Questions to Consider

• What is language?
• How do we get from a string of sounds or marks on a page to something meaningful?
• How do we go from thoughts to spoken language?
• How do we acquire language?
• How does human language differ from animal communication?
Introduction: A Simple Conversation

Consider the following scene in a local coffee shop. Bill is sitting at a corner table and is approached by another young man.

TED: “Hey Bill.”

BILL: (looks up) “Ted! What’s up?”

TED: (sits in the chair across from Bill) “Last week I was talking to Rufus and Elizabeth after the circus. Man, the stuff he knows is dangerous.”

BILL: “Dangerous? I think he just needs a day of rest and a box of tissues and he’ll be alright.”

TED: (initially looks confused) “Um ... no, not a stuffy nose, THE STUFF HE KNOWS. You know, all the knowledge about what happens in the future.”

BILL: (laughs) “I got it dude. What was he talking about this time?”

TED: “Well, he was reading a newspaper and I was reading over his shoulder. The paper must have been from some time in the far future.”

BILL: “Why do you say that Ted?”

TED: “Well, the headlines were really bizarre, but Rufus didn’t even bat an eye. He acted like everything was perfectly normal. Man, things in the future must be really weird.”

BILL: “Give me an example of the headlines.”

TED: “Well, it was last week, but I remember one of them. It was ‘Enraged cow injures farmer with ax.’ Can you imagine it? In the future, cows carrying axes!”

BILL: (laughs again) “Ted, I think that you misinterpreted that headline. I think that the farmer was probably the one with the ax.”

TED: “Man, Bill, I’m glad I ran into you. I’ve got to go call my uncle and tell him not to worry about his cow Betsy. Party on, dude.”

Ted then gets up and leaves the coffee shop, while Bill laughs quietly to himself.

Most of the time communication using language feels relatively effortless and easy. However, as the opening story exemplifies, sometimes we stumble and communication fails. Bill misinterprets some of what Ted says because two things sound similar. Ted misinterprets the newspaper headlines because there are two possible meanings depending on how we build the underlying grammar of the headline. The failures (and successes) of our language use reveal that the apparent ease belies an incredibly complex system of information and processes. This chapter provides an introduction to psycholinguistics, a subfield of cognitive psychology that examines how we use language. As the term suggests, this area is heavily influenced by concepts from the fields of linguistics.
(which examines the structure of human language) and psychology (primarily cognitive psychology). In some respects, our discussion of the processing of language mirrors our discussions of memory. In the chapters on memory (see Chapters 5, 6, and 7), we discuss the processes of encoding (getting information into memory), storage (holding and organizing memories), and retrieval (getting information out of memory). For language, we focus on similar processes: comprehension (understanding language coming in), the mental lexicon (our storage of language information), and production (mapping thoughts onto language and articulating them). In this chapter, we start with a discussion of what language is and how it is structured. We briefly review research and theory about how we use and acquire language. We close with a brief discussion of how human language use differs from other methods of communication used by both humans and animals.

### What Is Language?

Philosophers and linguists have offered a number of answers to the question, What is language? It is a hard question to answer. Most definitions agree that language is used for multiple functions, the primary of which is to exchange information between individuals. In the opening story Bill and Ted are talking with each other with the goal of passing along information between the two. Bill wants to know how Ted is doing, and Ted wants to relate his experience with Rufus. The same basic purpose is true of the textbook you are reading. Our goal, as authors, is to convey concepts about cognitive psychology. The medium through which we are trying to accomplish this is the words and sentences you are currently reading.

### Structure of Language

Most theories of language assume it consists of different kinds of linguistic domains: form (phonology and orthography), meaning (semantics), grammar (syntax), and use (pragmatics). Psycholinguistic theories have borrowed many of the concepts from these domains, proposing that language processing involves different levels of language elements. Consider our story. It is made up of several elements: sounds (or letters), words, phrases, and sentences. These elements are related to each other hierarchically. That is, words are made up of sounds, phrases are made up of words, sentences are made up of phrases, and our story dialogue is made up of sentences. The traditional psycholinguistic theoretical approach has been to assume that each of these levels of language consists of representations and rules (consistent with the representational approach to cognition discussed in Chapter 1). This approach allows for the productive nature of language. We are able to produce and understand a potentially infinite set of sentences, including those we have never heard before. The following subsections illustrate this approach for different levels of linguistic information.

### Language Form: Phonology and Orthography

Consider the sounds that make up the first spoken line of our story, “Hey Bill.” There are five distinct sound units, two vowels and three consonants: /h/, /eI/, /b/, /I/, and /l/. These sound elements are called **phonemes**. Different languages are made up of

---

**Phonemes**: distinct sound units that comprise a language
different sets of phonemes (e.g., American English has roughly forty phonemes, Native Hawaiian has as few as thirteen, while some African languages may have more than one hundred). In addition to the individual phonemes, languages have rules that specify how to put the phonemes together (e.g., rules for syllables). For example, in English we can put the /p/ and /t/ phonemes together in some contexts (e.g., captain) but not in others (e.g., English does not have a word that begins with both together as in the nonword ptain).

Notice that there is not a one-to-one correspondence between the sounds (phonology) and the letters that make up a language (orthography). Consider Bill’s mishearing of Ted’s “the stuff he knows” as “the stuffy nose.” Had Bill read Ted’s comment instead of hearing it, he would not have made the mistake. As speakers learning English as a second language will tell you, the spelling-to-sound rules are not simple. For example, consider the letter c; sometimes it is pronounced with the /s/ phoneme and sometimes the /k/ phoneme (circus has both). The differences between the two kinds of stimuli (e.g., letters are made up of visual lines and curves that we see; sounds are made up of vibrations in the air that we hear) illustrates the flexibility of our comprehension processes to handle two sets of representations to achieve the same basic communicative goals. We discuss this in greater detail later in this chapter.

Morphology: Language Interface of Form, Syntax, and Semantics
Above these form levels in the hierarchy is morphology. Morphological units (morphemes) are the smallest representations that convey meaning and grammatical properties. Consider Ted’s utterance about the headline: “Enraged cow injures farmer with ax.” Can you imagine it? In the future, cows carrying axes! In some cases, what we think of as “words” are single morphemes (e.g., cow and imagine), but in many cases “words” are made up of multiple morphemes (e.g., cows is made up of cow and the plural morpheme -s). Additionally, not all morphemes have the same properties, so they interact with the rules of morphology differently. For example, free morphemes can stand alone (e.g., cow), while bound morphemes must be attached to other morphemes (e.g., the plural -s). Sometimes morphemes are used to add grammatical features (e.g., the past tense inflectional morpheme -ed may be added to indicate that the event occurred in the past); in other cases added morphemes result in a change of the meaning or syntactic class (e.g., noun or verb) of a word (e.g., in farmer the -er changes the meaning of farm from “a place to grow food” to “a person who grows food”).

Our story also illustrates how the phonological and morphological levels may interact in interesting ways. Consider what happens when we add the plural morpheme to the following words: ax and cow. Listen to how you pronounce the plural morpheme in the two cases. For cows we use the /z/ phoneme, for axes we use the /lz/ (there is a third too: with bikes we use /s/). The rules change how the plural morpheme is phonologically realized depending on the phonological environment. Thus, the morphological level is the point at which form, syntax, and semantic information interact.

Syntax (Grammar)
The next level of representation in the hierarchy is typically syntax. At the most basic, this is the level of representations and rules that specifies the ordering of words. Consider two sentences with exactly the same words but two very different meanings: “Man bites dog” and “Dog bites man.” It is easy to see that the word order is important for the overall meaning. Consider too what happens if we reorder the words as “Bites man dog.” This ordering makes very little sense, and we easily recognize it as an illegitimate sentence. Syntactic structure is the abstract representation that specifies how the words are related, not by meaning but rather the grammatical properties (e.g., nouns and verbs) of the words. The elements and rules of syntax are similar to the grammar you may have

Morphemes: the smallest units of a language that contain meaning
Syntax: the rules structure of a language
learned back in your early language arts classes of your youth. Consider the following basic phrase structure rules for English:

1. A sentence (S) is made up of a noun phrase and a verb phrase \[ S: NP \rightarrow VP \]

2. A noun phrase (NP) is made up of a noun (N) that may be modified by an article, an adjective, and a prepositional phrase \[ NP: (\text{art}) (\text{adj}) N (\text{PP}) \]

3. A prepositional phrase (PP) is made up of a preposition followed by a noun phrase \[ PP: \text{Prep} + NP \]

So the reason “Bites man dog” does not make sense is that it does not follow the rules of English. (English does not allow a sentence that starts with a verb followed by two nouns.)

This underlying syntactic structure is not simply the linear ordering of the words (surface structure). In fact, it is not uncommon to have multiple underlying structures corresponding to the same surface structure. Consider the headline that caused Ted some confusion: “Enraged cow injures farmer with ax.” The two interpretations (the cow has the ax or the farmer has the ax) correspond to two different syntactic structures. Figure 9.1 shows a “tree structure” that shows how the different syntactic chunks (constituents) may be arranged. In Figure 9.1a, the prepositional phrase “with ax” is part of the verb phrase, modifying the verb injures. Figure 9.1b shows a syntactic structure in which “with ax” is part of the object noun phrase, modifying the noun farmer. So the overall meaning depends not only on the meanings of the individual words but also on the abstract syntactic structure represented in the figures by the tree structures above the sentences.

