The sidewalk is full of jostling people, the traffic in the street busy with midday cars and trucks. Through the crowd you spot a friend you haven’t seen in a while, wave, and catch up to her. She’s wearing a t-shirt (not her style) printed with the words “The Happy Pig” and a cartoony picture of a pig wearing an apron and carrying a tray of food. You ask her about this fashion statement. “Oh, it’s where I work now,” she says. “It’s, well, a . . . um, sort of Asian-fusion-deli place. They have terrific sesame noodles, totally awesome.” A bus roars by during the last of this, and you don’t actually hear all of your friend’s last sentence, but somehow you understand it anyway. “Where is this place?” you shout as another bus comes up. “It’s at the corner of—” Again, you can’t hear, and this time you have no idea what she said. Your friend fishes a takeout menu out of her backpack and shows you the address. “Great, I’ll try it!” you call as you’re swept away by the crowd.

Since the days of your early education, reading or hearing a sentence and comprehending its meaning has ordinarily been effortless and essentially instantaneous. (Fully appreciating the underlying concepts, of course, may pose more difficulty!) As the expert
1. The Nature of Language

Whenever you hear or read a sentence, you’re focusing on the meaning and relating it to information stored in your long-term memory. Despite the ease with which you now accomplish this, the cognitive processes that you perform to figure out the meaning of that sentence are actually very complicated. The discipline that explores the comprehension of language and the mental processes underlying it is psycholinguistics, the study of the comprehension, production, and acquisition of language. As its name suggests, the field draws both on psychology and on linguistics, the study of language and language structures.

1.1. Levels of Language Representation

Every sentence you hear or read is composed of many different kinds of information, among them sounds or letters, syllables, words, and phrases. These pieces of language fit together rather like an interlocking puzzle, so that the many components contribute to the overall meaning of a sentence. Language researchers refer to the pieces as different levels of language representation, and together they make up the grammar of the language. The term grammar often suggests rules of usage based on ideas like the parts of speech. Linguists and psycholinguists use the term differently. They use the term grammar to refer to the sum of knowledge that someone has about the structure of his or her language. Most of this grammatical knowledge is unconscious, but it underlies our ability to speak and comprehend a language with ease. Figure 12–1 diagrams the different levels of language representation that underlie the ability to understand the sentence “The chef burned the noodles.” At the top of the diagram is the level of discourse, which refers to a coherent group of written or spoken sentences. This level mentally represents the meaning of the entire sentence, beyond the meaning of the individual words. In the sentence “The chef burned the noodles,” an important part of the discourse representation is that “chef” is the agent performing the action and that “noodles” are the things being acted on. One way to represent this relationship is through propositions, assertions
made in clauses in sentences (Kintsch, 1998), as diagrammed as burn(chef, noodles) in Figure 12-1. A propositional representation (as shown also in Figure 1–4) concisely relates the action, the one doing the action, and the thing being acted on. A key part of language comprehension is arriving at this basic understanding of who did what to whom. The discourse representation also links the sentence meaning to the context in which it occurs (the conversation you’re having or the text you’re reading) and to information in long-term memory. This linkage allows you to relate the information in the sentence to prior knowledge (“The noodles were burned the last time we ate here, too!”) and to generate inferences (“Hmm, maybe we should try another restaurant.”).

Beneath the discourse level in the diagram is the level of syntax, which specifies the relationships between the types of words in a sentence (such as between nouns and verbs); syntax is a way of representing sentence structure, and many psychologists and linguists believe that it is part of our mental representation of a sentence as well. Here the sentence is composed of both a subject noun phrase (“the chef”), which at the discourse level maps onto the role of doer of the action; a verb phrase (“burned”), which describes the action; and another noun phrase (“the noodles”), which serves as the direct object and maps onto the role of the thing acted on. A standard way of representing the syntax of a sentence is a phrase structure tree, a diagram of a sentence that illustrates its linear and hierarchical structure (Figure 12–1). A phrase structure tree is a convenient way to talk about different components of a sentence but it’s also much more
than that. Many linguists and psycholinguists believe that in the process of understanding a sentence, we build a mental representation of the tree’s hierarchical representation of word relationships, and that this process is a key step in determining the meaning of the sentence. It is at the syntax level that comprehenders work out how the word order will relate to discourse information such as “doer of action.” For example, The chef burned the noodles and The noodles were burned by the chef both have the chef as the doer of the action at the discourse level, but the syntax of the two sentences is different.

A dramatic example of the importance of the syntactic level in language comprehension comes from studies of patients with brain damage. Patients who have had a stroke or other damage that affects parts of (typically) the left hemisphere of the brain may have aphasia, a language or speech disruption (so called from Greek words meaning “without speech”). Aphasia manifests itself in many different ways; one of them, which disrupts the syntactic level of representation, is called nonfluent aphasia or Broca’s aphasia, named for the French physician Paul Broca (1824–1880), who first described an aphasic patient with damage to a left frontal area of the brain now known as Broca’s area. This region is shown in Figure 12–2, which also shows

![Diagram of brain areas important for language](image_url)
other key areas important in language. Broca hypothesized that the region that now
bears his name was the location of language representations in the brain. We now
know that many other areas are important for language, and that the behavioral syn-
drome called Broca’s aphasia is not always tied to damage to Broca’s area, but the
names Broca’s area and Broca’s aphasia persist in part to honor this pioneer who
sought to link language and brain.

Patients with Broca’s aphasia have difficulty relating the discourse and syntactical
levels of representation. Thus, they may have a great deal of difficulty distinguishing the
meanings of *The chef burned the noodles, The noodles burned the chef,* and *The noo-
dles were burned by the chef.* Their difficulty is not in the meaning of the individual
words—these patients typically still know the meanings of words such as *chef* and
*noodles*—but with the relationships among them in the sentence. Because their knowl-
edge about how the world works is unimpaired, they would tend to interpret all these
sentences according to the most likely combination of the words *chef, noodles,* and
*burned,* and thus interpret all three sentences to mean that the chef burned the noodles.
The language disruption experienced by these patients highlights an important feature
of syntax, that recombination of words can produce sentences with different mean-
ings—sometimes even unexpected meanings, such as *The noodles burned the chef.*

Moving to the next level of the diagram in Figure 12–1, below syntax, are the *word*
and *morpheme* levels. These levels encode word meanings: for example, *chef* refers to
"someone skilled in cooking food." *Morphemes*, the building blocks of words, are the
smallest unit of meaning in a language. Some words, such as *the* and *chef,* are composed
of only a single morpheme, whereas others are built up from several morphemes.
*Noodles,* for example, is composed of two morphemes, the morpheme *noodle* plus the
plural morpheme, which we typically write as -*s*; and *burned* is also composed of two
morphemes, *burn* plus the past tense morpheme, typically written as -*ed*.

Compared to many other languages, English has a very simple morphological sys-
tem and very few morphemes, such as the plural and past tense forms, that attach onto
other morphemes (these are called *bound morphemes*). An example of a language with
a much richer morphology is American Sign Language (ASL), the language typically
used by the deaf in the United States. Figure 12–3 shows the verb meaning *give* in ASL
in its bare state (equivalent to the infinitive form *to give* in English) and with several
different bound morphemes added. The bound morphemes change the path of the ges-
ture used in signing the verb, and as you can see from the last example in this figure,
several bound morphemes can be combined during the production of the verb.

As well as classifying morphemes as bound or free, it is also useful to distinguis
morphemes that carry a good deal of meaning from ones that have relatively less
meaning but convey relatively more information about sentence structure. Mor-
phemes such as *chef* and *burn* that convey meaning but do not indicate much about
the structure of the sentence are called *content morphemes*. On the other hand, *func-
tion words* and *function morphemes*, such as *the* and the -*ed* past tense ending
in English, convey less meaning, but convey a relatively large amount of information
about relationships among words and about the syntactic structure of a sentence.
For example, *the* signals that a noun is coming up, and the past tense morpheme -*ed*
1. The Nature of Language

signals that, in this sentence, burn is a verb. Function morphemes (some of which are bound and some of which are free) link the levels of word and syntax. Interestingly, patients with Broca’s aphasia, who have difficulty with syntax, also have difficulty perceiving and producing function morphemes. The speech of Broca’s aphasics is halting and typically contains very few function words or morphemes. For example, when asked to describe the scene shown in Figure 12–4, a patient with Broca’s aphasia said “Boy . . . cookie . . . down . . . taking . . . cookie” (Goodglass & Geschwind, 1976). The only function morpheme here is the -ing suffix on the verb; other function words, such as a and the, are omitted.

In contrast, patients with Wernicke’s aphasia, also known as fluent aphasia, have a very different set of problems, which are at the word and morpheme levels. This type of aphasia often results from damage to Wernicke’s area, also shown in Figure 12–2, which is named for Carl Wernicke (1848–1904), the Polish-German neurologist and psychiatrist who described a patient with damage to this area.

Patients with Wernicke’s aphasia have generally good use of function morphemes, and their speech is typically fairly grammatical, with nouns, verbs and other parts of speech generally in the correct places in the sentence. But these patients can no longer produce content morphemes correctly, and the resulting speech is often nonsensical. In this example, a patient is attempting to describe the picture shown in Figure 12–4: “Well this is . . . mother is away here working her work out o’ here to get her better, but when she’s looking in the other part. One their small tile into her time here. She’s working another time because she’s getting, too” (Goodglass & Geschwind, 1976). Patients with Wernicke’s aphasia also have great difficulty comprehending content morphemes, with the result that they often have very little understanding of what is said to them.

FIGURE 12–3 American Sign Language has a rich morphology

(a) The uninflected form of the verb give, without any bound morphemes. (b) Adding a durational morpheme, yielding the meaning give continuously. (c) Adding an exhaustive morpheme, meaning give to each in turn. (d) Adding both those morphemes yields the meaning give continuously to each in turn.

(Based on What the Hands Reveal About the Brain by H. Poivner, E. S. Klima and U. Bellugi, MIT Press, 1987. © 1987 by the Massachusetts Institute of Technology. Reprinted with permission.)
The distinctions between Broca’s and Wernicke’s aphasias show two important points about how language is organized. First, the differences between the impairments of the two kinds of patients emphasize the distinct levels of how language is represented mentally and in the brain, and demonstrate how different levels can be affected to varying degrees. Second, the nature of patients’ impairments shows the degree to which these levels are interconnected: problems at one level, such as the disruption of function morphemes suffered by Broca’s aphasics, can contribute to difficulties at other levels, such as interpreting sentence syntax—which can then lead to difficulties in the comprehension of sentence meaning.

Look back one more time at Figure 12–1. The last level in the diagram shows phonemes, the smallest distinguishable units of speech sound that make up the morphemes in a given language. Spelling is not a precise enough system for representing speech sounds for a number of reasons. For one thing, writing systems vary in different languages (think of a page of the same text in Russian, Chinese, English, Hindi, and Arabic); for another, spelling rules in many languages have a number of exceptions that do not affect pronunciation (think of bear and bare, feet and feat); and, moreover, even native speakers of the same language may pronounce words differently. The solution is an alphabet of symbols for speech sounds—a phonetic alphabet—in which the speech sounds of all languages can be represented, independent of how they are spelled in any word or writing system. These are the symbols used to represent phonemes, as in Figure 12–1; you may be familiar with some of them from studying a foreign language or drama (many dialect coaches use

**FIGURE 12–4** The “cookie theft picture”
This picture is often given to aphasic patients to describe because its rich mix of people, objects, actions, and disasters provides the opportunity for many different kinds of descriptions.
the phonetic alphabet to help actors perfect an accent), or from the pronunciation key in a standard dictionary.

Like phrase structure trees, phonemes provide useful notation; they also lead to another claim about how language is represented mentally. Many language researchers believe that our knowledge of words includes representations of their phonemes. That is, while we are consciously aware of breaking a word down into letters for spelling, we also unconsciously represent words in terms of phonemes.

Patients with Broca’s aphasia offer further demonstration of how many of these different language levels relate to one another. We have seen that Broca’s aphasia is associated with poor comprehension and production of syntax, and also with poor comprehension and production of function morphemes. Some research links these deficits with the phoneme level, indicating that Broca’s aphasia may also include impairments in perceiving the function morphemes. Pairs of words were played to Broca’s aphasia patients (Bird et al., 2003). Some of them, such as pray and prayed, were identical except that the second member of the pair contained the past tense function morpheme -ed; in the phonetic alphabet, they are represented as [p ey] and [p eyd]. Other pairs, such as tray and trade (in the phonetic alphabet [t ey] and [t eyd] also sounded very similar, but were two different words with unrelated meanings. For each pair, the patients were asked whether they heard two different words or the same word said twice. The patients had great difficulty hearing the difference between pray and prayed; very often when they heard such pairs they thought the same word was being repeated. The patients were equally inaccurate in distinguishing tray from trade. The patients’ difficulty seems to be linked specifically to poor perception and comprehension of certain sequences of speech sounds, which in turn can lead to poor comprehension of function morphemes, and thence to problems interpreting syntax and ultimately understanding sentence meaning. The point is not simply that Broca’s patients have difficulty perceiving speech, but that language representations are interlocking, and failure at one level can have consequences that spread throughout the language system.

