An Introduction to Language Science

1

The rules aren't the ones we were taught in school.

IVAN SAG

One of my favorite language scientists is Daniel L. Everett, a former evangelical Christian missionary who has spent more than 30 years living among and studying the *Pirahã* (pronounced "pee-da-HAN"), a group of about 300 hunter-gatherers, who live alongside a river in a largely unspoiled and remote part of the Amazon rain forest. Everett went there originally to learn the Pirahã language so that he could translate the Bible and spread the gospel to the Pirahã. To do so, he had to overcome the heat, tropical diseases, jaguars, hostile traders, gigantic anacondas,[1] biting insects, snakes that drop from the ceiling, electric eels, piranhas, caimans,[2] a tiny fish that tries to swim up any unguarded body cavity,[3] and much more. You can read about his adventures in the autobiographical book *Don't Sleep, There Are Snakes*. More importantly, for our purposes, you can read about what he discovered about the language that the Pirahã speak, and the ways that it differs from languages that citizens of industrialized nations are more familiar with. It turns out that Everett's research touches on some of the biggest, most general, and most difficult questions that language scientists have attempted to tackle. What does it mean to know a language? How do languages work? Where do they come from? What made languages take their current form(s)? How is language related to thought? Are thought and language identical? This chapter examines these questions, too, not because they have clear answers (most of them do not), but because taking a run at

1

these questions can give us a deeper appreciation of what language is, how it got to be that way, and how our language abilities fit in with other cognitive (thinking) skills.

Part of Everett's research addresses one of the most fundamental questions in language science: What is language? What does it mean to know a language? This is the kind of *essentialist* question that *psycholinguists* (psychologists who study the mental and neural processes as well as the behaviors associated with language) tend to avoid whenever possible (Stanovich, 2009). However, the precise definition of language and a description of its component features greatly concerns researchers who want to know what mental abilities you need to use language, which of those abilities are used for language but not other kinds of cognitive tasks, and whether non-human animals share some or all of our ability to produce and understand language (Everett, 2005, 2007; Hauser, Chomsky, & Fitch, 2002; Jackendoff & Pinker, 2005; Pinker, 1994; Pinker & Bloom, 1990; Pinker & Jackendoff, 2005; Talmy, 2009).

Language Characteristics

Descriptions of language often appeal to Charles Hockett's (1960) design features. Let's focus on a subset of these features, because some of his proposed design features are not necessary for language (e.g., using the vocal channel for sending and receiving messages—sign language users do just fine without it), while others are not specific to language (e.g., cultural transmission—learning to make perogies or knit sweaters is also culturally transmitted). A set of central, possibly necessary, design features could include the following: *semanticity*, *arbitrariness*, *discreteness*, *displacement*, *duality of patterning*, and *generativity*. Let's consider each of these in turn.

Semanticity refers to the idea that language can communicate meaning, and that specific signals can be assigned specific meanings. This occurs at multiple levels in languages, as individual words can be assigned particular meanings, and so can longer expressions that contain more than one word.

Arbitrariness refers to the fact that there is no necessary relationship between actual objects or events in the world and the symbols that a language uses to represent those objects or events. For example, the word that goes with an object need not resemble the real object in any way. One result of arbitrariness is that names for objects can be completely different across languages (*koshka*, *gato*, *chat*, *neko*, and *mao* are all words for *cat*). The name could be changed as long as everyone agreed, and the name change would not affect the ability to express the concept in the language. Tomorrow, we English speakers could all start calling cats "lerps," and as long as everyone agreed, this would work just fine. Sometimes, people point to *onomatopoeia* (words like "moo" and "oink") in English as an example of a non-arbitrary relationship between sound and meaning. Sometimes people argue that the words for large objects have deep-sounding vowels made with the vocal cavity opened up to be big (*ocean*, *tower*), while words for small objects have high-sounding vowels with the vocal cavity closed down to be small (*pin*, *bitsy*). But *onomatopoeia* is not as systematic as people assume (the Dutch equivalent of "oink" is "knorr-knorr"), and there are plenty of counterexamples to the "big concept—big vowel" hypothesis (e.g., *infinity*).