**Figure 9.1 Different Syntactic Structures of a Newspaper Headline**

Two sentences with identical surface structures can have two different interpretations depending on the underlying abstract syntactic structure. In (a) there are six constituent nodes, and the prepositional phrase (PP) modifies the verb phrase (VP), resulting in an interpretation of “the cow injures with the ax.” In (b) there are seven constituent nodes, and the prepositional phrase is attached to the noun phrase (NP), modifying farmer and resulting in an interpretation of “the farmer with the ax.”
Semantics (Meaning)
Levels of representation above syntax are related to meaning—semantics. Most psycholinguistic theories make a distinction between linguistic elements like words and the mental concepts with which they are related (see Chapter 10 for a more detailed discussion of concepts). Shakespeare’s “a rose by any other name would smell as sweet” illustrates this distinction. The flower name rose is a word, with phonological (made up of phonemes /r/ /oa/ /z/) and syntactic properties (noun), while the concept corresponding to a rose might include features like these: scented flower from genus Rosa, comes in many colors and varieties, stems often have thorns. However, there are theoretical debates over how the linguistic and conceptual systems are related. Some theories don’t include separate semantic and conceptual representations. In these views, the semantic representations are the conceptual representations corresponding to words and sentences (e.g., Jackendoff, 1994, 2010). Other theories include semantic representations within the language processing systems separate from the conceptual system (e.g., Pavlenko, 1999). In these theories the semantic representations serve to map verbal labels to their corresponding concepts. For example, consider the case of a Spanish-English bilingual speaker (Francis, 2005). The speaker may have a single concept for ROSE, while having two separate semantic representations, one for English (rose) and one for Spanish (rosa). These semantic representations may include information about what kinds of roles the words may play or require in a sentence (e.g., who does what to whom). This information is particularly important for verbs. For example, the verb give requires an agent to give a theme to a recipient (e.g., “Mary gives the apple to John” sounds good, but “Mary gives the apple” sounds incomplete).

Pragmatics (Using Language)
So far we’ve discussed representations and processes involved with literal language, but we do not always use language literally. Paul Grice (1989) distinguished between sentence meaning (as just described) and speaker meaning (what the speaker intended to communicate). Consider Bill’s first line in our conversation, “Ted! What’s up?” The sentence meaning appears to be a question directed to Ted inquiring about what things are elevated, but Bill’s intention is probably to greet Ted and express a willingness to engage in conversation. So much of what is intended appears to be outside of the particular literal properties of the utterance (i.e., phonological, semantic, and syntactic). If this seems like an isolated case, take twenty minutes and listen to the conversations going on around you. Most likely you will find many examples of situations where people intend their meaning to be somewhat different from what they literally say (e.g., idioms like “he kicked the bucket,” metaphors like “my professor butchered my first draft,” and indirect requests like “can you pass the salt?”).

The subfield of linguistics that examines the use of language within particular contexts is called pragmatics. While pragmatics has a relatively long history of investigation within linguistics, psycholinguistic investigation of pragmatics (sometimes referred to as
Experimental pragmatics has been relatively rare and largely considered outside the mainstream (focused primarily on issues of figurative language like idioms and metaphors). However, within the past decade, research examining these issues is on the rise (Noveck & Reboul, 2008).

Language is more than a simple string of letters or sounds. Instead, this set of finite elements (letters and sounds) is combined in systematic ways to convey a potentially infinite set of meanings. In other words, language is a complex hierarchical system of abstract representations and rules for combination. As you might expect with such a complex system, the mental processes we use to process language are complex as well. The next section provides an overview of how we comprehend and produce language.

**How Do We Process Language?**

In 1861 Paul Broca examined a patient, Tan (the only word that he could speak freely), who had been unable to speak for twenty-one years. Tan seemed to retain his ability to understand language, but his ability to produce language was severely impaired. After his death, Broca performed an autopsy and discovered that Tan had suffered brain damage in the left inferior frontal region of his cortex (see Figure 9.2). This was the first documented case of expressive aphasia (also known as **Broca’s aphasia**). Patients with damage to this region of the brain have speech typically characterized as slow, effortful, and halting, lacking in most grammatical words (e.g., articles, prepositions).

Fifteen years later Karl Wernicke (1874) described a patient who apparently had the opposite problem. His patient had suffered damage to the posterior part of his temporal lobe and was described as having relatively fluent, syntactically intact production but impaired comprehension. Patients with damage to this area who exhibit similar deficits are diagnosed as having **Wernicke’s aphasia**. This early dissociation between language production and language comprehension processes shaped how later psycholinguists developed their lines of research and theories about how we use language. In particular, until relatively recently, most research has focused on either production or comprehension processes.

During the early half of the twentieth century psychological theory and research was dominated by behaviorism. Within this tradition, language processing was described with the same principles of behavior used to describe nonverbal behaviors. These views...
are best exemplified in B. F. Skinner’s (1957) book *Verbal Behavior*. The field of modern psycholinguistics is typically traced to the 1950s (see Levelt, 2012, for an excellent review of pre-Chomskyan psycholinguistic research). This early formative decade saw several interdisciplinary seminars and conferences that brought psychologists and linguists together as well as several key publications (e.g., George Miller’s textbook *Language and Communication*, 1951; Karl Lashley’s article, "The Problem of Serial Order in Behavior," 1951). In 1959 Noam Chomsky published a review of B. F. Skinner’s book (also see Chapter 1), arguing that language acquisition and processing cannot be adequately explained by behaviorist principles alone. This review and Chomsky’s book *Syntactic Structures* (1957) laid the groundwork that would radically change the fields of both linguistics and psychology. Research in the 1960s was largely focused on looking for evidence of the psychological reality of Chomsky’s proposed generative and transformational grammar. From the early 1970s to today, the field has become increasingly more interdisciplinary. Linguistic theory no longer plays the dominant role. Instead, the field has embraced theories and traditions in a wide variety of areas, including cognitive psychology, linguistics, artificial intelligence, philosophy, and neuroscience.

The following sections briefly describe research in language comprehension and production separately and then describe approaches that examine production and comprehension processes working together.

### Language Comprehensions

Someone understanding language (whom we will call the “comprehender”) is largely at the mercy of the language producer. The comprehender’s job is to try to reconstruct the intended meaning of the speaker (or writer). Normally this process feels very easy, but when you consider the potential for ambiguity, at nearly every level of representation, it is amazing that we ever understand anything at all. Consider what you have to do to understand the sentence “The cat chased the rat.” You must identify the sounds (or letters) that make up the words, retrieve their meanings, build the appropriate syntactic structure, and then combine that information into an overall interpretation of the sentence (see Figure 9.3). In the following subsections we briefly review these major subareas of research in language comprehension.

#### Figure 9.3 An Overview of Language Comprehension

![Diagram of language comprehension](Image)

To comprehend a sentence like “The cat chased the rat,” we must process the incoming information at several levels of linguistic analysis.
Language Perception

Most of the language that we try to understand is either spoken or written (leaving aside things like sign language or reading braille). Whether we are reading or listening to the sentence we usually come to the same interpretation, so it might be easy to assume that comprehension processes for spoken and written language are the same. However, there are some big differences between the two systems, at least at the initial stages. Consider the problem that Bill and Ted had with the sentence “The stuff he knows is dangerous.” When the sentence is written it isn’t ambiguous. However, in the story the sentence is spoken, which resulted in Bill’s misinterpretation of the sentence as “The stuffy nose is dangerous.” In many ways written language is much clearer than spoken language.

Written language is perceived by the visual system. It is typically persistent (outside of TV news crawls, the words stay visible on the page), letters are typically distinct from one another, there are spaces between words, and some words that sound identical are spelled differently. In contrast, spoken language comes in via the ears. It is transient (it unfolds over time and then fades away), and the phonemes and words don’t typically have gaps between them. In fact, the sounds often overlap to some degree, called coarticulation. Figure 9.4 shows the sound spectrograms for “The stuff he knows is dangerous” and “The stuffy nose is annoying.” Notice that the words all seem to blend together such that it is difficult to see where one word ends and the next begins. Despite this, we are able to recognize language spoken by different people with different voices speaking at different levels (e.g., whispering, talking, or yelling) and different rates. These variables result in different acoustic properties of the language (in the case of written language, consider different handwritings and fonts). It turns out that it is very difficult to identify particular core common features that correspond to particular phonemes. This is referred to as the invariance problem. So, given the complexity and variability of the language input, how do we identify the phonemes in the signal?

Figure 9.4 Sound Spectrograms of Speech

The top frame is a sound spectrogram of “The stuff he knows is dangerous,” while the bottom is “The stuffy nose is annoying.” Notice that the acoustic signal for the first part of the two sentences looks very similar. There is not much information in the signal to disambiguate the two different sentences. Notice also the lack of clear word boundaries. It is difficult to tell where one word ends and the next begins.

