specified sensing ligand*. However, this sentence is grammatical and would be understandable to a person who knew the meaning of all of the words. The opposite is also true. You might understand the meaning of a sentence that you judge not to be grammatical. Consider the sentence *The people is in the room*. This sentence is understandable but ungrammatical. Grammaticality does not depend on factualness. The sentence *The president of the United States is a three-year-old cat* is grammatical, but not factual.

The grammaticality of an utterance is not based on whether or not the utterance makes sense. A fluent speaker might even judge a sentence or longer utterance as being grammatical if it contained nonsense words. Lewis Carroll was famous for his nonsense language poem “Jabberwocky.” The following well-known passage from Carroll’s *Through the Looking-Glass and What Alice Found There* (1872) illustrates his ability to create an utterance that would be judged as grammatical by many

English speakers (and probably Jabberwocky speakers, too), even though it is built in large part on nonsense words.

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

Ambiguous sentences

In addition to making grammaticality judgments, fluent speakers usually can detect ambiguity in a sentence. A sentence is ambiguous if it has more than one meaning. Can you see why the following sentence is ambiguous?

The women appealed to all men.

Lexical ambiguity or **polysemantic ambiguity** refers to the situation in which a word or phrase can refer to more than one meaning.

Here, *appealed* is the problem: Does it mean *were desirable*

or *pleaded*? Such ambiguity, which involves a word that has more than one meaning, is called **lexical ambiguity** or **polysemantic ambiguity**. Lexical ambiguity is often consciously used to form puns, such as

Fish are really smart. They always are found in schools.

Structural ambiguity (or **syntactic ambiguity**) exists when the constituents of an utterance can be arranged in more than one way, yielding more than one meaning.

Lexical ambiguity is a semantic problem, and we will discuss it further in [Chapter 6](#). When the constituents of a sentence can be organized in more than one way, we refer to **structural ambiguity** or **syntactic ambiguity**. The following sentence can be organized in two ways:

Chris owns large dogs and cats.

This is ambiguous because it can mean

Chris owns large dogs and cats (of any size).

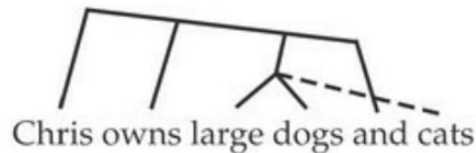
Here, *large* is linked to *dogs* but not to *cats*.



Or it can mean

Chris owns large dogs and large cats.

Here, *large* is linked to both *dogs* (solid line) and *cats* (broken line). This kind of ambiguity involves structural semantics, which we will discuss in [Chapter 6](#).



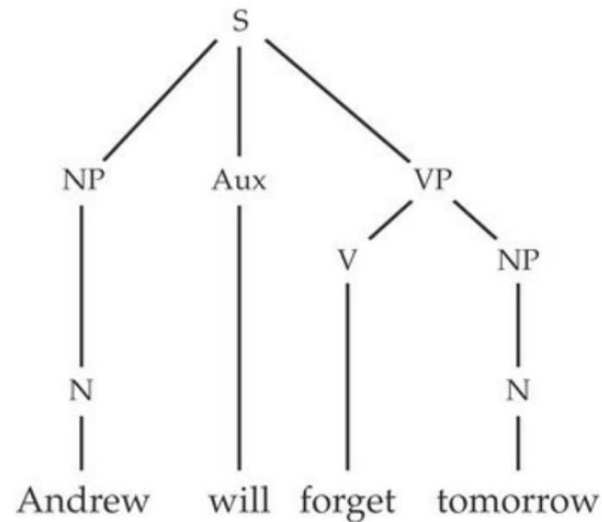
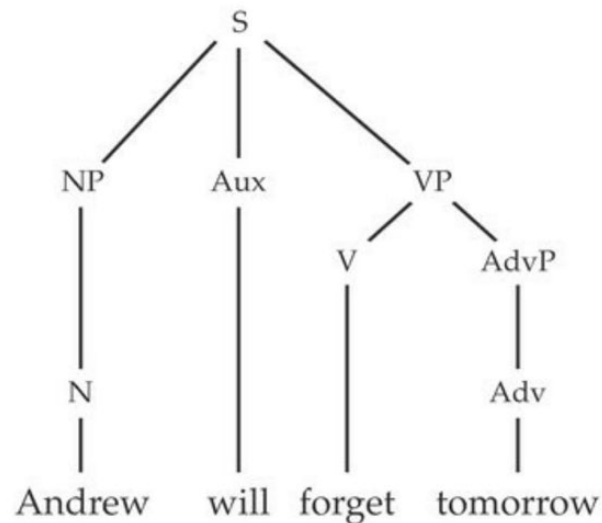
Part-of-speech ambiguity exists when a word in an utterance could be interpreted as belonging to different lexical categories; for instance, the word could function as either a noun or a verb.