1.2. Language versus Animal Communication

There are more than 5,000 human languages in the world, representing a huge reservoir of phonemes, morphemes, words, and syntax. With all this variety, what do human languages share that the communication systems of other animals lack? Answering this question would be an important step in the search to define what it means to be human. Many animals that live in social groups, including songbirds, many species of monkey, and honeybees, have complex systems of communication. The American linguist Charles Hockett (1916–2000) compared animal communication systems and human languages and identified a number of key and unique characteristics of human languages (Hockett, 1960). These include duality of patterning, that is, the property that meaningful units such as morphemes are made up of meaningless units such as phonemes, which can be recombined over and over again to make different words. For example, the phonemes [t], [k], and [æ] (a is the phonetic symbol for the short-a sound) can be arranged in different ways to make three different English
words: [kæt], [ækt], and [tæk] (spelled cat, act, and tack). Animal communication systems, such as the alarm calls of vervet monkeys, do not have duality of patterning: vervets have one call for a leopard predator and another for an eagle, but they cannot recombine the sounds in the calls to make new calls.

Another important characteristic of language is its arbitrariness: in general, the relationship between the sound (or spelling) of a word and its meaning is not predictable. There is nothing about the sound [kæt] (cat) that intrinsically means a feline—the word does not sound like a cat or look like a cat, and the fact that we use [kæt] to mean a small mammal with whiskers that purrs is an accident of the history of English.

Perhaps the most important feature of human languages is their generative capacity: we humans can recombine morphemes, words, and sentences to convey a potentially infinite number of thoughts. Your waitressing friend’s t-shirt had a drawing that you described to yourself as cartoony, a combination of the morphemes cartoon and y, an ending that often signifies “like.” Even if you had never seen or heard this combination before, your own generative capacity would allow you to create it. And if you were to write about this encounter, describing the picture as “cartoony,” your readers’ generative capacity to understand new combinations of morphemes would allow them to determine the intended meaning.

Similarly, words can be combined over and over to make an endless variety of sentences. An important component of this generative capacity in syntax is recursion, that is, the embedding of pieces of a sentence (or an entire sentence) inside other pieces or sentences. Figure 12–5 shows the syntactic structure of a sentence with recursion, in that the relative clause whom the manager hired yesterday is
1. The Nature of Language

Although we typically keep our sentences fairly short, the use of recursion gives us the ability to make sentences that have many embeddings and could in principle be indefinitely long. For example, you could easily keep adding clauses onto the end of a sentence, as in *The chef burned the noodles that were made from the wheat that was grown on the farm that sat on the hill that was near the forest that...*

The property of recursion has played an important role not only in psycholinguistics but generally in the development of cognitive psychology. Behaviorists, most notably B. F. Skinner, suggested that the syntax of sentences could be described as a chain of associations from one word to the next. Skinner suggested that behaviorist principles such as operant conditioning could explain how children learned language by being reinforced for adult-like speech. The linguist Noam Chomsky (1959) strongly criticized the behaviorist approach to language, arguing that the property of recursion could not be captured by any chain of associations. For example, in the sentence *Any chef who burns the noodles gets fired*, the verb *gets* and its subject (*chef*) are not adjacent because there’s an embedded clause about burning noodles in between. A simple chain of association between adjacent words or phrases would incorrectly associate the noun *noodles* with *gets* rather than the real subject, *chef*.

Chomsky’s position, that the behaviorist account is inherently incapable of accounting for human linguistic abilities, was a crucial step in the rejection of behaviorist accounts of all aspects of human abilities. (Operant conditioning does explain some aspects of emotional learning, though, as discussed in Chapter 8.)

Even though the communication systems of highly intelligent nonhuman animals do not have the properties that Hockett observed in human languages, a number of researchers have asked whether chimpanzees could learn a language system if it were taught to them. Because the chimpanzee vocal tract (the parts of the anatomy needed for making sounds, including the vocal cords, mouth, and nose) is incapable of making most human speech sounds, researchers have taught chimps sign languages (Gardner & Gardner, 1969; Terrace, 1979) or communication systems using shapes on computer keyboards (Savage-Rumbaugh et al., 1986). Researchers found that chimps were good at using symbols or signs to make requests for food or other desires (for example, “strawberry,” “tickles me”). However, many researchers agree that these animals’ linguistic behavior does not go much beyond this, and it pales in complexity even with that of a 2-year-old human child.

Figure 12–6 dramatically demonstrates the contrast between the utterances of the chimpanzee Nim, who was taught a sign language, and the utterances of several hearing children who learned English and deaf children who acquired American Sign Language. The graph shows growth of utterance length over time in all the human children, whereas Nim’s sign usage did not grow in length or complexity. Exactly why chimpanzees can be so intelligent in some respects and so unlike humans in their ability to use language is a source of continuing investigation and debate.

Some language researchers have suggested that syntactic recursion is the crucial property that separates human language capacities from other communication systems (Hauser et al., 2002). This approach places the crucial difference between humans and apes at the syntactic level of representation, but other suggestions have
been offered. For example, Seidenberg and Petitto (1987) argued for a difference between humans and chimps at the word level. They noted that chimps clearly can relate symbol use and getting a reward, but they do not seem to understand symbols as names for things. For example, a chimp can learn that pressing the symbol for “strawberry” on a keyboard often results in getting a strawberry to eat, but it does not realize that “strawberry” specifically names the tasty red fruit and does not refer to eating, nor to the location where strawberries are kept, nor to other objects and events associated with getting a strawberry. The chimps’ behavior clearly is an example of communication, but it does not fit the criteria of human languages that Hockett and others have identified.

The utterance length of Nim, a chimpanzee who was taught a signed language, appears initially comparable to that of an 18-month-old speaking human child; however, utterances of both speaking and signing children rapidly increased in length with age, but Nim’s did not.

(From NIM Herbert S. Terrace. Copyright © 1979 by Herbert S. Terrace. Reprinted with permission of Alfred A. Knopf, a division of Random House, Inc.)
2. Processes of Language Comprehension

Comprehension Check:

1. What are the levels of language representation, and how do they interact?
2. What features distinguish human language from animal communication systems?

2. PROCESSES OF LANGUAGE COMPREHENSION

The different levels of language representation illustrated in Figure 12–1 reflect information that we know implicitly about language. How is this knowledge put to use in comprehending what we hear and read, and in producing language? How can we look at the letters C, A, and T together and realize that this sequence of letters indicates a small feline mammal? Our mental representations of words are a key component of a wide range of processes: speech comprehension, reading, writing, typing, and speaking. The first question then to explore is how the representations of words are maintained and accessed in the service of comprehension and production.

2.1. The Triangle Model of the Lexicon

Language researchers use the term lexicon to mean the entire set of mental representations of words. (The word is derived from a form of the Latin word legere, “to read,” an obvious connection with meaning.) Often a lexicon has been described as a mental dictionary, a repository of what each of us knows about words, what they stand for, and how they are used. This comparison, though it can be useful, is not quite right: our mental representations are not lists of facts about a word’s pronunciation, part of speech, and meaning—and they are certainly not in alphabetic order, a key characteristic of dictionaries! In fact, for some time researchers have pointed out that the list-structure idea of word representation fails to capture the degrees of similarity among word meanings, such as the fact that a robin and a cardinal are more similar to each other than either is to a duck (Collins & Quillian, 1969). These investigators instead promoted the idea that mental representations of words could be better described as networks. Reading researchers pushed these ideas further, observing that reading can be thought of as relating spelling to sound and spelling to meaning; again, the emphasis is on lexical knowledge as mappings between one level of representation and another (Seidenberg & McClelland, 1989). These and other considerations have led many researchers to conceive of word representations as networks comprising at least three major components: spelling, sound, and meaning.

In this triangle model (Figure 12–7), speech perception involves relating the sound representation—the phonology—of a word (the bottom right point of the triangle) to its meaning representation, at the top. Similarly, reading involves relating the spelling, or orthography, of a word (at the bottom left), to its meaning. Producing language involves relating the meaning of a word to its sound representation for speaking it aloud, or to its spelling representation for writing.
The triangle model shows how researchers have hypothesized how different aspects of word knowledge are linked, but it does not explicate the actual processes involved in language comprehension and production. What it does, however, is provide a framework for exploring several different kinds of comprehension processes in terms of relating one part of the triangle (such as sound representations) to another (such as meaning representations).

2.2. Ambiguity: A Pervasive Challenge to Comprehension

Comprehension is a complex business. Part of the reason is that most of the relationships among sound, meaning, and spelling are arbitrary—nothing about the spelling or sound of the word “cat” is inherently feline. Another feature of language that contributes to the difficulty of relating different levels of linguistic representation is ambiguity, which in language is the property that permits more than one interpretation of a sound, word, phrase, or sentence.

Language carries a huge amount of ambiguity at every level (Table 12–1), and the ambiguities at each of these levels have to be resolved before we can understand the meaning of what someone is saying. Let’s consider ambiguity at just one level,
2. Processes of Language Comprehension

Take a look at the objects around you right now: you’ll certainly see this book, and probably a chair, a light, and a pen. All four of these words, *book, chair, light,* and *pen* have both noun and verb uses, so that every time you hear or read these words, you must determine whether the speaker is referring to an object (noun) or an action (verb). Some words, such as *pen,* have even more meanings, in this case as two nouns (a writing implement and an animal enclosure) and two verbs (to *pen* a sonnet, to *pen* up the pigs). Look around you again: you’ll likely see more objects with ambiguous names (maybe *table, floor,* and *page* among them) than objects with names that seem to have only one clear meaning.

This exercise demonstrates a basic characteristic of words: ambiguity is rampant in the common words of the language; and those words that do not have multiple meanings—for example, *ozone, comma,* and *femur*—are typically technical terms and other relatively unusual words. How does this relate to the triangle model? It means that a single spelling or sound maps on multiple word meanings at the top of the triangle. As a result, for most words most of the time, we must sort through multiple alternative meanings, even though we typically become aware of only one interpretation. (Puns and similar jokes are a rare exception: the humor depends on making us aware of another meaning of an ambiguous word.)

Studies of how we resolve ambiguities in language are important for several reasons. First, these ambiguities are much more prevalent than we often realize. Research that explores the conditions in which resolving ambiguity is easy and the conditions in which it is not may help to improve communication in difficult situations, such as when using walkie-talkies, cell phones, or other static- or interference-prone devices. Second, a better understanding of why we are so good at resolving ambiguity may help in the development of computer language programs, such as speech recognition systems. Third, and most important, studying how we cope with ambiguity provides a

<table>
<thead>
<tr>
<th>Type</th>
<th>Perception</th>
<th>The Ambiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word boundaries</td>
<td>You hear or read <em>skrim</em></td>
<td><em>Ice cream?</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>&quot;I scream?&quot;</em></td>
</tr>
<tr>
<td>Spelling/pronunciation</td>
<td>You read <em>&quot;wind.&quot;</em></td>
<td><em>&quot;Wind&quot; like a breeze?</em></td>
</tr>
<tr>
<td>and word meaning</td>
<td>You read <em>&quot;permit.&quot;</em></td>
<td><em>&quot;PERmit&quot; as in a license?</em></td>
</tr>
<tr>
<td>Spelling and word</td>
<td>You hear or read <em>&quot;bark.&quot;</em></td>
<td><em>&quot;parMIT&quot; as in &quot;allow&quot;?</em></td>
</tr>
<tr>
<td>stress</td>
<td></td>
<td>The outer layer of a tree?</td>
</tr>
<tr>
<td>Word meaning</td>
<td></td>
<td>The sound a dog makes?</td>
</tr>
<tr>
<td>Sentence structure</td>
<td>You hear or read &quot;Mary read the book on the Titanic.&quot;</td>
<td>Was Mary reading about the Titanic?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Was Mary reading a book while aboard the Titanic?</td>
</tr>
<tr>
<td>Pronouns</td>
<td>You hear or read &quot;Susan told Mary that she was going to win.&quot;</td>
<td>Does Susan think Mary will win?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Does Susan think Susan will win?</td>
</tr>
</tbody>
</table>
CHAPTER 12 Language

good testing ground for understanding how language is mentally represented and processed. Language comprehension is naturally so fast and accurate that it is often very hard to gain insight about the process from direct observations, even with the simplest forms of language. If instead we can create experimental situations in which participants misinterpret ambiguities, we may gain some insight into how we typically integrate the various levels to understand language correctly.

A common theme running through all the research on ambiguity resolution is the integration of bottom-up and top-down information (first discussed in Chapter 2). Bottom-up information comes directly from what we perceive. Right now, as you read, one source of bottom-up information is the print on this page. In the triangle model, bottom-up information moves from the two bottom points on the triangle, spelling and sound information, to the top point, meaning. Top-down information comes from information in the long-term memories of each of us that helps us interpret what we perceive, and from information in the context in which the bottom-up information occurs. In the triangle model, top-down information also includes influence from meaning to spelling representations during reading. Because bottom-up information, such as printed text, is a very different thing from top-down representations of meaning, context, and other information in long-term memory, it is not entirely clear how two such different forms of information are integrated with each other to aid perception. Indeed, different claims about how such information is integrated form some of the major controversies in language research today. In what follows, we look at the roles of bottom-up and top-down information in speech perception.