SOURCE: Copyright © 2014 members of the Audacity development team.

Coarticulation: an issue in language comprehension due to the overlapping of sounds in spoken language.

Invariance problem: an issue in language comprehension due to variation in how phonemes are produced.
Research suggests that we treat language-related stimuli differently from other stimuli. While most sounds are perceived along a continuum, speech is perceived as discrete categories (Liberman, Harris, Eimas, Lisker, & Bastian, 1961; Liberman, Harris, Hoffman, & Griffith, 1957). This process is called **categorical perception**. Eimas and colleagues (Eimas, Siqueland, Jusczyk, & Vigorito, 1971) demonstrated that infants (one- and four-month-olds) can distinguish between different phonemes. In fact, it is generally believed that infants can distinguish between most phonological units of the world's languages. Interestingly, as they experience language, they typically lose their ability to make distinctions between phonological contrasts not found in their native language contexts. For example, English makes a distinction between /r/ and /l/ phonemes, but Japanese does not. In Japanese, the /r/ and /l/ sounds are considered part of the same category of speech. Young Japanese infants can make the distinction; however, as they get older and experience Japanese in their environments, they lose that ability (e.g., Iverson et al., 2003; Kuhl et al., 2006).

Another processing feature of language perception is top-down contextual information. In other words, we use information we already know about words to help us interpret incoming language. Consider the “letter” in the top part of Figure 9.5. Is it an A or an H? Now consider it in the bottom half of the figure. Most people interpret it as an H when in THE and as an A when in CAT. Letters are easier to identify if they are embedded within words compared to nonword contexts or in isolation. This is called the word superiority effect (see Figure 9.6; Reicher, 1969). Contextual information can even be used to fill in missing information. Richard Warren (1970) presented listeners with the sentence “The state governors met with their respective legislatures convening in the capital city,” but he removed the first /s/ in legislatures and replaced it with a cough. The listener’s task was to report where the cough had occurred and whether he or she noticed anything else about the sentence. Not only were his participants unable to correctly identify the location of the cough, none of them noticed that the /s/ was missing. Instead, they used their knowledge of the word to “fill in” the missing input. This is known as the **phoneme restoration effect**.

Explanations of both the word superiority effect and the phoneme restoration effect rely on having mental representations of words. It is estimated that people generally know from 40,000 to 60,000 words (Aitchison, 2003). The collection of the representations of these words in our long-term memory is called the mental lexicon. The next section reviews some of the processes and effects involved with recognizing and retrieving information from the lexicon.

---

**Figure 9.5 Top-Down Effects in Letter Recognition**

In the top frame is an ambiguous A/H figure. When presented in different contexts, people typically perceive it as either an H (as in THE) or an A (as in CAT).

---
Lexical Recognition and Access

Imagine that you are reading something in a language that you don’t know, but you have access to a translation dictionary with 50,000 words in it. Looking up each word would make reading the sentence take a long time. Now consider how long it took you to read this sentence in a language that you know, using your own mental dictionary. You were much faster. Research has shown that it takes as little as 200 milliseconds to recognize a word (in the case of spoken language, this may be even before the end of the word is heard). Recognition of the word is only part of the issue; following recognition we access the word (Balota, 1990; Balota & Chumbley, 1984). If we consider a dictionary as our metaphor for the lexicon, recognition would be when we find the word. Access would correspond to reading the entry for the word, which would typically include how it is pronounced, what part of speech it is, and what its meanings are. The speed (and accuracy) with which we can accomplish this feat is a function of how we organize our lexicons and the extent to which we can use contextual information.

A typical dictionary is arranged in alphabetical order; however, research suggests that our mental lexicon is organized along many other dimensions (see Figure 9.7). Researchers have identified a wide variety of variables that impact word recognition. Perhaps the most powerful variable is how often a word is used—its lexical frequency. The more frequently a word is used, the faster that word is recognized (e.g., Monsell, 1991). “Neighboring” words (those that have similar orthographic spellings) also affect recognition. Generally, words with large neighborhoods (the set of words that differ by changing one letter) have more competition and take longer to recognize. Morphologically complex words take longer than morphologically simple words (hunter made up of hunt and -er takes longer to recognize than daughter, which is a single morpheme).

Stop and Think

9.4. What are the major differences between spoken and written language?

9.5. How are speech sounds processed differently from other kinds of sounds?

9.6. What processing features may be used to help understand degraded stimuli (e.g., reading a faded photocopy or understanding somebody speaking with a stuffy nose)?
Properties of the word alone are not the only factors that affect word recognition; the context in which they occur also matters. Words may be “primed” by other words (see Figure 9.8). Meyer and Schvaneveldt (1971) presented participants with a list of strings of letters and asked them to respond with a yes or no as to whether they were real words (a lexical-decision task). They demonstrated that people recognize a word (e.g., doctor) faster if it is preceded by a semantically related word (e.g., nurse) compared to an unrelated word (e.g., shoes). Similar results have been found for phonological and orthographic (spelling) primes.
Context also impacts access to a word’s meaning. David Swinney (1979) used a cross-modal priming task in which participants listened to a sentence while watching a display. At some point during the sentence, a string of letters appeared on the display and the participants performed the lexical-decision task. Consider the following critical passage: “Rumor had it that, for years, the government building has been plagued with problems. The man was not surprised when he found several spiders, roaches, and other bugs in the corner of his room.” Shortly after hearing the word bugs participants saw either ant (related to the insect meaning of bug, consistent with the context), spy (related to the listening-device meaning of bug, inconsistent with the context), or sew (unrelated to any meanings of bug). If the word appeared right after the word bugs, then both ant and spy were recognized faster than sew. However, if the word appeared 200 milliseconds after bugs, only ant was primed (spy and sew were the same). This shows that initially both meanings were accessed, but shortly afterward, only the contextually appropriate meaning remained. Further research (e.g., Rayner & Duffy, 1986; Simpson & Burgess, 1985) has found that the relative frequency of the word’s meanings turns out to be an important variable.

As we have seen, a number of variables impact how quickly we recognize a word and access its meaning. One of these important factors is the sentence context in which words appear. This shouldn’t be surprising since most words aren’t understood in isolation but rather as they appear in sentences. The next section describes some of the theory and research examining how we process sentences.

**Interpreting Sentences: Syntactic Analysis**

As discussed earlier, recognizing and accessing words is not the end of comprehension. Building the syntactic structure (called syntactic parsing) impacts how a sentence is understood. Consider our headline “Enraged cow injures farmer with ax.” As we hear or read this sentence, how do we decide which structure (see Figure 9.1 again) we build? Early approaches suggested that we primarily use syntactic information to make this decision.

Chomsky (1957, 1965) made a distinction between deep structure (derived from phrase structure rules like those discussed earlier) and surface structure (the linear order that actually gets produced). Transformations of the deep structure (e.g., adding, deleting, and moving syntactic constituents) result in the final surface structure. These processes were proposed to explain things like why the active sentence “the dog bit the man” and the passive sentence “the man was bitten by the dog” can mean the same thing. They both come from the same underlying deep structure, but the passive version has undergone a transformation to make it passive (the passive transformation rule would be something like this: move the second NP the man to the front, add was before the verb, add -en to the verb, and add by before the first NP the dog). While details of Chomsky’s theory have changed, it set the stage for much of the research on how we syntactically analyze a sentence.

One of the most influential advances in this area of research was the development of technology and procedures to measure eye movements during reading (see Photo 9.3). As we read, our eyes don’t smoothly scan across the page. Instead, they jump from fixation to fixation. Researchers began to use the pattern of these fixations and movements to measure “online sentence” comprehension. More recently, researchers have begun using electrophysiological techniques as well. Of particular interest is what happens when readers encounter positions within a sentence where the syntax is ambiguous.
Two general theoretical approaches have been considered. The syntax-first approach (e.g., Frazier, 1987; Frazier & Fodor, 1978) proposed that we construct one syntactic structure based on a set of parsing principles that focus on syntactic information alone. That structure is then evaluated against the semantics and context and revised if it does not make sense. A key prediction of this approach was that the parsing process computes the initial syntactic structure based entirely on syntactic information and that contextual and semantic information is only used afterward.

One of the syntactic principles proposed was that simpler structures are preferred over complex ones. Consider Figure 9.1 again. The syntactic structure on the left (9.1a) is considered the simpler structure because it has fewer branching points or nodes (six constituents), while the one on the right (9.1b) is more complex because it has more nodes (seven constituents). To test the predictions of this syntax-first approach, Rayner, Carlson, and Frazier (1983) had people read sentences similar to our ambiguous headline. Consider the following pair of sentences.

a. The spy saw the cop with the binoculars, but the cop didn’t see him.

b. The spy saw the cop with the revolver, but the cop didn’t see him.