Discreteness refers to the idea that components of the language are organized into a set of distinct categories, with clear-cut boundaries between different categories. For example, every speech sound in English is perceived as belonging to one of about 40 phoneme categories (e.g., a sound is either a /p/ or a /b/; it's either a /t/ or a /d/). For Pirahã speakers, every speech sound made by another Pirahã speaker will be recognized as one of 11 phonemes.[4] Think of how many different speakers a language has, how

different all of their voices are, how their speech can vary from occasion to occasion in how fast they talk, whether they speak clearly or not, and so on. Despite all of the vast differences between speakers, and differences within speakers over time, people who speak the same language will fit every sound made by every speaker into one of the available categories.

Displacement refers to a language's ability to convey information about events happening out of sight of the speaker (*spatial displacement*), about events that happened before the moment when the person speaks, and events that have not yet taken place as the person is speaking (*temporal displacement*). Different languages accomplish displacement in different ways. English has a system of auxiliary verbs (e.g., *will*, *was*, *were*, *had*) and affixes (e.g., *pre-* in *predates*; *-ed* in *dated*) to signal when an event occurred relative to the moment of speaking or relative to other events. Other languages, such as Mandarin, lack these kinds of *tense markers*, but use other means, such as adverbial expressions, to achieve the same means (so you would say the equivalent of, "Yesterday, the man goes" rather than "The man went"). Displacement is a ubiquitous feature of human languages, although the degree and scope of displacement may be more limited in some languages than others (Everett, 2008), but it is largely or completely absent in animal communication systems. Primates may call to one another to signal the presence of predators or food, as will bees, but these behaviors have more the flavor of a reflex, rather than being the result of a controlled, intentional desire to convey information (Tomasello, 2007).

Duality of patterning refers to the fact that we simultaneously perceive language stimuli in different ways; for example, as a collection of phonemes and as a set of words. The word *wasp* consists of four basic speech sounds or *phonemes* – /w/, /o/, /s/, and /p/. Normally, we "see through" the phonemes and the individual word-sounds to the meaning that a speaker is trying to convey, but each of these kinds of patterns, speech sounds (phonemes) and words, can be detected if we decide to pay attention to the form of the speaker's message, rather than its meaning.

Finally, *generativity* refers to the fact that languages have a fixed number of symbols, but a very large and potentially infinite number of messages that can be created by combining those symbols in different patterns. English has about 40 phonemes, but those 40 phonemes can be combined in an infinite number of ways. Similarly, the average high school graduate knows the meanings of about 50,000 different words, but can combine those words in new patterns to produce an unlimited number of meanings.

Language scientists agree that all of the preceding characterize human languages, but they do not all agree on other aspects of language. Many of these disagreements revolve around a component of language called *grammar* (or *syntax* by some theorists). At a very basic level, languages provide us the means to associate sounds with meanings (Hauser et al., 2002). Other animals are also able to associate arbitrary sounds with objects in the environment, similar to the way people associate sounds and meanings. Vervet monkeys make one kind of call when they see an airborne predator, and a different kind of call when they see a predator on the ground; and they respond in the appropriate way depending on which call they hear. If it's an eagle call, they dive into the bushes. If it's a leopard call, they head up into the trees. Vervets lack the capacity to combine sets of calls into longer messages (but see below for evidence that some apes have this ability). If vervets had a system of rules that enabled them to combine calls into more complex messages (e.g., "look at the size of that leopard!"), we would say that they have a *grammar*.

Grammar is one of the two chief components of a language. The other is the *lexicon*, the part of long-term memory that stores information about words (Sag, Wasow, & Bender, 2003). Languages need both of these components so that speakers can formulate messages that express *propositions* (statements of who did what to whom, roughly). To create such messages, a speaker searches for symbols in the lexicon that match the concepts that she

wishes to convey. The grammar tells her how to combine the symbols to create the appropriate signals (speech sounds) that will transmit her message to a listener.