2.3. Speech Perception

When someone speaks to you, fluctuations in air pressure strike your ear, and somehow you’re able to turn these sound waves into an understanding of what the speaker is saying. One key step in this remarkable feat consists of identifying the boundaries between the words the speaker is saying. This is an area in which reading and speech perception are dramatically different: in English and in most other writing systems there are clear white spaces between the printed words on the page, whereas in the speech signal boundaries between words are not marked by pauses. You may have the conscious perception of hearing individual words in speech, but in reality you’re hearing something that’s more like this: the words are all connected in a continuous speech signal. Jamming the print together on the page gives you some sense of the effect; an example of what speech actually looks like is shown in the sound spectrogram in Figure 12–8. A spectrogram is a two-dimensional visual display of speech in which time is shown on one axis, the frequency of the sound (which corresponds to pitch) on the other, and the intensity of the sound at each point of time and frequency is indicated by the darkness of the display (and thus white space indicates silence). The spectrogram in Figure 12–8 shows the spoken sentence “We were away a year ago.” Most of the words in the sentence are not separated by spaces, and there are some spaces in the middle of words, as in “ago.”

Without pauses to guide us, how do we find word boundaries? It appears that when we hear speech we unconsciously make educated guesses based on a mix of
2. Processes of Language Comprehension

Bottom-up and top-down information (Altmann, 1990). The bottom-up information includes cues from the speech signal directly, such as occasional stretches of silence when the speaker pauses to think. Top-down information includes knowledge about typical phoneme patterns, for example, that [b] and [k] do not typically occur next to each other in English words (so if you hear the sequence [bk], the [b] is probably ending one word and the [k] is starting the next) (McQueen, 1998). We have very detailed knowledge of this sort about our native language (or any other language that we speak well), but this knowledge is not helpful when we listen to someone speaking a language with unfamiliar patterns. Speakers of foreign languages seem to be speaking very fast in a jumble of sounds with no clear boundaries between the words. (The ancient Greeks referred to foreigners as barbaroi, “barbarians,” not because they were crude—they weren’t, necessarily—but because they spoke something that sounded like “barbarbar” instead of Greek.) In contrast, when we hear a language we do know well, we do not perceive the speech signal as a continuous stream, because our speech perception system is doing such a good job of guessing word boundaries; the result is the illusion that the boundaries, in the form of pauses, are actually in the physical signal.

A second key problem in speech perception is identifying the phonemes in the speech signal. There is a huge amount of variability in the way each phoneme is produced: every speaker has a different voice and a slightly (or not so slightly) different accent; and the clarity of articulation—the production of speech sounds—varies depending on the rate of speech, the speaker’s mood, and many other factors. The articulation of a phoneme also depends on what other phonemes are being articulated just before or just after it. Think about how you say the [k] sound in key and coo. Try preparing to say each word in turn, stopping just before you let the [k] sound come out of your mouth. How are your lips shaped in each case when you’re about to say the [k]? You’ll find the shapes are quite different, because even before you’ve gotten any sound out for the [k], your lips are already preparing to produce the following vowel. When the vowel is the “ee” of key, your lips are spread wide, but when the “oo” of coo is coming up, your lips are rounded. This overlapping of

FIGURE 12–8 “We were away a year ago”

This sentence is represented by (a) a spectrogram with (b) the approximate location of the words indicated below. In spectrograms, silence (pauses) in the speech signal appears as blank vertical bands; note that this sentence has no pauses between words, and the only silence is actually within a word, occurring at the “g” in the word “ago.”
phonemes in speech is called coarticulation, and it has a large effect on the sound of each phoneme. If you say the [k] in key and coo without the vowels but with your lips spread or rounded as if you were going to say each of these words, you will probably hear that the [k] in key is higher in pitch and sounds different from the [k] in coo. The phenomenon of coarticulation means that each phoneme is articulated differently depending on which phonemes precede and follow it.

The combined effects of coarticulation, variation across speakers, variation with speaking rate, and many other changes in the way people can speak mean that each phoneme is probably never articulated exactly the same way twice. This enormous amount of variation means that, in principle, it is extremely hard to identify which phonemes are in the speech signal. Nonetheless, we routinely accomplish this feat—how? Once again, part of the answer seems to be the use of top-down information, particularly information about the context in which a phoneme is pronounced. In this way a badly articulated, or even missing, phoneme is supplied through the phoneme restoration effect (Warren, 1970).

In deceptively simple studies of this effect, participants listened to spoken sentences on audiotape and reported what they heard. Unknown to the listeners, the experimenters had spliced out a small portion of the tape corresponding to a single phoneme in a word and inserted in its place a recording of a cough of exactly the same duration. (In the sentence “The state governors met with their respective legislatures convening in the capital city,” the asterisk indicates such a replacement.) Participants understood the sentences perfectly. Most did not notice that anything had been removed from the sentences; a common perception was that someone in the room had coughed during the recording. This illusion is even more dramatic in sentences such as the following (again, * indicates where a single phoneme has been spliced out and replaced by a cough):

- It was found that the *eel was on the orange.
- It was found that the *eel was on the axle.
- It was found that the *eel was on the shoe.

In these cases, the sound *eel is ambiguous among many different words, among them peel, wheel, and heel. Participants had no difficulty in comprehending the sentences, however, and they perceived *eel as a different word in each different context—they heard “peel was on the orange,” “wheel was on the axle,” and “heel was on the shoe.”

This result provides powerful evidence for the role of top-down information in phoneme perception. The word at the end of the sentence (orange, axle, or shoe) can influence the perception of *eel only after it has been partly recognized, its meaning partly retrieved, and its relationship to various words that end with a sound like eel considered. These sentences all start exactly the same way, and it is the top-down context occurring four words after *eel that biases listeners’ perceptions. These results also emphasize a point that has been confirmed in many studies: even though our conscious perception is that we instantaneously recognize the words we hear, very often in reality all the relevant information for recognizing a word does not arrive until after we hear it (Grosjean, 1985). Speech perception is extremely rapid, but it is not always quite as instantaneous as we feel it to be (see the Closer Look feature).
Another kind of contextual information in speech perception comes not from what we hear but from what we see. Someone who is hard of hearing may say, “I hear you better when I can see your face.” Someone who is not hard of hearing may say the same thing; whether we’re hearing impaired or not, a certain amount of our comprehension comes from lip-reading. Being able to see the speaker’s face provides additional information about which phonemes are being uttered because many phonemes are produced with characteristic mouth shapes. Mismatches between the speech sounds you hear and visual cues to articulation can be confusing—think of watching bad animation or a foreign film dubbed into English.

This confusion between what you see and what you hear is called the McGurk effect, after Harry McGurk, who discovered it by accident (Massaro & Stork, 1998; McGurk & MacDonald, 1976). McGurk and his research assistant, John MacDonald, were using video and audio tapes of mothers speaking to study speech perception in infants. When they dubbed the audio of a mother saying “ba” onto a silent video of her saying “ga” and played back the video tape, they were surprised to hear the mother on the tape suddenly saying a third sound, “da.” Eventually they realized that the problem wasn’t in the dubbing: if they closed their eyes and listened to the audio recording, they clearly heard “ba.” The perception of “da” when they watched as well as heard the tape was an illusion that arose because their perceptual systems combined cues from both the video and the audio recordings. The video tape conveyed “ga,” in which the consonant [g] is made by moving the tongue in the back of the mouth, and the audio conveyed “ba,” in which the consonant [b] is made with the lips in the front of the mouth. Speech perception processes combined these two conflicting signals to yield the intermediate perception of “da,” in which the consonant [d] is made in the middle of the mouth.

These examples of information integration explain why you could understand that your friend was saying that The Happy Pig’s sesame noodles are “totally awesome” even though the noisy traffic drowned out much of those last two words. You were able to supplement a very poor speech signal (the bottom-up information) because you were able to integrate top-down information from a variety of sources. You were getting some extra information from looking at your friend’s mouth during articulation—the [m] in awesome, for example, is made at the front of the mouth and is easily visible. You were also getting help from context (it seemed likely that she was going to offer some description of the noodles, which makes awesome, or another adjective, likely), and perhaps from long-term memory as well (maybe, irritatingly, she says “awesome” a lot). These sources together allowed your speech perception system to put together a very good guess about what was actually said. A minute later, however, when your friend was trying to describe the restaurant’s location, top-down information was scarcer. Mouth shapes might have helped a little, but there was little context—she could have been mentioning any of a number of streets or landmarks. As a result, when her speech was again masked by the traffic noise, you had no idea what she was saying.

Results we have discussed indicate the importance of integrating top-down and bottom-up information during speech perception, but they do not suggest how that information is integrated. It is thought that much of the recognition component of integration works via an unconscious process of elimination in which we consider a number of possible words, called a cohort, that match the speech signal we hear, and

Introduction
Words are not separated by silence in the speech signal, and thus the listener is faced with the problem of identifying the boundaries between words. One possible way to find the boundaries between words in speech is to try out many different hypotheses for words and word boundaries simultaneously. For example, someone who hears the sound sequence “rek-uh-men-day-shun” might briefly hypothesize that this corresponds to five separate words, wreck, a, men, day, shun, or three words, recommend, day, shun, but ultimately rejects these nonsensical combinations in favor of a single word, recommendation. However, except for occasional misperceptions, we do not have any conscious awareness of considering words that turn out to be wrong. Thus, it is important to seek experimental evidence for these unconscious processes in speech perception.

Method
Participants performed two tasks simultaneously: they listened to spoken sentences and at some point during every sentence, they saw a letter string on a computer screen and had to push a key indicating whether or not the letter string was a real word or not (a lexical decision task). Unknown to the participants, some of the spoken sentences and printed words had a particular relationship that was designed to address the question of whether listeners consider several hypotheses during speech perception. On these critical trials, participants heard spoken sentences containing a two-syllable word in which the second syllable formed a real word. For example, for the sentence He carefully placed the trombone on the table, the second syllable of trombone is the real word bone. For half of these sentences, the printed word in the lexical decision task was related to the embedded second syllable word (rib, which is related to bone), and on the other half of the trials, the word was unrelated (bun). If listeners were temporarily considering bone as a possible word in the sentence while they were trying to find word boundaries and recognize words, then activating bone as a possible word should prime related words such as rib. This priming of rib from bone would result in faster responses to rib than to the unrelated word bun in the lexical decision task. Neither rib nor bun is related to trombone, so if listeners immediately settle on trombone and do not consider bone during speech comprehension, then there should be no priming of rib and thus no difference in response times to bun and rib.

Results
In the lexical decision task participants evaluated words related to embedded second syllables (like rib, related to bone) faster than they evaluated unrelated words (bun).

Discussion
This result suggests that even though perceivers are not aware of considering several different word boundaries and words during speech recognition, they do activate possibilities (such as bone during recognition of trombone) that they rapidly reject. These results support the claim that speech recognition is a process of unconsciously trying out many alternatives and rapidly homing in on the one that is the best fit.
then gradually weed out those that don’t match the available bottom-up or the top-down information (Marslen-Wilson, 1984a). Thus, when you recognized the word awesome, you might have started out with a set of possibilities that contained words beginning with the same initial vowel sound: awe, awesome, awful, audit, audition, awkward, authentic, Australia, Austin, ... As soon as you heard the [s] in awesome, some of the words in the cohort no longer matched the speech signal (the bottom-up information) and dropped out of the cohort, leaving awesome, Australia, Austin, ... At the same time, you were also guessing about word boundaries, so your cohort may also have included a pair of different words, awe and some (Shillcock, 1990). Very rapidly, as more of the speech signal is perceived and as top-down information suggests that some possibilities don’t make sense (e.g., Austin is not an adjective), the remaining non-matching words drop out, and only one word, awesome, remains in the cohort.

Two key pieces of evidence support the view that speech perception involves a consideration of many possibilities from which the incorrect ones are weeded out. The first derives from the nature of the set of familiar words: although some words very rapidly diverge from other possibilities as more of the speech signal is heard, other words are similar to many other words throughout much of the speech signal. In the example of awesome, by the time you hear the au vowel and the s consonant, there are very few possibilities left in the cohort—awesome, Australia, Austin, awe, plus the start of a new word beginning with s, and not much else. However, the first two sounds of totally, the t consonant and the o vowel, leave many more possibilities in the cohort—totally, toast, tone, toe, told, Tolkien, toll, taupe, Toby, token, toad, and many others. Speech researchers describe these differences in terms of neighborhood density, the number of similar sounding words in the language. Awesome has fairly few neighbors, whereas totally is in a dense neighborhood of many similar sounding words.