The syntax-first approach predicted that when participants read these sentences, in both (a) and (b) the initial syntactic structure built should be the simpler one in which the prepositional phrase (“with the binoculars/revolver”) modifies the verb phrase (“the spy saw with the binoculars/revolver”). In the case of sentence version (a), this interpretation makes sense (the spy saw with the binoculars). However, this structure does not make sense in version (b) (the spy saw with the revolver), which should result in a slowing down of reading when getting to the “but the cop didn’t see him” part of the sentence. Rayner et al.’s results were consistent with this prediction (see Figure 9.9).

Central to the syntax-first approach is the idea that the initial structure is based on syntactic properties alone. However, a number of research findings led to the development of an alternative interactive approach that suggests that other variables may influence the initial parse (e.g., Altmann, 1998; Gibson & Pearlmutter, 1998). For example, Taraban and McClelland (1988) presented readers sentences with the same syntactic ambiguity that Rayner et al. (1983) investigated but with stronger semantic information biasing the more complex interpretation.

c. The police arrested the mastermind behind the hideout, but they forgot to read him his rights. (simpler structure)

d. The police arrested the mastermind behind the crimes, but they forgot to read him his rights. (more complex structure)

Using these sentences, they found the opposite pattern of results. Even though sentence (d) has a more complex syntactic structure, reading times were faster in (d) than in (c). Results like these demonstrate strong support for the interactive approach to syntactic analysis.

The review of research in this section demonstrates that understanding the meaning of a sentence involves more than just knowing the meanings of the words. The underlying syntactic structure plays an important role as well. Similarly, meaning may not end with the
meaning of isolated sentences. Most of the time we are not trying to comprehend sentences in isolation but are trying to understand a series of related sentences, paragraphs, and entire stories. The next section describes some of the processes we use to build structures within texts.

Beyond the Sentence: Texts and Discourse

Consider our opening story again. Understanding the entire passage involves building structures larger than individual sentences. At one point Ted says, “Man, the stuff he knows is dangerous.” Who does he refer to? Within that sentence there is not a good candidate. However, if you look back a sentence, you will see that there are two people mentioned. In this case he refers to Rufus and not Elizabeth because of the male pronoun used. But suppose that the preceding sentence contained “Rufus and Bach” instead. How do we decide which person he refers to in that case?

Arnold, Eisenband, Brown-Schmidt, and Trueswell (2000) used an eye-tracking procedure to investigate the kinds of cues used to figure out the correct pronoun antecedents. They had people listen to sentences like “Donald is bringing some mail to Mickey [or Minnie] while a violent storm is beginning. He’s [or She’s] carrying an umbrella, and it looks like they’re both going to need it.” Simultaneously, they were looking at pictures like those in Figure 9.10 (in all of the pictures the correct antecedent is the character with the umbrella). They examined two cues: gender (if both characters are the same or different gender) and accessibility (typically things mentioned first are more accessible than things mentioned second). They monitored what people looked at in the pictures as they heard the sentences. Their results, presented in Figure 9.11, showed that people use both gender and accessibility information to quickly determine the correct antecedent for the pronouns.

Stop and Think

9.9. What is the difference between deep and surface structure? What are syntactic transformations?

9.10. How do the syntax-first and interactive approaches differ with respect to resolving syntactic ambiguity?
Using a pronoun to refer back to something in another sentence (called anaphoric inference) is one way in which we use inferences to bind sentences together into cohesive texts (things make sense from one sentence to the next) and coherent structure in discourse (the whole text makes sense with what we know about the world). Inferences are interpretations of the story that go beyond what is actually stated. They help tie the story together, forming a cohesive whole, rather than a list of disconnected sentences. There are many types of inferences. Consider the following:

Dick and Harry got the picnic basket out of the trunk. The beer was warm. Dick accidentally shot his hunting partner later that afternoon.

We may make the following inferences: the beer was warm because it was in the trunk; the trunk was hot because it was a hot day; Dick and Harry drank the beer; Dick had a gun; and Dick and Harry went hunting together. Not all types of inferences are automatically generated as we comprehend the text, and others may only be made when we later recall the story (McKoon & Ratcliff, 1992; Singer, 1994). Furthermore, a comprehender’s goal (e.g., why somebody is reading a particular text) has also been demonstrated to have an impact on what inferences are made (VanDenBroek, Lorch, Linderholm, & Gustafson, 2001). For example, if you read a book to study for an exam, you may emphasize drawing inferences that draw connections within the text. In contrast, when reading for pleasure, you may instead emphasize drawing connections with your own experiences or background knowledge.
The end product of comprehension is a mental representation of what the entire discourse is about. This representation is typically called a mental or situational model (Johnson-Laird, 1983; Zwaan & Radvansky, 1998). Think back to our opening story. What is your interpretation? Is it an image of two guys sitting in a coffee shop, sipping their drinks, discussing events that one of them had experienced earlier? One way to think of a situational model is as a mental simulation of the events evoked by the language that you understood. This is somewhat different from the scripts and schemata discussed in Chapters 7 and 10. Those are mental representations of stereotypical events. The situational model is a mental representation of the current interpretation, which may be influenced by inferences drawn from a schema. For example, our story does not describe the coffee shop where Bill and Ted are talking, but your situational model may contain inferences drawn from what you expect at a typical coffee shop (e.g., Bill and Ted are sitting at a small table with two chairs, Ted has a cup of coffee in front of him).

Support for this approach comes from findings suggesting that language comprehension is tightly connected to the perceptual representations of the situations described. Zwaan, Stanfield, and Yaxley (2002) had comprehenders read a sentence that implied something about the shape of objects (e.g., “The egg was in the carton” or “The egg fried in the pan”) and then presented a picture of an object to be named (see Figure 9.12). They found that naming times were slower if the implied shape in the sentence mismatched the pictured shape than if they matched. Glenberg and Kaschak (2002) found evidence that the situational model may include action components. They presented comprehenders with sentences that implied action in a particular direction (e.g., “close
the drawer” implies action away from you, while “open the drawer” implies action toward you). Their comprehenders had to decide whether the sentences made sense and indicate their decision either by pressing a button close to them or away from them. Sentences that implied action in a direction consistent with the response direction (e.g., if “yes it makes sense” was a button close to the body and the sentence was “open the drawer”) were verified faster than those that were inconsistent.

Recent neurophysiological results also support this approach. For example, using fMRI data (a brain imaging technique; see Chapter 2 for more details), Hauk, Johnsrude, and Pulvermüller (2004) found overlapping brain area activation (see Figure 9.13) when people read action words (e.g., lick, pick, and kick) and when they performed related actions (instructed to move their tongue, finger, and foot).

As the preceding section suggests, understanding language involves a complex set of procedures. In many ways we are at the mercy of the speakers and authors. They know what thoughts they are trying to convey, and our job, as comprehenders, is to try to recover that meaning from what they produce. As readers we have to be able to recognize different fonts and handwriting. As listeners we have to identify sounds spoken by different people at different speeds and with different accents. We have to identify words, select the appropriate meanings of those words, and then figure out how those are syntactically related to each other. Then we need to piece together the sentences, filling in details to make sense of

Stop and Think

9.11. What processes are used to combine separate sentences into a cohesive and coherent structure?

9.12. What is a situational model? What evidence is there that we generate these models during comprehension?
what we've heard or read. Yet despite this complexity, we do so with remarkable speed and without much conscious effort. In the next section we flip the language system upside down and examine the processes involved in producing language.

**Language Production**

It is tempting to assume that language production works the same way as language comprehension in reverse. However, on closer consideration we see that the processes may be different. As we saw in the previous section, comprehension largely is a case of resolving ambiguity (e.g., which sound did you hear, which word was it, which meaning, which syntactic structure). In contrast, when producing language, you are typically in control of the situation. You know what ideas to convey and who your audience is; you can select which words to use, the order to put them in, and the rate at which to speak. There is much less ambiguity to resolve during production. However, there is an interesting paradox. Even though we are in total control, when we make mistakes producing language, it is typically at the expense of meaning, while maintaining the correct form. In other words, even though the purpose of producing an utterance is to convey meaning, the mistakes we make often show disruptions in meaning but appear to obey the rules corresponding to other levels of representation (e.g., syntax and phonology). The following section describes some of the ways in which language production processes have been examined.

**Making Mistakes: Speech Errors**

The processes of language production begin with mapping our thoughts (or message) onto linguistic representations. This presents a challenge to doing research because

---

Figure 9.13 fMRI Images From the Hauk et al. (2004) Study

The frame on the left shows brain activation for three different movements. The frame on the right shows brain activation for reading action-related words.