Still another type of ambiguity occurs when neither words nor structure are ambiguous, but various constitu-

ents can be labeled as different parts of speech or lexical categories. This is called **part-of-speech ambiguity**. Examine the following sentence:

Andrew will forget tomorrow.

The sentence can mean that by tomorrow Andrew will forget something, or that Andrew will forget a specific day or event that is here labeled as *tomorrow*. Note that *tomorrow* has the same basic meaning in each interpretation and that the constituents of the sentence cannot be organized in more than one way. The ambiguity stems from the fact that *tomorrow* can be an adverb modifying *will forget*. It answers *when* Andrew will forget. Or it can be a noun. In this case, it is the object of the verb *forget*, and tells what is going to be forgotten. It is easier to see the possible arrangement of constituents in a phrase marker. The sentence could generate the following two phrase markers.



Notice that these diagrams are the same except for the function that *tomorrow* serves in the two sentences.

In addition to usually being able to detect when an utterance might be interpreted in more than one way (be ambiguous), native speakers of a language also know when words or sentences of the languages that they know mean the same or different things. No one will confuse

I am going to the ball game

with

Your house is on fire.

Exercise 5 What we know about our language

1. What is wrong with each of the English sentences below? Are there word order, word relationship, incompleteness, or ambiguity problems, or a combination of problems? Correct each sentence. **Example:** He had too much work had. The linear word order is incorrect. The sentence should read: He had had too much work.

c. Sentences that are structurally ambiguous because a constituent can function as different parts of speech

Synonymous sentences

We can also distinguish when two sentences have the same meaning. This can simply be due to the sentences using two different words for the same thing. For example,

Felines are more fastidious than canines

should be interpreted as meaning the same as

Cats are cleaner than dogs.

These sentences are synonymous because of word synonymy; *feline* and *cat* are generally synonymous, as are the word pairs *cleaner* and *more fastidious* and *canine* and *dog*. (They can carry some different connotations.) You will also know that the following sentences are synonymous:

Linguists often use large words.

Large words are often used by linguists.

Although it seems like a simple matter to recognize this synonymy, it is a complex mental skill. Here, we have two sentences that you may have never read or heard before. The first sentence is in the active voice and the second one is in the corresponding passive voice. Yet you understand that both are synonymous. This illustrates the productive feature of language.

Other sentences for which words have been reorganized turn out not to be synonymous for various reasons. A sentence used earlier

Andrew will forget tomorrow

loses one of its possible meanings if reorganized to its passive form:

Tomorrow will be forgotten by Andrew.

And a sentence such as

The dog ate a biscuit

becomes nonsense if rearranged as

A biscuit ate the dog.

Note that from a structural point of view, *A biscuit ate the dog* is a sentence. Sentences do not have to be true or logical to be sentences.

Also, a component of a sentence might lead to more than one type of ambiguity. The sentence

The French student was late to class

is generally ambiguous for structural reasons. The sentence could be restructured as

The student taking French was late to class

or

The student from France was late to class.

But the word *French* also has two slightly different meanings. In one case, it refers to the French language; in the

other it refers to the people from France. So the original sentence is characterized by both structure and lexical ambiguity.