If researchers are right in believing that we recognize words by initially considering a cohort of possibilities and eliminating incorrect ones, it stands to reason that the more neighbors a word has, the longer it is going to take to eliminate them and get to the point of recognizing the word that was actually spoken. This has been demonstrated experimentally: many studies have shown that participants are faster at recognizing words, such as awesome, that have few neighbors, than words such as totally, that have many neighbors—which thus confirms the neighborhood density effect (Luce & Pisoni, 1998).

A second piece of evidence about cohorts and the elimination process comes from observation of our involuntary responses. We have no conscious feeling that we are considering many possibilities during speech perception, but the cohort model suggests that the candidates in the cohort must evoke a greater degree of activation than words that are not considered. If so, we should be able to observe some consequence of this activation, and that, too, has been demonstrated experimentally. For example, in one study participants were shown objects on a table and told to follow the instructions they heard, such as “Pick up the beaker” (Allopena et al., 1998). In some conditions, the objects on the table included a beaker, a toy beetle (beetle overlaps with the first sounds of beaker), a toy stereo speaker (speaker rhymes with beaker but does not overlap in the first sound), and various other objects that had no sound overlap with beaker. The investigators...
monitored the participants’ eye movements to the various objects on the table, which indicated that while recognizing beaker, participants were also considering (glancing at) both beetle and speaker (but not objects that had no sound overlap with beaker). Similarly, other researchers found that when hearing words such as trombone that contain another word (bone), both words (bone and trombone) are considered (Shillcock, 1990). These results indicate that even though we have no conscious awareness of weighing alternatives during speech perception, we do in fact activate a cohort of possibilities on the way to recognizing words.

Why would we overlook mismatches at the start of a word and consider alternatives such as speaker for beaker or bone for trombone? One possible reason is the difficulty of recognizing word boundaries. We must guess about word boundaries at the same time we are developing candidates for the words we hear. If we are not sure about word boundaries, then we cannot be sure which phonemes are actually at the start of a word. Thus, it makes sense to consider many partially overlapping words in the cohort, even ones that have different initial sounds. The process of speech recognition is like working on several different intersecting answers in a crossword puzzle at the same time—you’re guessing word boundaries, which words you’re hearing, and which phonemes you’re hearing, all at the same time. A good guess in one of these areas can rapidly make guesses in the others much easier.

2.4. Representing Meaning

Identifying words is only the beginning of comprehension, and getting to the actual meaning of what the speaker is saying is the ultimate goal. In the triangle model (see Figure 12–7), computing the meaning of individual words is represented as the mapping between the phonological level and the meaning representation. Researchers often think of the mental representation of meaning as a network of interconnected features.

Some evidence for this non-dictionary view of lexical meaning comes from studies of patients who have sustained damage to the temporal lobes of the brain. These patients previously had normal language abilities, but their temporal lobe damage (typically either bilateral or predominantly in the left hemisphere) leaves them with impaired knowledge of word meaning (Figure 12–9). Some of these patients have category-specific impairments, that is, they have more difficulty activating semantic representation for some categories than for others (mentioned in Chapter 4). Researchers who study these participants ask them to indicate the meaning of pictures, either by naming the object in the picture or in some other way, such as by choosing among a set of pictures in response to an instruction like “Point to the banana.” Many different kinds of pictures have been used in these studies, spanning a range of living and nonliving things—animals, birds, plants, tools, musical instruments, vehicles, and others. Researchers have found a striking result: some patients are much worse on pictures in some categories than in others, and across patients, it is not always the same pictures or categories that prove to be especially difficult. Patients with this kind of brain damage tend to fall into two broad categories: those who are relatively more impaired at recognizing living things (animals, fruits, birds) and those who are relatively more impaired at recognizing
2. Processes of Language Comprehension

Manufactured objects (tools, vehicles, musical instruments). Occasionally patients have difficulty with a narrow category, such as fruits and vegetables, and less difficulty with other kinds of living things; see also Figure 4–16 and the accompanying discussion.

Some researchers have suggested that these patterns of impairments imply a semantic representation (a mental representation of meaning) that rests on various combinations of perceptual information (especially visual information) and functional information—information about what the thing is used for (Warrington & McCarthy, 1987; Warrington & Shallice, 1984). In this view, there is a broad division between living things and manufactured objects on these two dimensions. To the nonzoologist, living things are distinguished from one another mainly by perceptual features—a zebra has black and white stripes, an antelope and a deer have different shapes of horns or antlers. Manufactured objects, such as tools, writing implements, and furniture, have some important perceptual properties, but their function is typically more important. A pencil or a hammer or a car can be any color or pattern and still be a pencil or a hammer or a car, but if you altered a picture of a zebra to remove its stripes, most people would think it was some type of horse and not a zebra at all. Patients with greater impairment to the parts of the brain that process perceptual information thus will have more difficulty recognizing living things compared to manufactured objects, whereas greater impairment to brain areas

FIGURE 12–9 Regions of the brain involved in representation of word meaning

Damage to these areas affects understanding the meanings of certain words, sometimes in the form of category-specific impairments. The patterns of impairment suggest that words are represented in semantic networks that include various types of information, including perceptual aspects. Some of this information is also represented in similar areas of the right hemisphere.
processing functional information will produce greater difficulty in recognizing manufactured objects as opposed to living things. One exception to this pattern is musical instruments—patients who have difficulty identifying living things often also have difficulty with musical instruments. But the exception may not be so “exceptional”: true, musical instruments are manufactured objects, but fine detailed perceptual information is important in distinguishing them (think about how you would be able to know whether a picture is of a guitar or a violin).

The patterns of impairments argue against a semantic representation in which each word is its own distinct lexical entry, or even a scheme in which each category is stored in a separate brain area. Those sorts of organization do not predict what has been observed: particular semantic clusters become impaired together; for example, difficulty identifying musical instruments often accompanies difficulty identifying animals. Instead, the patterns of impairment suggest that the meanings of words are represented by combinations of perceptual information, functional information, and probably other types of information as well. The problems of the patients just discussed suggest that such networks of functional and perceptual information may be represented in different parts of the brain. That is, patients with greater difficulty recognizing living things would be expected to have more damage to areas of the brain that involve integrating perceptual features such as color and shape, whereas patients with greater difficulty recognizing objects would be expected to have more damage to areas of the brain related to function, particularly to motor areas, (because the function is often implemented by the way we manipulate an object).

Support for this hypothesis comes from neuroimaging studies of normal participants while they accessed semantic representations. In one such study, participants were presented with words and asked to think silently of color words appropriate to the presented words (such as yellow for the word banana, thus activating perceptual information), or action words (such as eat for banana, activating functional information) (Martin et al., 1995). The investigators found that thinking about color activated an area near brain regions involved with color perception, and thinking about related action activated an area near brain regions that control movement. These results, and others, suggest that meaning representations are distributed across multiple brain regions in networks coding various aspects of meanings, including perceptual features, movement features, and emotional associations. (A similar study was described in Chapter 4.)

If words are represented through networks of features, what happens when a word has several different meanings? Virtually every common word in English has more than one meaning. The problem is not so great in the “mental dictionary” account of meaning representations; a dictionary entry lists all the different meanings of an ambiguous word. If we think about meaning as emerging from feature networks, however, how are the different meanings of a word represented, and how do we activate the right meaning when we encounter an ambiguous word? A study that sought to answer this question presented participants with words that have equally frequent noun and verb meanings, such as watch and tire (Tanenhaus et al., 1979). These ambiguous words were presented in contexts in which the syntax of the sentence forced either a noun interpretation of the word (for example,
I bought the watch or the verb interpretation (for example, I will watch). The ambiguous word was always the last word of the sentence. Participants heard sentences of this sort, and after each sentence, they saw a word on a screen to read aloud. Sometimes the word to be read aloud was related to the noun meaning of the ambiguity, such as clock for the noun meaning of watch, and other times it was related to the verb meaning, such as look for the verb meaning of watch. The investigators compared reading times in three conditions: (1) when the onscreen word was consistent with the meaning of the ambiguous word as used in the sentence (for example, clock following I bought the watch), (2) when the onscreen word was consistent with the other meaning of the ambiguous word (clock following I will watch), and (3) a control condition in which the word that appeared on the screen was unrelated to either meaning of the ambiguous word. The investigators also varied the elapsed time between participants’ hearing the ambiguous word at the end of the sentence and the appearance of the (related or unrelated) word to be read aloud.

When the word to be read appeared immediately following the ambiguous word, the result was surprising: reading times for both of the related words (in this example, clock and look) were shorter than when the word was unrelated to either meaning of the ambiguity—even though one of the two related words made no sense in the sentence context. When the word to be read aloud appeared 200 milliseconds after the sentence ended, however, the participants read the appropriate related word faster than the inappropriate word: for example, the participants read clock faster following I bought the watch than following I will watch.

These results provide information about the relative timing of bottom-up and top-down information in the resolution of lexical ambiguity. It appears that immediately after an ambiguous word is heard, several of its different meanings are activated, but within 200 milliseconds top-down information from context has suppressed all but the meaning that matches that context. Some researchers have interpreted these results to imply that the resolution of ambiguity relies on a two-stage process: first, there is a bottom-up stage in which all meanings of an ambiguous word are accessed independent of context; then a second stage, in which top-down information about context is used to pick the correct meaning (see, for example, Swinney, 1979).

Subsequent studies manipulated variables such as the strength of the context and the relative frequencies of alternative meanings of ambiguous words (bank as a financial institution, for example, is more frequent than bank as in “bank of the river”), and the results have shown that these and other factors affect the extent to which multiple meanings are activated for ambiguous words.

A neural-network model has been devised to interpret these findings: in this model the activation of each meaning of an ambiguous word depends on the strength of its connection to the spelling or sound of the word (Kawamoto, 1993). For instance, two meanings of watch are initially activated because there are strong pathways from the spelling and phonology of watch to the two different meanings. Thus, activating either the spelling or phonology in the triangle model rapidly activates the two meanings. For words with one frequent and one relatively rare meaning, however, such as bank, the higher frequency meaning has been perceived and produced more often, and thus has a stronger connection between the meaning and
the spelling and phonology. The result is that the higher frequency meaning is activated more quickly than the lower frequency meaning.

In addition, top-down context effects are slower than these bottom-up effects when the contexts are not strongly associated with one or the other meaning. In the sentence "I bought the watch," there is nothing about the context "I bought the" that is particularly associated with timepieces; rather, only after the word "watch" appears can the context have a role in the interpretation of this ambiguous word. Other contexts, such as "I thought the alarm clock was running fast, so I checked my watch," are more strongly associated with the concept of time even before the word "watch" is encountered, and the effects of context are seen more rapidly in cases like this one.

This discussion of lexical ambiguity resolution in context has emphasized the timing of different kinds of information activation from bottom-up and top-down processes. It also relates, however, to the idea that words are represented by overlapping networks encoding different aspects of word meanings, such as the perceptual and functional features of a noun, such as banana. Banana is certainly not an ambiguous word in the sense that watch and bank are, yet activation in completely different brain regions was observed when participants were asked to think of the perceptual or the functional aspect of its meaning. Thus, every word has different aspects of meaning that can be emphasized to greater or lesser degrees in different contexts, and there is variation in the extent to which these different meanings are related. At one end of this distribution are words like watch, which have various meanings that are distinct from one another. At the other extreme are highly specific technical words like ozone. Because context can dramatically emphasize different aspects of word meaning, most words fall somewhere in between. Even for words with single definitions, such as banana, there may be core aspect—function or perceptual characteristic—of the object that is activated every time the word is heard. And although words, such as watch, that can signify different things show some early activation of different meanings, after the first few hundred milliseconds, sentence context contributes greatly to determining the features of the word that are active during comprehension.