---
Table 9.1 Types of Speech Errors

<table>
<thead>
<tr>
<th>Type</th>
<th>Examples: Intended → Produced Error</th>
<th>Unit Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exchange</td>
<td>York library → &quot;lork yibary&quot;</td>
<td>Phoneme</td>
</tr>
<tr>
<td>Anticipation</td>
<td>Reading list → &quot;leading list&quot;</td>
<td>Phoneme</td>
</tr>
<tr>
<td>Perseveration</td>
<td>Beef noodle → &quot;beef needle&quot;</td>
<td>Phoneme</td>
</tr>
<tr>
<td>Addition</td>
<td>Blue bug → &quot;blue bugh&quot;</td>
<td>Phoneme</td>
</tr>
<tr>
<td>Shift</td>
<td>Black boxes → &quot;back boxesh&quot;</td>
<td>Phoneme</td>
</tr>
<tr>
<td>Deletion</td>
<td>Same state → &quot;same state&quot;</td>
<td>Phoneme</td>
</tr>
<tr>
<td>Morpheme Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exchange</td>
<td>Thinly sliced → &quot;slicely thinned&quot;</td>
<td>Stem</td>
</tr>
<tr>
<td>Anticipation</td>
<td>My car towed → &quot;my tow towed&quot;</td>
<td>Stem</td>
</tr>
<tr>
<td>Perseveration</td>
<td>Explain … → &quot;rule exsertion&quot;</td>
<td>Prefix</td>
</tr>
<tr>
<td>Addition</td>
<td>Some weeks → &quot;somes weeks&quot;</td>
<td>Inflectional suffix</td>
</tr>
<tr>
<td>Shift</td>
<td>Gets it → &quot;get its&quot;</td>
<td>Inflectional suffix</td>
</tr>
<tr>
<td>Word Errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exchange</td>
<td>Writing a letter to my mother → &quot;writing a mother to my letter&quot;</td>
<td>Noun</td>
</tr>
<tr>
<td>Anticipation</td>
<td>Sun is in the sky → &quot;sky is in the sky&quot;</td>
<td>Noun</td>
</tr>
<tr>
<td>Perseveration</td>
<td>Class will → &quot;discussing the class&quot;</td>
<td>Noun</td>
</tr>
<tr>
<td>Addition</td>
<td>These flowers are purple → &quot;these purple flowers are purple&quot;</td>
<td>Adjective</td>
</tr>
<tr>
<td>Shift</td>
<td>Something to tell you all → &quot;something all to tell you&quot;</td>
<td>Quantifier</td>
</tr>
</tbody>
</table>

SOURCE: Dell (1986).

manipulating thought in carefully controlled experimental designs can be difficult (Bock, 1996). However, while the input to language production is difficult to manipulate, the output of the process, produced language, provides a rich set of data for analysis. Sit down and listen to a conversation going on nearby. Generally you will find that our productions are remarkably fluent and accurate. However, they do contain false starts, hesitations, “ums” and “uhs,” and from time to time mistakes (Erard, 2007). While speech errors are often referred to as “slips of the tongue,” analyses of the pattern of errors suggest that most of them result from higher-level processes rather than the motor control of the tongue. Collections of speech errors have provided the foundation of most theories of language production (e.g., Dell, 1986; Fromkin, 1971; Garrett, 1975).

Table 9.1 gives a sampling of the wide variety of speech errors that have been collected and analyzed. Errors appear to involve all of the linguistic units we have discussed in this chapter. Researchers have noted several regularities in the errors we make.
For example, errors resulting from the interaction between two representations appear to be constrained to interacting elements of the same type. In other words, most word errors involve words from the same syntactic category. In sound errors, most vowels interact with other vowels and consonants with other consonants. Regularities like these suggest that words, syntax, and sounds are likely to be processed separately during production. Analyses have also revealed that sound errors typically involve elements relatively close together, while word errors may involve words farther apart, again suggesting that lexical and phonological processing may occur separately. Another regularity is that sound errors result in actual words more often than expected by chance. Furthermore, when these sound errors result in nonwords, the nonwords typically conform to the phonological rules of the speaker’s language. These regularities suggest that the interaction between lexical and phonological processes may be complex.

These kinds of regularities led Merrill Garrett and others to propose that language production proceeds through a series of processes: conceptualization, formulation, and articulation (e.g., Garrett, 1975, 1988; Levelt, 1989). Conceptualization is the level of thought, where we piece together the nonverbal situational model (the “message”) we want to talk about (see Figure 9.14). Formulation is the grammatical processing stage during which the message is mapped onto linguistic units. Garrett proposed that this happens in two stages. First is the functional stage of processing where we select semantically appropriate lexical items and assign them to functional roles (e.g., subject, verb, object). Following this is the positional stage, during which a syntactic structure corresponding to the functional roles is built and lexical items are inserted into the syntactic structure. Following the positional stage, the form information (e.g., sounds, spellings) of the words is specified, and finally we articulate (say or write) the utterance. The theory predicts a separation of semantic,
syntactic, and phonological processes during production. The separation of the meaning from the syntax and form explains the production paradox. Since meaning processing is completed earliest, errors resulting late in the production process may disrupt meaning while conforming to syntactic and form constraints.

Separation of Semantics, Syntax, and Form

As in language comprehension research, much of the research in language production has focused on the nature of the relationship among semantic, syntactic, and form processing. You may already have had an experience supporting the idea that semantics and form processing are separated during language production. Think about a time during which you found yourself having trouble remembering the name of a famous movie star or the word for some obscure object. Often when in this state you may know that you know the word, and remember some details of the word, but you just cannot retrieve it at that moment. This feeling is called the tip-of-the-tongue state (James, 1890). It is thought to reflect a state in which you have accessed the semantic and syntactic representations of a word but not the phonological form of the word. Vigliocco, Antonini, and Garrett (1997) demonstrated that when Italian speakers were in this state, they had access to semantic and syntactic features of words (Italian specifies the grammatical genders masculine and feminine for nouns) but not the letters and phonemes of the words.

Schriefers, Meyer, and Levelt (1990) provided an experimental demonstration of the separation of semantic and phonological processing during production. They presented speakers with object pictures to be named, while simultaneously ignoring a word they heard over headphones. They manipulated two variables: the relationship between the
pictures and the words (if the picture was a dog, then the interfering words could have been *dot*, phonologically related; *cat*, semantically related; or *ship*, unrelated) and when

the interfering word appeared relative to the picture (150 ms before the picture, at the same time as the picture, or 150 ms after the picture). The results of their experiment are presented in Figure 9.15. When the distractor word was presented early, there were clear effects of hearing a semantically related word but not a phonologically related word (relative to the unrelated control word condition). The pattern is different when the distractor was presented later: no effect of a semantically related word but an effect of a phonologically related word. However, further research has suggested that this division may not be so distinct (e.g., Cutting & Ferreira, 1999; Damian & Martin, 1999; Peterson & Savoy, 1998).

Experiments examining agreement processes (e.g., subject-verb agreement, pronoun gender agreement) have been used to explore the separation of semantic and syntactic levels of processing in production. Consider the following sentence.

a. The knife is on the table.

At first it may seem straightforward. The subject of the sentence is singular (referring to a knife), so you need to use the singular verb form *is*. But now consider the sentence in (b).

b. The scissors are on the table.

In this sentence the subject refers to a single pair of scissors, but the correctly agreeing verb form is *are.* Bipartite words (*things having two parts*, e.g., *scissors, pliers, and pants*) demonstrate that the property of the plural form has both a semantic and syntactic component to it. To examine how these semantic and syntactic components are used in production, Kathryn Bock and colleagues (e.g., Bock & Eberhard, 1993; Bock & Miller, 1991; Bock, Eberhard, Cutting, Meyer, & Schriefers, 2001; Humphrey & Bock, 2005) used a sentence completion task in which speakers were presented with the beginning of a sentence up to but not including the verb.

c. The cutting board under the knives...
d. The cutting board under the scissors...

The speaker’s task was to repeat and complete the sentence fragment. When presented with sentences like those presented in (c) and (d), speakers made errors and completed the sentences using a plural verb (e.g., “The cutting board under the knives are getting old”). Instead of using a verb that agreed with the subject noun (cutting board), the participants sometimes used a plural verb agreeing with the plural “local noun” (knives or scissors) from the prepositional phrase. The results, shown in Figure 9.16, suggest that notional and grammatical number agreement processes operate differently at separate stages during production.

Research examining production and comprehension of language suggests that many of the same levels of representation may be involved in both behaviors. However, because comprehenders and producers start with different inputs (spoken or written language versus thought), the processes involved may operate differently. Most research has focused on either comprehension or production as independent processes. Recently, some researchers have begun examining how production and comprehension processes are directly related to each other.

---

**Stop and Think**

9.13. What is the paradox of production?

9.14. What evidence has been used to suggest that semantic and phonological processing are separated during language production?
Dialogue: Production and Comprehension Together

Up until this point our discussion has focused on either language comprehension or production in isolation. However, if we look back to our opening story, it is a dialogue between Bill and Ted. This is the typical situation in which language is used, with multiple individuals taking turns as comprehenders and producers. Herb Clark (1996) characterized using language as a joint action akin to dancing. Like dancing, language users need to coordinate their linguistic actions, often rapidly switching between roles as speakers and listeners, to successfully communicate ideas. So how are comprehension and production processes related to each other?