Both ambiguous utterances and synonymous utterances can be explained in terms of the concept of deep and surface structure. With ambiguous utterances, the same surface structure is being derived from different deep structures. Synonymous utterances are derived when a specific same deep structure can result in two or more surface structures because of different transformations that took place to get to the surface structure.

Exercise 6 Synonymous sentences

Below are two pairs of synonymous sentences. Why are the sentences in each pair synonymous (or nearly so)? And in what way do the pairs differ from each other in the way they create synonymy?

Pair 1

1a. My psychiatrist thinks that I am mentally ill.

1b. My shrink thinks that I am nuts.

 Pair 2

2a. The will mentioned seven heirs.

2b. Seven heirs were mentioned in the will.

syntax. Syntax is the study of the rules to combine morphemes and words into linguistic units larger than words; morphology is the grammar of words; phonology is the grammar of speech sounds.

A sentence can be defined by its structure (the number of independent and dependent clauses it contains) or by its function (declarative versus imperative, for example). A sentence can also be either active or passive. A clause has both a subject and a predicate; a clause that cannot stand by itself as a sentence is called a dependent clause. A phrase is any constituent of a sentence that does not have both a subject and a predicate. Phrases are named after their main constituent, the head of the phrase, such as a noun. Phrases serve various functions in a sentence. For instance, a noun phrase may be the subject of the sentence, the object of a verb, and so on.

In the 1950s, Noam Chomsky revolutionized the study of syntax with his concept of transformational-generative grammar. This idea states that before we speak we have formulated an idea of what we are going to say. Universal grammar, a basic prewiring of the brain that predisposes all people to encode experiences linguistically in a specific way, converts those ideas into phrase structure rules. The universal phrase structure rules lead to the deep structure. To be understandable to others, the deep structure must be encoded into the specific grammar of

Summary

The study of grammar is the study of the rules of language. Language is a system of symbols that is rule-governed. Language is also a multilevel system of rules. Although most people see grammar as synonymous with the study of sentence structure and other forms larger than words, this is only one level of grammar, called

the language that one speaks. Once the experience is encoded in the deep structure, it is transformed by moving, deleting, substituting, or inserting various elements until a grammatical utterance is formed (surface structure). The same deep structure will therefore have different surface structures in different languages, or even within the same language, depending on the style of the speaker or the circumstance under which the utterance is spoken.

There have been numerous offshoots of Chomskian linguistics, and Chomsky himself made alterations to his earlier ideas, including substituting other concepts for deep and surface structure. In 1995, he published *The Minimalist Program*, in which he made additional changes to his previous hypotheses about language, including dropping the concept of deep and surface structure. His ideas about language continue to evolve, as do nongenerative alternative hypotheses such as cognitive-functional linguistics that is discussed in [Chapter 10](#), and challenges some of the most basic hypotheses of generative grammarians.

Fluent speakers of a language possess a linguistic competence that usually lets them make grammatical judgments about whether or not the rules of the language are being followed when a person is speaking to them. This is how we detect a person as being a foreign speaker. Some-

times foreign speakers will not use a complete sentence. They may leave out a verb or a preposition, for instance. In these cases, they are often using second-language words combined with the grammar of their first or native language. They may also use the wrong word order because word order differs in different languages. They may not understand idioms or why something that they said is ambiguous. The fact that fluent speakers of a language can usually detect deviations from grammatical rules is proof of the rule-governed nature of language.

¹Sylvia Bongard and Andreas Nieder, “Basic Mathematical Rules Are Encoded by Primate Prefrontal Cortex Neurons,” *Proceedings of the National Academy of Sciences* 107 (January 19, 2010), 2277–2282.

Suggested reading

Adger, David, *Core Syntax: A Minimalist Approach*, New York: Oxford University Press, 2003.

Carnie, Andrew, *Syntax: A Generative Introduction*, 3rd ed., Malden, MA: Wiley-Blackwell, 2013.

Chomsky, Noam, *The Minimalist Program*, Cambridge, MA: MIT Press, 1995.