2.5. Sentence Comprehension

Sentences provide contexts that can shade the meanings of individual words; they also of course have meaning themselves. Part of this meaning comes from the meanings of the words contained in the sentence, part from the syntax of the sentence—the relation of those words to one another. "Man bites dog" means something different from "dog bites man." But nothing is simple: sentences as structural wholes, as well as the words within them, can carry ambiguity. For example, consider the sentence "The spy saw the cop with binoculars." This sentence can be interpreted in two ways, depending on which of two possible sentence structures is assumed. Each of the possible structures (indicated by brackets) leads to a different meaning. If the sentence structure is "The spy saw [the cop with binoculars]," the prepositional phrase "with binoculars" is describing something about the cop, namely, that this is a cop who has some binoculars. But if the sentence structure is "The spy saw [the cop]
with binoculars,” then “with binoculars” is describing something about the manner of seeing, namely, that the spy is using the binoculars as an aid to seeing. This example illustrates structural ambiguity: the linear string of words that is heard or read is consistent with more than one syntactic structure and sentence meaning. The speaker or writer intended only one structure and meaning, and the listener, or reader, must figure it out, reconstructing it from the string of words. Structural ambiguities are extremely common in speech and writing, but nonetheless we generally manage to find our way to the correct interpretation. How do we do it? The occasional failures are quite revealing about how comprehension works. These failures are often the basis of jokes, in which the humor comes in part from lulling the listener (or reader) into one interpretation of a sentence and then springing another—the intended—interpretation on him or her. The great comedian Groucho Marx was a master at this. One of his most famous jokes of this nature is his line in the movie Animal Crackers: “One morning I shot an elephant in my pajamas. How he got in my pajamas, I don’t know.” The joke works because of a structural ambiguity in the first sentence, the same ambiguity as in the binoculars sentence above. The audience initially interprets the first sentence as “I shot [an elephant] in my pajamas” and is caught when the following context reveals the structure to be “I shot [an elephant in my pajamas].” This kind of ambiguity is called a garden path sentence because the listener or reader is first “led down the garden path” to an incorrect interpretation before being allowed to reanalyze the sentence and find the correct interpretation. Garden path sentences reveal a very basic property about sentence comprehension: its immediacy—we interpret words as we encounter them (Just & Carpenter, 1980). In principle, you could avoid dealing with many ambiguities by waiting until you’d heard the entire sentence, or even more sentences, before making any decisions about what the words mean and what the sentence structure is. That way you wouldn’t be surprised however an ambiguity resolved, because you would delay your interpretation until you had heard enough so that the context resolved any ambiguity. The fact that garden path sentences surprise us shows that comprehension proceeds as soon as we can make a reasonably good (unconscious) guess about what we’re perceiving. This means that we often have to make these guesses about correct interpretations from only partial information. Presumably we get to the right interpretation of ambiguities most of the time without ever being consciously aware of alternative meanings because these early guesses are either often right or are fixed very rapidly, before we have a chance consciously to notice alternative interpretations. We previously observed this same phenomenon in the phoneme restoration effect, in which people did not notice when a phoneme in a word was replaced by a cough, as in “the *eel on the orange.” The context from orange was integrated so rapidly that listeners believed that they had clearly heard the word peel. The same sort of rapid integration operates to resolve syntactic ambiguities. Whereas ambiguities in the speech signal must be investigated with spoken language stimuli, researchers studying syntactic ambiguity resolution typically use measures of reading time to test their hypotheses about how we arrive at the right interpretation of ambiguities. By presenting written sentences, researchers can measure reading time at each word (for example, by using a device that measures readers’ eye
movements) and thereby track at which point a sentence becomes difficult to comprehend (the eye pauses). Having a measure of difficulty at each point in the sentence is important in understanding ambiguities, because patterns of reading can reveal when readers have misinterpreted an ambiguous sentence. As we have just seen, structural ambiguities are temporary, lasting only until a later portion of the sentence makes the intended interpretation clear. This is typically the case in language comprehension. For example, consider these sentences:

1. Trish knows Susan... (structure: ambiguous)
2. Trish knows Susan from summer camp. (structure: subject–verb–direct object–prepositional phrase)
3. Trish knows Susan is lying. (structure: subject–verb–[sentence [subject–verb]])

Sentence 1 contains a temporary structural ambiguity concerning the interpretation of know and therefore of any words that might come after it. Know can mean “be acquainted with” and we can interpret the word that way, in which case the material that comes next in the sentence is simply a noun indicating who is known. Sentence 2 assumes this interpretation: Susan is the direct object of know, and the sentence means that Trish is acquainted with Susan. Alternatively, know can mean to be aware of the truth of something, in which case what follows know is generally an entire embedded sentence stating that truth. Sentence 3 is of this sort, and Susan is the subject of the embedded sentence “Susan is lying.”

The point in the sentence at which the structure and, therefore, the intended interpretation are made clear is known as the disambiguation region. In sentence 2, the disambiguation region is “from summer camp”; in sentence 3, it is “is lying.” Observations of reading times in the disambiguation region can reveal comprehension difficulties caused by ambiguities. Participants who read an ambiguous sentence and then encounter a disambiguation region that does not match their initial interpretation slow down in the disambiguation region (their eyes fixate on this region for an extended amount of time). At this point they realize that they have been led down the garden path and have to reanalyze the sentence, which takes more time (Rayner & Pollatsek, 1989).

Researchers have considered two general hypotheses about how we make early guesses about temporarily ambiguous sentences and, as in other problems of comprehension we have discussed, they involve different amounts of bottom-up and top-down information. One hypothesis holds that a syntactic structure is chosen first with only bottom-up information, and only later checked against top-down information (Frazier, 1987). According to this idea, a component of the language comprehension system, the parser, takes the written or speech input and builds a syntactic organization for the incoming sentence, much like the tree shown in the syntactic level of Figure 12-1. When a structural ambiguity is encountered, the tree could be built in two or more ways, so the parser chooses the simplest option, the one with the fewest possible nodes and branches. In the case of the ambiguous sentence about Trish and Susan, sentence 1, the first interpretation is the direct-object structure like that in sentence 2, because it is simpler than the embedded-sentence structure in sentence 3. As a result,
when Susan is encountered, it is immediately interpreted as a direct object. This choice is initially made without regard to any context or even to the meanings of the words in the sentence. In a second stage of comprehension, if the meaning does not make sense with the structure chosen by the parser, we realize we have been led down the garden path and hastily beat a retreat.

The parser hypothesis suggests that the strategies for resolving ambiguities of sentence structure are very different from lexical ambiguity resolution. Only one sentence structure is considered at a time, whereas we saw good evidence earlier that several different interpretations of an ambiguous word are activated during lexical ambiguity resolution. The difference comes from different conceptions of the lexicon and syntax—word meanings are stored in the lexicon, but syntactic structures are generated anew each time a sentence is heard. Thus, it is not computationally taxing to activate several alternative meanings of a word, but building several alternative structures simultaneously is thought to be too difficult to accomplish in the time we generally take to understand a sentence.

The alternative hypothesis is that we cope with ambiguities in sentence structure in basically the same way that we cope with lexical ambiguities (MacDonald et al., 1994; Trueswell & Tanenhaus, 1994). In this view, syntactic ambiguity resolution is a guessing game guided by both top-down and bottom-up information. Advocates of this view make the point that structural ambiguities also involve lexical ambiguities, as in “be acquainted with” and “have a belief about,” alternative meanings of know in sentences 1–3. This kind of lexical ambiguity, just like other lexical ambiguities, should yield activation of alternative meanings of the word that depend partly on the frequency of these alternative meanings and the context. And just as with lexical ambiguities, context effects on the resolution of structural ambiguity will in normal circumstances tend to be weaker than bottom-up information. This view is appealing because it emphasizes the interlocking nature of the word and syntactic levels of language representations, and it allows a consistent characterization of ambiguity resolution at many different levels: whether it is an ambiguity at the level of the speech signal, word meaning, or sentence structure, or some combination of these, the comprehension system unconsciously rapidly integrates whatever information it can to interpret the input in a way that fits best with the available evidence.

Researchers are currently working on testing these two hypotheses, but as of this writing the jury is still out.

2.6. Figurative Language

Figurative language is by definition ambiguous, in that it is the deliberate use of one word to mean another, by metaphor or simile. Your friend “fished” a takeout menu out of her backpack, but the use of this word is not meant to suggest that she produced a fishing rod, baited the hook, and made a spectacular cast in the middle of a busy sidewalk. Instead, it concisely evokes a picture of rummaging around in a container to look for something. Figurative language presents another problem of comprehension: we must decide whether a literal meaning or a figurative meaning is intended. Like other kinds of ambiguities, figurative language is amazingly frequent in ordinary speech; some analyses suggest that speakers use figurative language about
six times per minute of speech (Pollio et al., 1977), and figurative language is especially common in descriptions of emotions and abstract concepts (Gibbs, 1994).

Very often we do not even consciously realize that our language is figurative. Suppose someone says to you, “Registration was such a zoo—I spent two hours in line and then they said I didn’t have the right forms.” What’s the figurative language here? Zoo, of course, used as a metaphor to evoke a crowded and chaotic situation. But in this sentence there’s also a second, less obvious metaphor: in the phrase “spent two hours,” time is described as if it were money. Metaphors of this sort pervade our thoughts and emerge in a number of different expressions—in the case of time and money, we have metaphoric expressions about wasting time, spending time, saving time, spare time, and investing time.

Figurative language fills our speech to such a large degree that it is tempting to consider it as just another example, albeit a special case, of the multiple meanings of words and sentences. However, evidence from neuropsychology tells us that there is something more to the story. Although several sites in the left hemisphere are crucial for most aspects of language comprehension, the interpretation of figurative language appears to rely to a much greater degree on processing in the right hemisphere. Patients with right-hemisphere damage often have particular difficulty understanding figurative language, and neuroimaging studies have shown that normal participants have more right hemisphere activation while comprehending metaphors than while comprehending literal language (Bottrini et al., 1994).

The exact role of the right hemisphere in the comprehension of figurative language is not yet well understood, but it may have to do with the interpretation of sentence intonation, the “melody” of a sentence—the rise and fall of pitch, the variations in stress. Consider the sentence “Jim’s a really nice guy”: by uttering it with different intonations, you can make this sentence a statement of literal fact, a question, or a sarcastic comment indicating that you think that Jim is really not nice at all. The right hemisphere is highly involved in interpreting sentence intonation (Buchanan et al., 2000). The connection? Sarcasm, irony, jokes, and some other types of figurative language often rely on intonation. But the right hemisphere’s role in the interpretation of figurative language interpretation cannot rest entirely in fine-grained analyses of intonation, given that it plays a similar role in the interpretation of spoken and written speech (and there is no intonation of written speech).

2.7. Reading

When you couldn’t understand what your friend was saying, she gave you a takeout menu so you could see the address of the restaurant where she worked. The restaurant’s name, The Happy Pig, and the picture of the pig were at the top of the menu. You identified the picture as a pig very rapidly, and you read the words The Happy Pig very rapidly too. The processes that allowed you to identify the pig—that is, object recognition processes described in Chapter 2—include putting together perceptual information from the picture such as areas of light and dark and the location of angles so as to identify the object. In other words, you used visual information in the picture to activate semantic information—in this case the meaning “a pig wearing an apron
and carrying a tray of food.” As with interpreting a picture, the goal of reading is to translate visual information—the words on the page—into semantic information about the meaning of the words and the text.

2.7.1. Reading Pathways

In thinking about the task of translating print to meaning, it’s useful to refer again to the triangle model of the lexicon in Figure 12–7. Notice that there are two possible routes from the printed word to its meaning. The first is spelling → meaning, the route from the spelling of the printed word at the bottom left of the triangle up to meaning at the top, a route much like that for the recognition of objects. In both cases, you take information about patterns of light and dark on the page, angles, and other features and relate this visual input to stored representations of meaning. The alternative route is spelling → phonology → meaning: the print is first related to the phonological representation (that is, there is mapping between the two bottom points of the triangle), and then the phonological code is linked to meaning, just as in speech perception. When you read, you may have the sense that a voice in your head is saying the words to you; this effect appears to be the result of activating phonological codes from print as we read. Clearly, this phonological route is used when we sound out an unfamiliar word—the spelling is translated into a pronunciation. The phonological route is the basis of the “phonics method” of teaching reading. If you remember being explicitly told to sound out words when you were learning to read, this is how you were taught. Now that you are a skilled reader, do you still sound out words, or do you skip the pronunciation and use the spelling → meaning route? How much information flows through each of these two alternative routes?

It is hard to answer these questions just by thinking about them, because reading is so fast and automatic by the time most readers reach adulthood. The questions are therefore a topic of research, and answers to them have shifted dramatically during the past several decades (Rayner et al., 2001). Many researchers initially thought that skilled readers were, or at least should be, using only the spelling → meaning route. They thought that once readers had sufficient practice and no longer needed to sound out words, reading via the spelling → phonology → meaning route added a needless extra step that would slow them down.

Other researchers have taken a very different view of the reading process. They have noted that although the spelling → meaning route looks conveniently short in the triangle model, in fact the mapping between spelling and meaning is arbitrary. Again, there’s nothing about the letters C, A, T that look anything like a cat. Computing these arbitrary mappings is relatively hard, but the mapping between spelling and pronunciation has many regular patterns that make this mental computation easier. For example, the letter C is very often pronounced like the [k] sound, as it is in cat. Once this computation is complete, all that remains is the mapping between phonology and meaning. It’s true that this relationship also is arbitrary, but it is a route that is already highly practiced and is already used for speech perception in children who are beginning to learn to read.

According to this view, then, there is no reason to believe that reading via the phonological route is any slower than reading through the direct route from spelling...
to meaning. It is also important to remember that the “routes” here are not mutually exclusive, and in neural network models that implement the ideas behind the triangle model, the meaning representations can receive activation directly from spelling and through the spelling→phonology→routes simultaneously (Seidenberg & McClelland, 1989). Thus, it is possible to integrate these two positions and suggest that both routes may be used to varying degrees, and the extent to which one route or the other dominates at any given time may depend on various factors, including the skill of the reader and the properties of the word being read.