As noted in the previous section, our utterances are generally fluent and accurate. One proposal to account for this is that we use our comprehension system to monitor our ongoing productions, not only after we have said them but also before their actual articulation. Levelt (1983) called this approach the perceptual loop. The theory is that even before we articulate our planned utterances, we run our “inner speech plan” through our comprehension system to look for errors so that we can make corrections before articulation. Zenzi Griffin (2004) reported findings consistent with this proposal. She monitored speakers’ eye movements while they described pictures. She found that speakers gazed at named objects longer when they named them incorrectly and then corrected the error. However, recent evidence suggests that comprehension alone may not be sufficient for monitoring errors (e.g., Huettig & Hartsuiker, 2010; Nozari, Dell, & Schwartz, 2011; Vigliocco & Hartsuiker, 2002). Even our disfluencies

Figure 9.16: Results of Bock et al.’s (2001) Study

Proportion of completions using a plural noun (e.g., finishing the sentence fragment “The cutting board under the scissors” with “are dirty.”)
(e.g., “ums” and “uhs”) may be meaningful. Fox Tree (2001) has suggested that the “ums” and “uhs” we produce may be used to signal comprehenders of upcoming production difficulties. Research findings like these provide strong support for the notion that we may use our production and comprehension processes together during language use.

Garrod and Pickering (2004) proposed a framework to describe how production and comprehension processes interact during dialogue. Their theory, called the alignment theory, proposes that the goal of conversation is the alignment of the representations of the speaker and listener. That is, successful communication involves aligning the phonological, syntactic, semantic, and situational models in both participants for successful communication. They proposed that the mechanism through which alignment occurs is priming, like that discussed earlier in the chapter. As people converse, they often repeat sounds, words, syntax, and meanings. Over the course of the dialogue, via priming, the sets of representations activated in the people become more and more coordinated.

Mounting evidence demonstrates that participants in dialogues coordinate semantic, syntactic, and lexical representations. Garrod and Anderson (1987) had pairs of people verbally talk each other through a maze. They found that their participants tended to use the same terms and phrases to refer to where they were in the maze. Branigan, Pickering, and Cleland (1999) demonstrated that people tend to repeat the syntactic structures they use.

Richardson and Dale (2005) provided particularly striking evidence of the coordination between speakers and comprehenders. They monitored the eye movements of speakers describing their favorite scene from the television show Friends while presented with an array of the characters from the show. They also measured the eye movements of a separate group of participants, who viewed the same character array while they listened to recorded descriptions provided by the first group. Results showed a large overlap in the gazes of the two groups. In other words, the listeners typically looked at the same characters at the same time as those who produced the language.

In most situations in which we use language, we act as both a producer and a comprehender. Given that we have the capacity to do both highly related activities, it may not be surprising that the processes are interrelated. Consider that point in our lives when we do not have the capacity to be full-fledged participants in dialogues, when we are infants learning language. How do we acquire our language representations in the first place? The next section provides a brief discussion of research examining how we acquire language.

**Acquiring Language**

After reading the previous sections of this chapter describing how complex language and the mental processes that underlie its use are, you may be amazed at how quickly and effortlessly we use language. What is perhaps even more amazing is how we acquire language in
the first place. If you’ve ever tried to learn a second language as a teenager or an adult, think about how long and hard the process was. Compare that with infants and children learning their first language (shown in Photo 9.5). How do they do it and make it seem so easy? We begin this section with a brief description of typical language development and follow with a summary of theoretical and empirical approaches to the investigation of this behavior.

**Typical Language Development**

There is great uniformity in the pattern of language development across languages and cultures. Newborns enter the world without being able to use language, but evidence suggests they have some experience and knowledge very early on, perhaps even before birth. DeCasper, Lecanuet, Busnel, Granier-Deferre, and Maugeais (1994) had mothers read stories to their fetuses during pregnancy. At the thirty-eighth week a new or old story was read while the fetuses’ heart rates were measured. Heart rates slowed in response to the old stories, demonstrating that even in the womb, fetuses could distinguish between the two stories. Mahler et al. (1988) demonstrated that four-day-old infants could distinguish between spoken French and Russian. As already mentioned earlier in the chapter, one-month-old infants can distinguish most of the phonological contrasts of all the languages of the world. By their sixth month, infants typically can recognize their own names and respond to "no." From six to twelve months they can recognize names of familiar objects, foods, and body parts (Bergelson & Swingley, 2012). From age one to two years, children can point to objects and pictures when named and understand some requests or questions (e.g., "Push the truck" or "Where's the horsey?"). At this age children often exhibit overextension, applying the words they know to more things than adults do (e.g., doggie may be used to refer to all four-legged animals) and underextension (e.g., using car to refer to a particular car rather than all cars). It is estimated that children typically understand nearly three times more words than they produce at this stage. Vocabulary growth continues rapidly, and by the third year the vocabulary gap between production and comprehension narrows. Also by the third year they can answer who, what, and where questions.

Language production typically lags behind language comprehension, in areas other than vocabulary as well. Early vocal behavior consists of nonlinguistic vegetative sounds (e.g., crying, burps, sucking noises), but as early as six weeks, infants begin cooing vowel sounds. By four to five months they begin babbling and producing clear consonant vowel clusters (e.g., ba, gi), followed by reduplicated babbling (e.g., baba, gigi). By ten months infants begin to show more complex babbling, combining sounds into incomprehensible wordlike utterances (e.g., dab gogotah), and by twelve months their utterances may be showing evidence of the phonological rules of their environmental language. At this point they may begin to use their first words, often pointing at things to which they refer. These words are usually unique to the child, rather than fully formed adult forms (e.g., baba for bottle). By their second birthday they may use two hundred to three hundred words (typically focused on the "here and now," like important people and objects that can be moved or manipulated), and by their third birthday they can use one thousand words. Nouns typically appear before verbs. Children’s utterances initially consist of single words, but in their second year they start to combine words to produce longer “telegraphic” speech, leaving out grammatical words (e.g., articles like the and a and prepositions like by and for). By their third
year their utterances continue to get longer and more complex. They typically use full sentences and can form questions, make negative statements, and use grammatical morphemes.

Infants don’t learn language in isolation; instead they are typically actively engaged in highly interactive and social contexts. Parents talk to and play with their children (Hart & Risley, 1995, estimated that parents direct three hundred to four hundred utterances an hour to their children). Speech directed at infants (called child-directed speech) typically differs from that between adults. It typically has fewer words, less complex syntax, more repetition, and exaggerated prosodic structure (higher pitched, slower, longer pauses, and distinctive contours). These speech simplifications provide infants important cues that assist in their language learning (Dominey & Dodane, 2004). Eye contact and smiling provide strong social cues. Snow (1977) has argued that as early as one month infants learn some basic turn-taking rules of conversation through the playful “dialogues” with their caregivers (e.g., “Ooh, was that a burp? Are you burping at me?” The parent then pauses and waits until the infant makes another vocalization. “You were! You were burping at me.”). Parent-infant interactions provide a rich environment, full of linguistic information to help the child learn language.

As children get older, their language use gets more sophisticated. They continue their vocabulary explosion (e.g., by the age of six they may have a vocabulary of 14,000 words), their utterances get longer, and their syntax grows in complexity. Additionally, they may begin to learn a new medium of language use: reading. Overall, the speed and apparently effortless nature of our ability to use language is amazing. The next sections review some of the theoretical approaches taken to explain how we are able to do it.

**Nature or Nurture: Mechanisms for Learning Words and Syntax**

How we acquire language is a matter of ongoing debate. Some approaches place the theoretical emphasis on experience (nurture) while others focus on biological predisposition (nature). The behaviorist approach, as advanced by B. F. Skinner (1957), theorized that language learning could be explained through principles of reinforced imitation. Chomsky (1959) argued against this explanation because it could not account for the infinite productivity of language; children can comprehend and produce utterances they have never heard before. Chomsky instead proposed that we come innately prewired with knowledge about language and that language acquisition is a maturational process, like learning to walk. Children learn language in a predetermined way when in an appropriate language context. These two approaches exemplify two extremes of the nature versus nurture debate. Somewhere between these two extremes are the interactionist approaches (e.g., Golinkoff, Mervis, & Hirsh-Pasek, 1994; Markman, 1989), which propose that language learning is the result of the interaction between experience and biological predispositions for language and cognition.

For example, Golinkoff, Hirsh-Pasek, and colleagues proposed the emergent coalition model (Hirsh-Pasek, Golinkoff, & Hollich, 2000). The model hypothesizes that early word learning begins associatively but transitions to social and cognitive constraint-driven processes. They argue that infants are born with biases to attend to and integrate attentional (e.g., perceptual salience, temporal contiguity), social (e.g., eye gaze, social context), and linguistic (e.g., grammar, intonation) cues when learning words. Over time, the relative importance of these cues may change. In a series of studies (Hollich et al., 2000; Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006), they presented two objects to infants (ten, twelve, and twenty-four months old) with a person situated between the objects.
During the learning phase of the experiment, the researchers manipulated the perceptual salience of the objects: one object was very interesting (e.g., brightly colored, moving parts) while the other was less interesting (e.g., dull color, stationary). They also manipulated the social cues by having the person naming the object (e.g., “Look a modi!”) stare at one of the objects (see Figure 9.17). This allowed the researchers to see how the attentional and social cues interact during learning. Following the learning phase, they tested whether the infant had learned the name. This was done in three ways: (1) they presented the two objects and asked the child to look at the object with the label (e.g., “Can you find the modi?”), (2) they presented a “new label” to see if the infant would look away, and (3) they mentioned the original label to see if the infant’s gaze returned to the object. The results showed that ten-month-old infants used only the perceptual salience to connect the
name to the object, twelve-month-old infants learned the name only when the perceptual and social cues aligned, and twenty-four-month-olds learned the names using only the social cues. Brandone, Pence, Golinkoff, and Hirsh-Pasek (2007) used a similar methodology to examine verb learning. They found that two-year-olds were able to learn new verbs when perceptual cues (whether an action produced a result like a sound or a light) and speaker cues (both linguistic and social cues) matched but not when they mismatched.