Of course, the arguments for different routes are difficult to assess without data. One study designed to investigate the extent to which phonology is used in reading had participants make judgments about the meanings of words (Van Orden, 1987). The task was set up in such a way that the fastest and most accurate way to perform it was to use the spelling→meaning route, ignoring phonology. The logic of the experiment was that if effects of phonology appear in a task for which they are not helpful, and possibly even counterproductive, then this is good evidence that readers routinely rely heavily on the spelling→sound→meaning route. Specifically, participants first read the name of a general category, such as FOOD or ANIMAL, on a screen. The category name was then removed from the screen and a new word appeared. Participants were instructed to press a key as quickly as possible to indicate whether or not the new word was a member of that category. For example, for the category FOOD, participants would press a key labeled YES for the word meat; but if the word heat appeared, they would press a key labeled NO. On some trials, homophone words such as meet appeared. (These are words that are pronounced in the same way as another word but are spelled differently.) Participants who saw the category FOOD and the word meet should answer NO, but some participants pressed the YES key, and most took a long time to make their judgments. The results point to participants’ use of the spelling→phonology→meaning pathway: the spelling meet activated the phonology of the word, and the phonology was used to activate meaning. Because the phonology maps to two different meanings, one of which is a food and one of which is not, participants had more difficulty producing a correct answer.

This result shows the importance of the spelling→phonology→meaning pathway in reading, but it does not mean that the spelling→meaning pathway is not also used. It is possible that pathway use can vary to some degree across readers or for different words, a hypothesis that was tested with a meaning judgment task similar to the one just described (Jared & Seidenberg, 1991). The investigators found that homophone interference, such as that illustrated with meet, was limited to relatively infrequent words in the language and did not occur for higher frequency words. They interpreted this result to indicate that the degree to which the spelling→meaning pathway is used depends on the amount of practice readers have had with this pathway: For high-frequency words, for which readers have a great deal of experience, the spelling→meaning pathway becomes heavily used; readers have had far less practice with lower frequency words and thus for these words they rely more on the spelling→phonology mapping and go from there to meaning.

Neuroimaging studies have produced corroborating evidence for this theory of how people use the different reading pathways. fMRI has been used to examine the amount of brain activity in readers of many ages and with many different levels of
2. Processes of Language Comprehension

reading skill as they read many different types of words (Pugh et al., 2001). These studies indicate that two different areas in the left hemisphere are crucial for fluent reading. One of these is in the temporoparietal areas of the brain, near regions that are important for word meaning and phonology, and the other area is the occipitotemporal system (Figure 12-10). The researchers suggest that when we learn to read, the temporoparietal system initially predominates and is responsible for learning the relationships in the spelling→phonology→meaning pathway. The occipitotemporal system develops later and becomes more important as reading skill increases. This system appears to relate visual information (that is, spelling) to meaning information directly. It is too soon to tell how definitively these brain regions relate to the spelling→phonology→meaning and spelling→meaning pathways as laid out in the triangle model, but fMRI research offers an exciting way to gain insight into the nature of the information processing that underlies reading.

2.7.2. Connected Text

A major part of reading is the recognition of individual words and the interpretation of sentences, but reading involves more than that if you are to make sense of this or any other connected text. What you see before you is not one sentence, but many connected sentences. How do we take in large amounts of text? The first point we must make is that there is a motor aspect to the act of reading: you're moving your eyes across the page so that you can see all the words. These eye movements come in quick jumps, called saccades, that alternate with periods in which your eyes are still while you're fixating on some specific point in the text. About 90 percent of reading time is spent fixating, and you make two or three saccades each second to new parts of the text.
When you are fixating on a word, the image of that word falls on the fovea, the part of the retina with the greatest visual acuity. The farther away from the fovea the image falls, the poorer the visual acuity. The sentence set off below has one word printed in boldface type. When you fixate directly on this word, its image will fall on your fovea. Now, still keeping your eyes only on the boldface word, try to identify the other letters in the sentence.

Tyc amksp birxu roz ulvdp walk gboh the pqzy gvlwn clg.

If you were careful to keep your eyes on walk and not fixate anywhere else on the line, you probably found that you could identify or make a good guess about the letters just before or after this word, but more distant letters were just a blur. An obvious question about this example: why were all the letter combinations other than walk nonsense words? The reason is that most participants in demonstrations like this tend to glance at the entire display while looking for the one boldface word. If you had done that and the sentence had made sense, you would have been able to use top-down processing to help you guess the letters on either side of walk. Having nonsense on either side of walk reduced the use of top-down information and gave you a good sense of how poor your visual acuity really is outside the fovea.

Many studies suggest that skilled readers fixate on most but not quite all words when they are reading material of roughly the same difficulty as this textbook. Some words, especially long ones, are fixated more than once, and some short words can be perceived even when they are not directly fixated. The word the is very often not fixated; it is short enough (and frequent enough) that it can be perceived even if the fixation is nearby. Moreover, when a word is fixated the eye movement to it is often precisely planned so that the fixation is in the middle of a word, so that all or most of the word is clearly visible in a single fixation. This brings up the interesting question of how you are able to plan your eye movements to skip certain words and focus precisely in the middle of others if you haven’t read the text yet and don’t know what’s coming up. In other words, how can you skip the if you haven’t seen it yet?

The answer is that in some cases you use top-down information from your accumulated general knowledge of English vocabulary and sentence structure to guide your planning of eye movements. And you also get quite a lot of bottom-up information from the text. It has been shown that even though readers have a clear image of only about six characters around the point of fixation, they can get some gross visual features of letters and words that are seven to twelve characters away. Look again at the nonsense sentence above. While fixating on walk, try to see the shape of the second “word” to the right. You probably found that you could identify, if not the actual letters, then at least that this was a short “word” that began with some tall letters. This sort of partial information is not enough for perfect identification of upcoming words, but it is enough to allow eye movement programming processes to guess that the word the may be coming up, and sometimes skip it.

2.7.3. Speed Reading
As a skilled reader, you read very quickly, every second identifying several words. Many people who must do a great deal of reading often wish they could read even
faster, and they invest in “speed-reading” courses. Does speed reading instruction work? No, not really.

Every speed-reading program is a bit different, but most make similar assumptions about how reading does or should work. One common—and outdated—assumption relates to the triangle model: speed-reading programs often suggest that efficient reading should relate spelling directly to meaning, avoiding the pronunciation route. These programs suggest that readers have got into the lazy habit of activating pronunciations while reading and could greatly improve their reading speed if they eliminated this “extra” step and read via the direct route of spelling → meaning. As we have seen, this view previously dominated the teaching of reading to children, but there is now good evidence that activation of pronunciation is in fact a natural component of skilled reading.

Most speed-reading instruction also encourages readers to move their eyes across the page faster. Because it is impossible to program and execute saccades any faster than our natural rate, the only way to get across a page faster is to make longer saccades. But this doesn’t help: because words outside the foveal region are not well perceived, the consequence of longer saccades is that some words will never be fixedated, or may even never be near the point of fixation, and thus won’t be seen. In other words, speed reading is a lot like skimming: you zip through some parts of the text and skip over others.

The reading patterns and comprehension abilities of trained speed readers have been compared with those of college students who had no speed-reading experience (Just & Carpenter, 1987). The college students read texts in two different ways. In one condition, they were instructed to read the text; in the other, they were instructed to skim. The investigators found that the speed readers’ eye movements were very similar to those of the college students who were skimming.

The study also explored the effect of skimming on the participants’ understanding of the material. (An important claim of speed-reading programs is that reading speed can be increased by their techniques without sacrifice of comprehension.) With easy texts, all three groups—speed readers, skimmers, and normal readers—were fairly accurate. With more difficult material, however, the skimmers and the speed readers had poorer comprehension than the college students who were reading normally. These results reinforce everything that we have seen about comprehension: essentially, that it is a multifaceted process in which many different levels of information are integrated with one another. Skipping over pieces of the material, whether by informal skimming or by following a speed-reading method, inevitably leads to missing key portions of context and other information that are crucial to developing an accurate meaning representation of the material.

Comprehension Check:

1. Why do researchers believe that when hearing speech, we initially consider many possible words (a cohort) and then reject possibilities that don’t fit?

2. What is the relationship between the triangle model of the lexicon and the phonics and whole-word methods of reading instruction?
3. PROCESSES OF LANGUAGE PRODUCTION

You’re taking an exam and your pen drops and rolls into the next row of seats. You tap the person sitting in front of you and say, “Could you please get my pen? It’s under your chair.” You’ve again effortlessly completed a complex feat. You’ve taken a goal to have your pen back, an abstraction that doesn’t initially have any language associated with it, and you’ve turned that goal into a linguistic representation and then into series of muscle movements to utter the words. (This is similar to how we convert a motor goal—lifting a pen—into a series of movements, as discussed in Chapter 11.) There are a host of other ways you could have translated this thought into language—perhaps “My pen fell under your seat. Could you please get it?” or “Would you please give me the pen that rolled under your seat?”—but for some reason you made this particular choice of wording. Studying language production means studying the processes by which we turn nonlinguistic thoughts into language and then develop an actual plan for the utterance.

Compared to what we have learned about language comprehension, we know relatively little about how humans produce language, primarily because of several methodological challenges. In studies of comprehension, researchers typically present language stimuli and measure variables such as comprehension time, accuracy, and patterns of brain activations. Measurements of these variables can be very precise, in part because we can control exactly when the stimulus is presented, and we can time our measurements from that moment. In language production, however, the start of the process is the creation of a set of nonlinguistic internal representations, such as those that underlie your desire to have your pen back. This different kind of starting point is a real challenge for investigators, because it is much harder to have precise experiment control. Indeed, much of the early research in language production was not experimental at all but instead observational, recording speakers’ errors in production. It is also possible to examine errors during production of signed languages and during written language production (handwriting or typing). However, the vast majority of research has been on errors during speaking, owing in part to the ease with which speakers’ errors can be observed.

One of the first researchers to use this method was Victoria Fromkin (1923–2000), a linguistics scholar who argued that the patterns of errors that speakers make are informative about the underlying processes of language production (Fromkin, 1971). She carried a small notebook with her everywhere, and whenever she heard an error in speech, she’d whip out her notebook and write down the error and the intended utterance. The collections of errors that Fromkin and other researchers gathered revealed that the patterns of speech errors are not random but group together in specific ways. These error patterns formed some of the first data on language production processes.

Exchange errors occur when two elements of a sentence are transposed. In word-exchange errors, such as I wrote a mother to my letter, and tune to tend out, words in a phrase or sentence exchange places. The exchanged words are typically from the same grammatical class, so that nouns exchange with nouns, verbs with verbs, and so forth. The exchanged words are often fairly far away from each other in the sentence.
In contrast, sound exchange errors, in which two sounds exchange places, typically occur in nearby words, from similar positions within the words. These errors are often called spoonerisms after the Rev. William Archibald Spooner, head of New College, Oxford University, from 1903 to 1924, whose speech was peppered with these errors. They include these possibly apocryphal reprimands to students: “You have hissed all my mystery lectures!” and “You have tasted the whole worm!”

The very different characters of word and sound exchange errors were important early evidence in the development of language production models, such as the one shown in Figure 12–11 (Garrett, 1975; Levelt, 1989). In this model, production processes proceed through three distinct levels before arriving at articulation. First, at the message level, the speaker (or writer) formulates the message to be conveyed. At this point the message is still nonlinguistic, with no words or sentence structure attached to it: an example is your desire to have your pen back. The next stage,
CHAPTER 12 Language

grammatical encoding, contains two different processes. One process involves choosing words with which to convey the message, the other involves developing the syntactic structure of the sentence to be uttered. All this information is then passed to another stage, phonological encoding, in which the phonological representation of the utterance is developed. Finally, the message is articulated. Let's take a more detailed look at some of these stages.

3.1. Grammatical Encoding

The process of choosing words during language production, one aspect of grammatical encoding, involves relating semantic information from the message level to individual words in the lexicon. Research in language production has focused on how those semantic representations are translated into particular word choices. For example, when you’re trying to get your pen back, you could describe it as being either under the seat or under the chair in front of you. Both are reasonable choices given the furniture and the message you’re trying to convey. This suggests that the meaning of your message partially activates both these word choices, and typically the one with the greater activation is selected. Occasionally two words are both selected, resulting in an error that is a blend of two different words, as in “My pen is under your cheat—I mean seat.” (Freud believed that errors of this sort reflected suppressed thoughts or desires, perhaps in this case a belief that the person in front of you—or you yourself—might try to cheat on the exam. Embarrassing mistakes like this one are often called Freudian slips. Some of these errors do seem to stem from thoughts that speakers don’t intend to utter, but most reflect the basic workings of the language production system more than our hidden motives.)

The other aspect of grammatical encoding is the development of a syntactic structure for the spoken or written utterance. As with word choices, myriad different sentence structures will do the job of conveying most messages. How do you settle on a single choice? It turns out that a great deal of the unconscious decision making for sentence structures depends on your word choices, specifically on how rapidly you are able to determine the various words that will go in the sentence. The process of choosing words starts for many words at once, so that while chair and seat are both being activated, so is pen, and probably get and give are competing with each other.