Results like these suggest that children have cognitive biases that interact with a rich linguistic and social environment in which they learn language. What does this suggest about the uniqueness of language for humans? Can animals use language? The final section of this chapter examines this question.

**Human Language and Animal Communication**

We started the chapter by asking the question “What is language?” Part of the answer is that it is a way to exchange information or communicate. Humans and animals use many ways to communicate (e.g., pheromones, gestures, facial expressions, body language). Many of us probably talk to our pets but realize that interaction is not the same as talking with another person. In fact, most researchers believe that full-fledged language use is unique to humans. This final section of the chapter begins by comparing human language to animal communication and ends with recent attempts to teach animals human language.

**Comparing Human Language to Animal Communication**

There have been many attempts to define the unique characteristics of human language. Hockett (1960) outlined a set of thirteen design features of communication (see Table 9.2 for a complete list). He proposed that although different systems of animal communication may include some of these features, only human language includes all of them. These features include aspects of language related to issues we have discussed earlier in the chapter: productivity, semanticity, arbitrariness, duality of patterning, and traditional transmission. Hauser, Chomsky, and Fitch (2002) have proposed that the minimum distinguishing characteristic of human language is recursive syntax. Recursion occurs when a rule calls for a version of itself. For example, consider the phrase structure rules we discussed earlier. A noun phrase includes a noun that may be modified by an article, adjective, or prepositional phrase. A prepositional phrase is made up of a preposition and a noun phrase. Recursion results because the noun phrase can contain a prepositional phrase, which can in turn contain a noun phrase. So how do systems of animal communication stack up?

As every dog owner knows, dogs bark, but are they “saying” anything? Most researchers agree that the functions of barking are primarily for warning, territory marking, defense, and protest. Pongrácz, Molnár, and Miklósi (2006) found that people are able to use acoustic properties of dog barks to categorize them as aggressive or happy and playful. While this evidence suggests that barking may serve a communicative role, it falls far short of the complexities exhibited by language. Perhaps surprisingly, bees exhibit a system that shares more features with humans. Honeybees dance to communicate the location of nectar sources (von Frisch, 1967). The angle of the dance indicates the direction, and the rate of looping indicates the distance. The bee system of communication exhibits some features (e.g., displacement, semanticity, and productivity) but not others (e.g., discreteness, arbitrariness, and duality of patterning). Perhaps the system of animal communication that comes closest to human language is that of songbirds. Many birds use calls to signal particular behaviors (e.g., warning alarm, coming in for a landing); others also
use songs. Songs, typically limited to males, are used to attract females and repel other males of the same species. The songs are structurally complex, made up of individual notes combined into ordered subparts. However, whereas a hallmark of human language is how word order and syntax are associated with meaning, variations in birdsongs have not been demonstrated to reflect differences in meaning. Gentner, Fenn, Margoliash, and Nusbaum (2006) have demonstrated that European starlings could be trained to distinguish between song sequences containing recursive and nonrecursive structures. However, the ability to distinguish recursion does not demonstrate that starlings can use recursion in their songs (Corballis, 2010).

There is little convincing evidence that the communication systems of animals meet the currently accepted definitions of human language, offering strong support for the notion that language use is unique to human beings. Does this mean that animals can't learn to use language? We address this next.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Bee Dances</th>
<th>Bird Song</th>
<th>Human Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocal-auditory channel</td>
<td>Using the sounds to communicate</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Broadcast transmission and directional reception</td>
<td>Sounds are broadcast widely but may be used to localize the speaker</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rapid fading</td>
<td>The signal disappears quickly after production</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Interchangeability</td>
<td>Users may be both producers and comprehenders</td>
<td>limited</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Total feedback</td>
<td>Speakers have access to their own productions</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Specialization</td>
<td>The speech organ has been adapted for the task</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Semanticsity</td>
<td>The signals have meaning</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Arbitrariness</td>
<td>No direct connection between the sound and the meaning</td>
<td>×</td>
<td>in part</td>
<td>✓</td>
</tr>
<tr>
<td>Discreteness</td>
<td>The signal is made of separate units</td>
<td>×</td>
<td>✓</td>
<td>if semantic</td>
</tr>
<tr>
<td>Displacement</td>
<td>Can communicate about things removed from time and space</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>Productivity</td>
<td>Can use unique combinations that have not occurred before</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>Traditional transmission</td>
<td>Can be taught and learned</td>
<td>×</td>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>Duality of patterning</td>
<td>Combinations of meaningless parts result in meaningful units</td>
<td>×</td>
<td>?</td>
<td>✓</td>
</tr>
</tbody>
</table>

Attempts to Teach Animals Human Language

The human vocal system has evolved to allow for speech (Liberman, 1984). No other animals’ vocal systems are adapted for this capability. Parrots can be taught to mimic human-sounding speech, but mimicking speech isn’t the same as using language. Irene Pepperberg (2009) attempted to teach Alex (see Photo 9.6), an African grey parrot, language. With thirteen years of language instruction, Alex was able to demonstrate some remarkable abilities. Alex had a vocabulary of nearly eighty words, could distinguish between things of different colors and composition, and demonstrated the ability to make some unique combinations of words. Chaser, a border collie, was trained to recognize and distinguish the proper names of more than one thousand objects (Pilley & Reid, 2011). Her trainers have argued that she has an awareness that maps words onto referent objects. Sofia, a mixed-breed dog, can reportedly respond to requests resulting from unique combinations of action and object terms (Ramos & Ades, 2012). But perhaps the most famous and intensive attempts to teach language to animals have involved chimpanzees.

Washoe, a female chimpanzee, was brought up as a human child and taught to use American Sign Language (Gardner & Gardner, 1969). From morning to night, all communication between Washoe and her caregivers was with sign language (sign language was also used between caregivers when in Washoe’s presence). Using daily records of Washoe’s signing, the experimenters estimated that she could use from 150 to 200 signs, from many different syntactic classes. Caregivers argued that she demonstrated behaviors similar to those of human children learning language, including overgeneralization of words, and could create new signs generatively (e.g., combined signs for “water” and “bird” to refer to a duck). Fouts, Fouts, and Van Canfort (1989) reported that Washoe’s adopted son Loulis (she cared for a ten-month-old chimpanzee following the death of her own newborn) learned to use sign language from other signing chimpanzees. Another chimpanzee, Sarah, was taught an artificial language consisting of plastic symbols of different shapes, sizes, and textures (Premack, 1988; Premack & Premack, 1972). Sarah had a “reading” and “writing” vocabulary of nearly 130 words. Researchers claim she was able to understand the words in the absence of their referents, suggesting that she was able to demonstrate key characteristics of language (e.g., semanticity, arbitrariness, and displacement). Sarah could also follow simple written instructions like “insert banana pail” as well as more complex ones like “insert apple pail banana dish” (meaning put the apple into the pail and the banana onto the dish). Kanzi, a male bonobo chimpanzee, learned to communicate with a special keyboard labeled with geometric symbols (Savage-Rumbaugh, 1993; Savage-Rumbaugh, Fields, and

Stop and Think

9.20. What are the characteristic features of human language?
9.21. How do the dances of bees compare to human language?
9.22. How does the performance of chimpanzees taught to use language compare to human children learning language?
As you read the following summary of a research study in psychology, think about the following questions:

1. What aspects of language are being examined in this study?
2. What is the independent variable in this study?
3. What is the dependent variable in this study?
4. What alternative explanations can you come up with to explain the results of this study?

**Study Reference**


**Note:** Experiment 1 of this study is described.

**Purpose of the study:** The authors wanted to determine whether hearing one side of a cell phone conversation is more distracting than listening to the entire conversation. It was hypothesized that hearing only one-half of a conversation puts the listener into a less predictable state, which would in turn impair his or her ability to pay attention and perform a concurrent task.

**Method of the study:** Participants were instructed to complete two tasks: track a moving dot with a computer mouse and respond to letters on a computer screen.

![Figure 9.18 Results of Emberson et al.'s (2010) Experiment 1]

Participants performed more poorly in both the tracking and letter reaction time tasks in the halfalogue than in the dialog conditions. The authors argued that these results were due to conversations that were less predictable in the halfalogue conditions.

SOURCE: Emberson et al. (2010).
Results of the study: Performance on both tasks was worse for halfalogues than dialogues. These results are presented in Figure 9.18.

Conclusions of the study: The authors concluded that because conversations are coordinated behaviors, speech in a halfalogue is less predictable than a complete dialogue. They argued that the decreased predictability of the halfalogue automatically pulled away attentional resources, which resulted in fewer resources and thus poorer performance on the two tasks.

CHAPTER REVIEW

Summary

- **What is language?**
  Language is a system constructed from multiple levels of representations to convey meaning. Each level of representation uses rules to combine elements together to form other representations. These levels of representations include form (spelling and sounds), grammar (syntax), and meaning (morphemes and semantics).

- **How do we get from a string of sounds or marks on a page to something meaningful?**
  The major problem in language comprehension is to resolve potential ambiguities to recover the intended meaning of the producer. This process is accomplished through a series of processing stages using information in the signal as well as contextual information about the words, grammar, and world knowledge.

- **How do we go from thoughts to spoken language?**
  Language production involves levels of representations similar to those in comprehension; however, the system has evolved not to resolve ambiguity but rather to get the form of the output correct. In dialogue, perhaps the most typical way in which we use language, both language production and comprehension processes are involved. Alignment theory proposes that successful communication arises when the participants’ linguistic and situational model representations are aligned. Alignment is achieved largely through automatic priming mechanisms.

- **How do we acquire language?**
  Infants and children learn language rapidly and without explicit instruction. Production abilities tend to lag behind comprehension initially, but the gap is typically closed by the second year. Patterns of acquisition appear to be relatively stable across different individuals and cultures, suggesting to some that humans have an innate ability to learn language. Others believe the acquisition of language results from interactions between cognitive biases and language experience.

- **How does human language differ from animal communication?**
  Animals use systems of communication that share some of the features of human language but not all. Attempts to teach animals to use systems of human language have had limited success.

Chapter Quiz

1. Enter the letter for the correct definition next to the terms below.
   (a) the smallest unit of language that has meaning
   (b) perceiving a continuous stimulus as discrete categories
   (c) a representation of what a text is about
   (d) chunks of syntactic representations
   (e) the sound representations that make up human languages
   (f) building the grammatical structure of a sentence
   (g) the characteristic that words have meaning
2. What does it mean that language is hierarchically structured?
3. What are the phoneme restoration and word superiority effects? What process do they illustrate?
4. What is the syntax-first approach to parsing?
5. What is an inference? How is it used to help with language comprehension?
6. What is the “paradox” in language production?
7. What design feature of language corresponds to the use of unique combinations of representations to produce an infinite number of utterances?
   (a) duality of patterning
   (b) productivity
   (c) innateness
8. Washoe was
   (a) an African grey parrot.
   (b) a child raised in a language-free environment.
   (c) a chimpanzee taught to use human language.
   (d) a speech error demonstrating categorical perception.
9. In the sentence “Connor teased Daphne” the -ed is a
   (a) phoneme.
   (b) bound morpheme.
   (c) free morpheme.
   (d) syntactic constituent.
10. Evidence suggests retrieval of words from the mental lexicon is affected by
    (a) lexical frequency.
    (b) orthographic neighborhoods.
    (c) morphological complexity.
    (d) all of the above.

Key Terms

<table>
<thead>
<tr>
<th>Anaphoric inference 234</th>
<th>Invariance problem 227</th>
<th>Semantics 224</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broca’s aphasia 225</td>
<td>Morphemes 222</td>
<td>Surface structure 231</td>
</tr>
<tr>
<td>Categorical perception 228</td>
<td>Phoneme restoration effect 228</td>
<td>Syntactic parsing 231</td>
</tr>
<tr>
<td>Coarticulation 227</td>
<td>Phonemes 221</td>
<td>Syntax 222</td>
</tr>
<tr>
<td>Deep structure 231</td>
<td>Pragmatics 224</td>
<td>Wernicke's aphasia 225</td>
</tr>
</tbody>
</table>

Stop and Think Answers

9.1. Identify the phonemes in the sentence “Ted quietly chatted with Bill.”
   /t/ /e/ /d/ /k/ /aɪ/ /eɪ/ /æ/ /t/ /eɪ/ /t/ /d/ /æ/ /t/ /l/ /θ/ /b/ /l/ /l/ /l/

9.2. Identify the morphemes in the sentence “Ted quietly chatted with Bill at the coffee shop.”
   Ted quiet –ly chat –ed with Bill at the coffee shop

9.3. What are two different interpretations of the sentence “Groucho shot an elephant in his pajamas”?
   How are the different interpretations linked to syntax?
   In one case, Groucho is wearing his own pajamas. In the other interpretation, the elephant is wearing Groucho’s pajamas. The difference syntactically has to do with what the prepositional phrase “in his pajamas” modifies (either “shot in his pajamas” or “elephant in his pajamas”).

9.4. What are the major differences between spoken and written language?
   Written language is typically persistent, with clear delineations between letters and words, and is processed by the visual system. Spoken language is transient (it fades rapidly over time), without clear boundaries between phonemes and words, and is processed by the auditory system.

9.5. How are speech sounds processed differently from other kinds of sounds?
Most sounds are perceived as continuous. However, speech sounds are perceived as discrete categories.

9.6. What processing features may be used to help understand degraded stimuli (e.g., reading a faded photocopy or understanding somebody speaking with a stuffy nose)?

In addition to using information from the signal itself, contextual information about words and meaning are used to resolve potential ambiguities.

9.7. What factors impact how quickly a word is recognized?

Lexical frequency, morphological complexity, orthographic and phonological neighborhood size, and semantic priming.

9.8. What factors are important in accessing the appropriate meaning of a word?

The context in which a word is used and the frequency of alternative meanings.

9.9. What is the difference between deep and surface structure? What are syntactic transformations?

Deep structure is the syntactic structure formed from meaning through the use of phrase structure rules. Surface structure is the final linear ordering of words in a sentence that result after transformations of the deep structure.

9.10. How do the syntax-first and interactive approaches differ with respect to resolving syntactic ambiguity?

The syntax-first approach uses only syntactic information to build the initial syntactic structure of a sentence. Semantic and contextual information is used afterward to build a new structure if necessary. Interactive approaches use other sources of information (in addition to syntactic information) to build the initial structure.

9.11. What processes are used to combine separate sentences into a cohesive and coherent structure?

Inferences are used to connect sentences together and integrate world knowledge into the ongoing understanding of the text.

9.12. What is a situational model? What evidence is there that we generate them during comprehension?

A situational model is a dynamic representation (a simulation) of the interpretation of the text. Research suggests that the situational model may be perceptual (e.g., orientation information) and action (e.g., direction of movement) aspects inferred by the text.

9.13. What is the paradox of production?

If the producer knows the meaning of what he or she wants to say and is in control of the situation, then why do most speech errors appear to obey syntactic and form regularities at the expense of disruptions in meaning? The answer appears to be that meaning is processed separately and earlier than syntactic and form information.

9.14. What evidence has been used to suggest that semantic and phonological processing are separated during language production?

The tip-of-the-tongue state is an example in which the semantic but not form information has been accessed. Experiments using the picture-word interference task show an early stage of primarily semantic processing followed by a later stage of phonological processing.

9.15. How might we use our comprehension processes to aid our productions?

Evidence suggests that we may use comprehension to monitor what we plan to say, allowing us to detect and repair faulty utterances.

9.16. What is some of the evidence that speakers and listeners align their linguistic representations during dialogue?

The repeated use of words and syntax between participants engaged in dialogue suggests the alignment of our linguistic representations. This is also supported by the coordination of gaze durations between speakers and listeners watching the same visual array of photos during the description of a television show.

9.17. What are the major linguistic milestones of a six-month-old infant? A twelve-month-old infant? A two-year-old child?

From six to twelve months infants can recognize names of familiar objects, foods, and body parts. Six-month-olds typically produce reduplicated babbling, while twelve-month-olds begin to produce their first words. From age one to two years, children can point to objects and pictures when named and understand some requests or questions. By their second birthday, they typically produce two hundred to three hundred words and are beginning to combine the words into short “telegraphic” utterances.

9.18. Why do Chomsky and others propose that much of language acquisition is driven by innate knowledge of language?

Language acquisition appears to follow the same basic pattern across different languages and cultures. This suggests that it may be a maturational rather than
learned process (like walking). Additionally, language is productive, meaning that we can understand and produce sentences we have never experienced before, suggesting that reinforcement of past experiences is not sufficient for language learning.

9.19. **How does the emergent coalition model describe the process of word learning in infants?**

The model proposes that infants initially attend primarily to perceptual and attentional cues early. However, as they get older they use other linguistic and social cues (either in combination or alone). This reflects a developmental shift in the use of relevant cues.

9.20. **What are the characteristic features of human language?**

Table 9.2 lists the thirteen characteristics of human language proposed by Hockett. More recently Chomsky and others have proposed that the presence of recursion in syntax is the hallmark of human language.

9.21. **How do the dances of bees compare to human language?**

The bee system of communication exhibits some features (e.g., displacement, semanticity, and productivity) but not others (e.g., discreteness, arbitrariness, and duality of patterning).

9.22. **How does the performance of chimpanzees taught to use language compare to human children learning language?**

Attempts to teach animals to use systems of human language have had limited success. While animals may learn some words (many fewer than do human children), animals fail to learn to use all but the simplest syntax.