But even though the word-choice process starts for a number of words at about the same time, it is not necessarily completed at the same time for all the possibilities. There are a variety of reasons why word choice may be faster or slower at different times or for different words. For instance, certain parts of the message may be more important than others and thus the corresponding representations may be activated more strongly. Also, rare words may take longer to activate. These variations in lexical accessibility—the ease with which a word can be retrieved and readied for production—have a large effect on choice of syntactic structure for an utterance. The words that are chosen first are put toward the start of the sentence to be uttered, and the syntactic structure that develops is one that can accommodate these word choices.

This influence of lexical accessibility on structure choice was demonstrated in a series of ingenious language production studies that were disguised as memory
experiments (Bock, 1982). In one study, participants described simple pictures; they were told that saying a sentence out loud about each picture would help them to remember the pictures better. The syntactic structure that participants used in the picture descriptions was the behavior of real interest in the research. An example picture, shown in Figure 12–12, could be described with an active sentence such as “The lighting is striking the church” or a passive sentence such as “The church is being struck by lightning.” Bock hypothesized that the relative accessibility of the words lightning and church would affect the choice of active or passive syntactic structure, and she manipulated the accessibility of these words by briefly presenting a priming word, which the participant simply had to read aloud, before the picture appeared. (Participants were not aware of the connection between the task of word reading and that of sentence production.) In one condition the priming word worship appeared on the screen, in another condition the priming word storm appeared. Picture descriptions using passive sentences, such as “The church is being struck by lightning,” were more frequent when the prime word had been worship than when it had been storm. The prime word had preactivated some meaning
information, so that when the picture appeared and participants began planning their utterance, choosing a word for the primed concept was completed more rapidly than for the unprimed concept. This first-chosen word was then placed into an early position in the developing sentence, resulting in the choice of active or passive sentence structure as needed.

Putting the more accessible words first has real advantages in conversation, allowing speakers to set up the plan for the utterance early, before all words have been selected. As the beginning words are selected and the structure of the sentence starts to be developed, these early words move into the next stage of language production, phonological encoding. Speakers can therefore plan many different parts of the sentence at once. Because the earliest parts of the sentence are the ones in the lead in planning, these parts will be ready to be uttered while the speaker is still working on later parts. This has the conversational advantage of speeding up turn-taking: you don’t have to wait until the last detail of a sentence is planned before you begin uttering it.

The complicated interleaving of lexical selection and sentence structure planning sometimes goes awry. One of the consequences can be word-exchange errors such as *I wrote a mother to my letter*. Errors of this sort appear to result from a mistake in inserting words into the sentence structure, and therefore these misassignments are not completely random but typically happen within a particular word type. Nouns (such as *mother* and *letter*) exchange with each other, verbs exchange with verbs, and so on.

Another breakdown can occur if the speaker finishes uttering all the planned parts of the sentence before later parts are fully ready to be uttered. For this reason, speakers may find themselves suddenly at a loss for words (a common experience), at which point they often slow down or add filler words and vocalizations like “um” at this point. Remember your friend’s saying, “It’s, well, a . . . um, a sort of Asian-fusion-deli place”? Here the speaker was having trouble choosing words that adequately described this novel restaurant. She started uttering “It’s” before word selection and planning were complete, leading to the insertion of pauses and fillers while she planned the rest of the sentence and made it ready for articulation.

3.2. Phonological Encoding

As words are selected and pieces of the sentence structure are planned, these pieces of the utterance are sent off to the next stage of language production, phonological encoding. Here speakers retrieve the pronunciation representations that are necessary for articulating the words in the utterance. Remember that pronunciation representations are distinct from meaning (the triangle model of the lexicon again).

Indeed, there is a great deal of evidence in language production that choosing a word and retrieving its pronunciation are distinct stages. Usually these two processes happen so quickly that they are difficult to distinguish, but it is occasionally possible to isolate each stage. One way to do this is to study what happens when a speaker is having trouble thinking of a name or other word, in a “tip-of-the-tongue” state. We all have been in these states at one time or another, and a little reflection about
what’s going on in these situations reveals that “can’t think of the word” isn’t an accurate description—what you can’t access is the pronunciation of the word. When you can’t think of someone’s name, you know perfectly well whom you’re talking about. You have accessed the semantic component of the word quite satisfactorily, but for some reason you’re having trouble getting from the semantic representation to the pronunciation. These tip-of-the-tongue (TOT) states thus slow down a process that is usually too rapid to be observed easily, and they therefore provide researchers with valuable evidence about how phonological encoding proceeds. Obviously, following participants around and waiting for them to enter a tip-of-the-tongue state is a very inefficient way to collect this information, and so researchers have developed a procedure to induce these states. (If you want to try out a short version of this procedure yourself, or try it out on some friends, use the instructions and definitions in Figure 12–13.) In these studies, participants read or hear definitions of rare words and try to say the word that’s being defined. Sometimes they know the word right away, and sometimes they have no idea what word is being defined. But other times the experimenter gets lucky: the participant says, “Oh that’s a . . . um . . .” and is clearly in a tip-of-the-tongue state. The experimenter then asks the participant questions such as “Can you tell me what sound the word starts with?” or “How many syllables do you think it has?” People in a TOT state often can report several things about the word’s pronunciation, such as the first or last phoneme of the word or the number of

<table>
<thead>
<tr>
<th>Definition</th>
<th>Your Answer of Hint</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A Japanese robe with a wide sash.</td>
<td></td>
</tr>
<tr>
<td>2. An illusion in which a non-existent pool of water seems to be visible in the desert.</td>
<td></td>
</tr>
<tr>
<td>3. A device used to keep time in music.</td>
<td></td>
</tr>
<tr>
<td>4. A mammal that kills snakes, the main character in Kipling's &quot;Rikki Tikki Tavi.&quot;</td>
<td></td>
</tr>
<tr>
<td>5. A large upholstered footrest.</td>
<td></td>
</tr>
<tr>
<td>6. A type of giant redwood tree that grows in northern California.</td>
<td></td>
</tr>
<tr>
<td>7. A single lens placed over one eye to improve vision.</td>
<td></td>
</tr>
<tr>
<td>8. A single-celled organism that moves by means of temporary pseudopods.</td>
<td></td>
</tr>
</tbody>
</table>
syllables, or can name some other word that sounds similar. This partial knowledge of a word is not unique to spoken languages: deaf users of signed languages occasionally find themselves in a “tip-of-the-fingers” state. Like users of spoken languages, the signers in this state cannot quite come up with the sign they want, but can report partial information about it (such as the basic handshape of the sign they are trying to produce), without being able to demonstrate the movement required to produce the sign (Thompson, Emmorey & Gollan, 2005). These results suggest that someone in a TOT (or tip-of-the-fingers) state certainly knows the word and has some small amount of the phonological representation activated; they also demonstrate that the choice of a word and the activation of its phonological form (whether for speaking or signing) are distinct processes.

For most people, TOT states are just an occasional annoyance; a more devastating form can emerge from certain kinds of brain injury. In contrast to the patients with semantic impairments discussed earlier, other patients with brain damage may appear to have normal semantic information but severe difficulty retrieving the phonological representation of words they have selected for an utterance (Kay & Ellis, 1987). This deficit in naming things is often termed anomia, from Greek words meaning “without” and “name.” Anomia is a common consequence of injury to many language-related areas of the brain. Patients with anomia may look at a picture and be unable to retrieve the phonological representation, although they can demonstrate their understanding of the meaning in other ways: for example, a patient who sees a picture of a hammer and cannot retrieve the pronunciation for the word will readily mime hammering a nail. The deficit is more frequent for rare words, suggesting that the frequency with which the meaning—phonology pathway is used for a given word affects the ease of phonological retrieval. Like tip-of-the-tongue studies with unimpaired speakers, observation of anomic patients’ difficulties serves to underline the distinction between word choice and phonological encoding.

Phonological encoding, not surprisingly, is the birthplace of spoonerisms and other sound exchange errors. The exchanging sounds typically are in nearby words and are often located at similar positions in words (such as the first sounds of two words), and so they appear to represent slight glitches that occur when several words are planned simultaneously.

3.3. Integrating Grammatical and Phonological Stages

We know that information from grammatical encoding flows to phonological encoding; is there any evidence that phonological encoding can affect grammatical processing? Originally researchers assumed that such feedback didn’t happen. One reason for this belief was the nature of the error patterns that researchers observed in speech. Word exchanges seemed to occur over long distances and to depend on word type, not pronunciation, whereas sound exchanges occurred between sounds in nearby words, without regard to whether the words were nouns, verbs, or other word types. These patterns suggested that word exchanges happened before phonological encoding, and that sound exchanges happened only within phonological encoding, independent of grammatical encoding.
More recently, however, evidence has emerged that supports a more interactive system in language production. It is becoming clear that although there is a clear initial ordering of events in language production, so that speakers, signers (of American Sign Language), and writers begin to assemble sentences and choose words before they do much work on planning the motor movements that will enable them to speak, sign, write, or type, there is also some interaction between the later processes and the earlier ones. Some of the first evidence for this interaction again came from analysis of speech errors. Investigators examined a collection of word exchange errors and noted whether or not the exchanged words contained similar phonemes (Dell & Reich, 1981). For example, writing a mother to my letter is an error in which the exchanged words (mother and letter) sound fairly similar, whereas the error in The shirts fall off his buttons is an exchange between two phonologically dissimilar words (shirts and buttons). This examination revealed that the exchanged words contained similar phonemes more often than would be expected by chance—as though the similarity of sounds adds confusion and leads to misassignment of words into the sentence. This result indicates first that the process of phonological encoding of the words was well under way while the words were being arranged into their sentence, and second that phonological encoding was affecting the arrangement of words in the syntactic structure. Thus, this result argues for some interaction between grammatical and phonological encoding, because these phonological effects on word exchange errors should not happen if grammatical encoding proceeds completely independently of phonological encoding.

Similar evidence comes from experiments that try to induce speakers to produce exchange errors. In one such study, participants saw short phrases such as darn bore on a computer screen (Baars et al., 1975). The participants' task was to prepare to say this phrase if in a few seconds it was replaced by a GO signal. However, if the phrase was replaced by a different phrase on the computer screen (such as dole beam), participants were to ignore the previous phrase and say the new one. In some cases, participants saw many phrases before they got a GO signal. The phrases that appeared before the GO signal all began with the same pair of letters (e.g., da_bo_). Then came a GO signal, and a phrase appeared with the same initial letters, but this time reversed (e.g., what were the initial letters of the first word were now the initial letters of the second word and vice versa). After planning their responses on the basis of the initial series, participants found it difficult to switch, and this is where the sound exchange errors appeared. Sometimes the reversed sounds produced another pair of words: darn bore could become barn door. Sometimes the sound exchange produced nonsense, as when dart board became bart doard. Here's the interesting wrinkle: sound exchange errors occurred 30 percent of the time with darn bore—and only 10 percent of the time with dart board, which reverses to nonsense. This result provides more evidence for the interactive nature of language production: if phonological encoding were operating independently of other levels, then the rate of speech errors shouldn't vary as a function of whether the result is a real word or nonsense. Because speech errors that make real words are more common than errors that make nonsense, researchers have concluded that the phonological encoding level is interacting with the word selection processes.
Language production processes (like the general motor production processes considered in Chapter 11) are complex and multilayered. The data illuminating these production processes initially consisted largely of analyses of speech errors, but more recently experimental methods have been developed to study production in the laboratory. Data from both analyses of errors and results of experiments show clearly that speakers and signers accomplish a great deal of planning before beginning an utterance, and after beginning, they are still planning upcoming parts of an utterance while actually uttering the earlier parts. This simultaneous planning at multiple levels results in some interaction between the different levels.

Comprehension Check:
1. What are the stages of language production?
2. Why has it traditionally been more difficult to study language production in an experimental context than to study language comprehension?

4. LANGUAGE, THOUGHT, AND BILINGUALISM

Imagine that there are three students sitting in the row in front of you, one from the United States, one from Mexico, and one from Korea. Do they think about the world exactly as you do? Of course not. How much of that difference is from the accidents of language, that one is a native speaker of English, one of Spanish, and one of Korean? How much is from other differences in experience? Moreover, for the two students who have learned English later in life, how has the experience of learning a second language changed them?

4.1. Language and Thought

Languages differ dramatically from one another in terms of how they describe the world. Because language is the major vehicle for expressing our thoughts, scholars since the time of Aristotle have tried to understand to what extent the languages we speak shape the way we think. This question is most closely associated with the writings of Benjamin Lee Whorf, who argued that languages strongly shape the way their speakers perceive and conceive of the world (Whorf, 1964). Although not all differences between languages result in differences in thinking, researchers have found striking differences in how people who speak different languages perform a number of cognitive functions, such as how they navigate in and reason about space, perceive colors, and reason about objects, substances, events, time, numbers, and other people (e.g., see Gentner & Goldin-Meadow, 2003; Levinson, 2003). Many of the studies have found cross-linguistic differences even in surprisingly elementary processing (e.g., the ability to orient in space, judge duration, or perceive differences between different shades of colors). Beyond showing that speakers of different languages think differently, these studies have found that linguistic processes are pervasive in
many fundamental domains of thought. That is, it appears that ‘thinking’ is a complex set of collaborations between linguistic and non-linguistic representations and processes.

4.2. Bilingualism

The experiments designed to assess claims about the relationship between language and thought are faced with many complications: for instance, people who speak different languages also come from different cultures, making it hard to attribute any observed differences in thought to language differences alone. One “natural experiment” is the study of bilinguals, people who speak two (and despite the name, sometimes more than two) languages. The question here is whether one’s thought processes in one language are influenced by the ability to comprehend, produce, and think in another language. The question of bilingual language representations actually extends well beyond questions of language and thought. For every language process discussed in this chapter—speech perception, reading, ambiguity resolution, language production, and all the others—researchers have asked whether these processes operate in the same way for bilingual and monolingual speakers. For example, do bilingual speakers have one lexicon or two, one for each language? There seem to be some situations in which bilinguals’ two languages appear to be quite separate, and others in which they interact and even interfere with each other. Working out the factors that govern when two languages interfere and when they do not may help language researchers understand not only bilingualism but also all aspects of language representation in the brain.

In the United States, many people know little or nothing of any language except the one they grew up speaking, although this is changing. In most other parts of the world, however, most literate people speak two or more languages with some fluency. What are the consequences of being able to produce and comprehend in two languages instead of only one? The answers to this question can depend on how early the two languages are learned (see the accompanying Debate box) and how much each language is used—children need to hear and speak a language at least 25 percent of the time in order to become really proficient in it (Pearson et al., 1997).

Let’s consider the case of children who grow up hearing and using two languages roughly equally often and compare their situation to that of monolinguals. The bilingual child must learn to map each concept (such as dog, run, yellow) onto two different phonological forms and must learn the syntactic structures for two different languages, whereas the monolingual child is learning only one. This additional learning burden initially results in slower vocabulary development in each language in bilingual children when each language is considered individually. For example, consider a typical group of bilingual English–Spanish children, a group of monolingual English children, and a group of monolingual Spanish children. If we tested the vocabulary scores of the bilingual children and the monolingual children, the bilingual children would tend to have smaller vocabularies—they would know fewer English words than the monolingual English children and fewer Spanish words than the monolingual Spanish children. However, if we added up the total
CHAPTER 12 Language

Unlike so many other skills, learning a second language is easier when begun in childhood. Why is this? The answer may lie in the notion of a sensitive period for language learning: a period during development during which a child readily acquires an ability; before or after this limited period, acquisition is notably more difficult. Claims about a sensitive period raise many additional questions that do not yet have firm answers. For example, it is unclear exactly why the window of opportunity closes: is it because of a biological change, much like puberty, or is it an environmental effect—does the very act of learning one language thoroughly create interference for learning another?

We can ask how long the sensitive period is—when do you have to start learning a foreign language in order to be as fluent as someone who began learning it at birth? Researchers have offered various answers, some suggesting that you could wait until the age of 13 to start learning a language and still become as fluent (and accentless) as a native speaker. It’s unlikely that there is any one magic age cutoff, however, but most research suggests that for complete fluency without any trace of a foreign accent, it’s important to start learning a second language very early in childhood. It has been shown, for example, that although 6-month-old babies can perceive phonemes from many languages that they have never heard, they start to lose this ability by the age of 10 months (Werker & Tees, 1984).

Relating perceptual abilities that have a sensitive period to actual language mastery is complicated, however. An obvious experiment is to compare individuals who begin to learn a foreign language at different ages; immigrant populations offer a pool of potential participants for experiments of this sort. We could, for example, select people who first started learning English at age 10 and compare them to those who started learning at age 20, and test them after all the participants had been in the United States for the same number of years. This has been done, and essentially all studies of this sort show that the group that started learning English earlier performs much better in knowledge of English phonology and syntax (Flege et al., 1999; Johnson & Newport, 1989). However, it is difficult to know how to interpret these data, because the two groups likely have had very different language experiences and amounts of language practice: people who came to the United States at age 10 typically enroll in English-language schools, make friends with English-speaking classmates, and tend to get much more practice using English than those who arrive in the United States at age 20. Compared to children, adult immigrants are overall less likely to spend time speaking and listening to English, so that even though the two groups have been in the United States the same length of time, the quality and amount of their English experience tend to differ.

Analyses of language skills in immigrant populations that have tried to take these differences into account (Flege et al., 1999) have found that the sheer amount of language use, not the age at which language learning starts, is the best predictor of syntactic knowledge. Thus, the studies noted above may not be clear evidence for a sensitive period for learning syntax. By contrast, knowledge of phonology, as measured by accent, does appear to have a sensitive period, in that people who begin learning a second language early have better accents than those who learn late, even when amount of practice is very carefully equated in the two groups.

If there is a sensitive period for acquisition of language phonology, we can also ask whether the limits of the period are hard and fast. Researchers have investigated whether even a little early experience with a foreign language can stave off later declines in the ability to perceive and pronounce phonemes like a native speaker (Au et al., 2002; Oh et al., 2003). One study tested groups of English speakers who had spoken Korean for a few years before switching to speaking English full time; another tested English speakers who had spoken no other language but who had heard Spanish frequently
the number of phonology–meaning mappings that the bilingual children know in both Spanish and English, their total vocabulary scores would tend to be at least as high as those of the monolingual children (Pearson et al., 1993). Similar results have been found for learning the syntax of a language: the burden of having to learn two different sentence structures initially slows down bilinguals’ syntactic development in each of the languages (Müller & Hulk, 2001). Thus, children whose language comprehension and production experience is spread across two languages at least initially pay some price compared to children whose experience is concentrated on one language.

Although bilinguals eventually acquire both their languages extremely well, it is interesting to ask about the consequences of bilingualism in adulthood. A bilingual who routinely uses both languages will on average have half as much practice producing and comprehending each language as a monolingual. The effect of this split practice can be seen in word production, and bilinguals have been shown to have more tip-of-the-tongue experiences than monolinguals (Gollan & Acenas, 2004), even when the bilinguals are tested in the language that they use more often.

Another difference between bilinguals and monolinguals is that bilinguals must (consciously or unconsciously) choose which language to produce every time they speak or write, whereas monolinguals have no such choice to make. Whenever a bilingual uses one language, the other language must be inhibited, or language production
could turn into an incoherent mix of two different vocabularies and sentence structures. Several studies have suggested that this constant need to inhibit one language and focus attention on the other increases the effectiveness of executive processes (the topic of Chapter 7). Bilingual children and adults appear to do better on nonverbal tasks that draw heavily on memory and cognitive control resources, such as playing a game in which a sequence of lights and sounds must be remembered and reproduced (Bialystok et al., 2004; Bialystok & Martin, 2004). Thus, the experience of being bilingual has many consequences, both in and beyond language use itself.

Comprehension Check:
1. What is the Sapir-Whorf view on the influence of language on thought? What is a more moderate position?
2. How do bilinguals differ from monolinguals in their rate of language acquisition?

Revisit and Reflect
1. What are the different levels of language representation and how do they fit together?
The levels of mental representation of language discussed are the discourse, syntax, word, morphology, and phoneme levels. These fit together in a complex way because the various levels are distinct from one another (a phoneme is clearly not the same thing as a proposition at the discourse level), yet the levels are not completely independent. For example, certain sentence structures (syntax) are used to convey certain kinds of propositions in discourse, certain kinds of morphemes occur with certain kinds of sentence structures, and so on. Moreover, all elements of language can be recombined over and over; for example, the phonemes and the morphemes can be recombined to generate new words and sentences.

Think Critically
- Broca’s aphasics have difficulty comprehending and producing function morphemes. Many European languages, such as Italian and Russian, have more bound function morphemes (appearing as suffixes on words) in a typical sentence than does a typical sentence in English. Might this difference in the rate of function morphemes affect the nature of Broca’s aphasia in the different languages?
- The generative capacity of language implies that we can produce literally a potentially infinite variety of sentences, and that any sentence could in principle be indefinitely long. In practice, however, even long sentences don’t go beyond a few dozen words in length. Why do we seem to use only this small fraction of our generative capacity? Does the limitation seem to be in our production ability, in our comprehension ability, and/or in other types of cognitive processing?
2. How does language comprehension proceed at these different levels?

Language comprehension is a specific example of the perceptual processes discussed more generally in Chapter 2, and the same principles that emerged in that chapter can be seen in language as well. For example, language comprehension has both bottom-up and top-down processes. A key component of language comprehension is ambiguity resolution, because ambiguity is everywhere in language—such as in words, sentences, word boundaries, and phonemes. The ambiguities can interact, so that decisions about where a word boundary is can affect interpretation of what phonemes were heard and vice versa. We appear to cope with these multiple ambiguities by unconsciously making decisions that favor the most likely choice given all available information.

Think Critically

- How are the problems of ambiguity in understanding language similar or different to the problems of ambiguity in other perceptual processes that were discussed in Chapter 2?
- We appear to be able to activate multiple meanings of ambiguous words. Is this finding another example of a cohort—activating many possible words that partially match the speech signal? If so, could there be a kind of neighborhood density effect (perhaps the number of alternative meanings) in the interpretation of ambiguous words?

3. What are the similarities and differences in comprehension processes for spoken language and for reading, and what are the similarities and differences in comprehension processes for words and sentences?

An obvious difference in spoken and written language is that listeners must identify word boundaries in the speech signal, whereas in many writing systems, including that of English, the word boundaries are marked with spaces. Another important difference between hearing and reading is that the speech signal is fleeting, whereas readers can reread text if necessary. A final major difference is that reading is a skill that is taught, whereas spoken-language comprehension is acquired at a very young age without explicit instruction. Still, reading and speech comprehension have important similarities, in that we appear to interpret both speech and writing through a process of integrating top-down and bottom-up information to arrive at the most likely interpretation of the material.

Comprehension of words and sentences appears to rely on many of the same processes, which is not surprising given the fact that sentences are themselves made up of words. In particular, both words and sentences contain a great deal of ambiguity, and a substantial amount of research suggests that in both cases ambiguity resolution involves integrating top-down and bottom-up information to develop the most likely interpretation given what has been heard or read up to that point in the sentence.
CHAPTER 12 Language

Think Critically

- Some writing systems, including Chinese, do not put spaces between the words in writing. Do you think that reading in Chinese might be more similar to listening in Chinese than reading English is to listening to English?
- In the Groucho Marx joke that begins I shot an elephant in my pajamas, either the person doing the shooting is in the pajamas or the elephant is in the pajamas, but not both. How is this like the duck–rabbit ambiguity discussed in Chapter 2?

4. How do language users plan and produce language?

Language production involves translating a nonlinguistic message (a thought or goal that does not yet have any language attached to it) into an utterance. The planning of the utterance begins with grammatical encoding in which representations of words are selected to encode the message, and a syntactic structure is chosen to allow the most accessible words (those the language user is most ready to articulate) to go first. Next comes phonological encoding, in which the pronunciation of the words is retrieved and a plan for articulation is developed. The basic flow of information is from message to grammatical encoding to phonological encoding, but there is also some evidence for interaction between the levels, so in some circumstance planning articulation can affect grammatical encoding.

Think Critically

- What is the relationship between the accessibility of words and the tip-of-the-tongue (TOT) state?
- Think of some common phrases or titles that have two nouns joined by and or or, such as salt and pepper, “The Pit and the Pendulum,” or Pride and Prejudice. How often is the second noun longer than the first? Could the length of a word affect its accessibility and choices for word order in these phrases?

5. What is the relationship between language and thought?

Researchers have long debated to what degree one’s language shapes thought. The strongest position, most closely associated with the work of Sapir and Whorf, is that the language you speak shapes many of your thoughts. A more moderate position is that language shapes primarily your language-based thoughts, but it does not have a strong influence on early perceptual processes or cognitive processes that do not rely extensively on language.

People who grow up bilingual are initially slowed in their language development in each of their languages compared to monolingual children. This stands to reason, because a bilingual child will on average get only half the amount of practice with each language compared to a monolingual child. These effects of different amounts of language practice extend into adulthood, and both positive and negative effects of being bilingual have been found that can be traced to the different language comprehension and production experiences that bilinguals and monolinguals have.
Think Critically

- Eyewitness memories of events have been shown to be susceptible to leading questions, so that we will tend to "remember" a blue car as green if someone asks, "How fast was the green car going when it hit the tree?" (see Chapter 5) How does the fallibility of eyewitness memory here relate to the influence of language on thought?

- Bilinguals often engage in code switching, the use of some words from one language while speaking another language. For example, Spanish–English bilinguals who are speaking Spanish may include an occasional English word in the conversation. Could code switching in this way be related to lexical accessibility during utterance planning?
<table>
<thead>
<tr>
<th>Answers to TOT definitions, page 532</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. kimono</td>
</tr>
<tr>
<td>2. mirage</td>
</tr>
<tr>
<td>3. metronome</td>
</tr>
<tr>
<td>4. mongoose</td>
</tr>
<tr>
<td>5. ottoman</td>
</tr>
<tr>
<td>6. sequoia</td>
</tr>
<tr>
<td>7. monocle</td>
</tr>
<tr>
<td>8. amoeba</td>
</tr>
</tbody>
</table>