Before we go any further, we need to get straight a common misunderstanding of the word *grammar*. When people hear "grammar," they often think of "grammar school" or the system of rules that your 8th grade English teacher tried to get you to memorize so that you could speak and write standard English. Like me, you probably failed to internalize many of your 8th grade English teacher's lessons. This is partly because 8th grade English is unbearably boring and partly because the principles that your 8th grade teacher was trying to foist on you are completely arbitrary and artificial. For example, Mrs Heidemann tried to get me to believe that you cannot end a sentence with a preposition.[5] But then, there's this kid whose dad always reads him the same story at bedtime. One night, when dad turned up with the same old horrible book, the kid said, *Hey, Dad! What did you bring that book that I didn't want to be read to out of up for?* Five prepositions at the end, perfectly interpretable.[6] Mrs Heidemann was trying to teach me *prescriptive grammar*. Prescriptive grammars are collections of artificial rules. If you follow the grammar teacher's prescription (like you follow a doctor's prescription), your language will sound like that used by members of the upper class in England's home counties.

The vast majority of language scientists are not interested in prescriptive grammar. The kind of grammar we are interested in is <u>descriptive grammar</u>, which is the set of rules or principles that governs the way people use language "in the wild." That is, how people naturally and normally think and behave. Here is an example of a <u>descriptive rule of grammar</u>: "Each clause can only have one main verb." You already know this rule, even though nobody, not even Mrs Heidemann, ever tried to teach it to you. As a result, you would never say, *Mrs Heidemann brewed drank the coffee*. Similarly, English descriptive grammar says, "Put verbs in the middle, not at the beginning of sentences." Again, you already know this rule, because you never say things like *Drank the coffee Mrs Heidemann*. So when this book talks about grammar, remember that it is talking about <u>descriptive</u> grammar (the natural kind) not <u>prescriptive</u> grammar, the Mrs Heidemann kind. Language scientists who study grammar greatly prefer studying descriptive grammar because most of us are interested in the human mind and, as Ivan Sag and colleagues noted (2003, p. 42), "A theory of grammar is a theory about the mental representation of linguistic knowledge."

Descriptive grammars explain why language takes the form that it does. Steven Pinker and Ray Jackendoff (2005) suggest that grammars regulate the combination of symbols into messages in three crucial ways. First, the grammar determines the order that symbols appear in expressions. In English, adjectives come before nouns (*red wine*). In French, the adjectives mostly come after the nouns (*vin rouge*), with a few exceptions (e.g., *grand dame*, "great woman"). Second, the grammar dictates different kinds of *agreement*. Agreement means that certain words in a sentence must appear in a specific form because of the presence of another word in the sentence. In English, we have number agreement (*girls like* but not *girls likes* or *girl like*, as in *Girls like books* but not *Girls likes books*). Other languages have other kinds of agreement, such as Spanish *gender* agreement (*el toro* not *la toro*). Finally, the grammar determines *case marking*, where words must appear in particular forms depending on what grammatical functions they fulfill. English has lost most of its case marking, but it still has some in its system of pronouns (<u>He</u> *left* not <u>Him</u> *left*; *I like* <u>him</u> but not *I like* <u>he</u>). Russian has tons of case marking, as nouns and other words appear in different forms depending on what role they play in the sentence (e.g., vodk<u>a</u> changes to vodk<u>u</u> as the noun moves from subject to object; Водка здесь *Vodka zdes'* "Here is the vodka," but not Водку здесь *Vodku zdes'*; Я пил водку *Ya pil vodku* "I drank vodka," but not Я пил водка *Ya pil vodka*).

To figure out what rules of grammar people actually carry around in their heads with them, linguists spend a great deal of time and effort observing people speaking spontaneously and recording the details of how they combine words into longer expressions. They then

take these records and try to determine why words appear in specific parts of phrases and sentences, and why they appear in particular forms. This type of analysis allows them to deduce the rules behind the patterns that appear in transcripts of speech. When this type of analysis is done on English, it leads to a number of conclusions about English grammar. For example, English is a *subject-verb-object* language. In declarative statements, the grammatical *subject* of the sentence, which is normally the focus of attention or the topic of the discourse, appears at the beginning of the sentence. The verb appears in the middle. The grammatical *object*, which normally is the thing that is acted upon, comes last. Other languages order these elements in different ways. Japanese, for example, puts its verbs at the end. Languages like Russian have free word order and make much greater use than English of different versions of nouns to express who is initiating the action and who is being acted upon. To figure out which system a language has, you actually have to go out and watch people use the language. Sometimes, doing that produces big surprises.

Based on observations of English and other languages, Chomsky and his colleagues have proposed that *recursion* is a core property of the grammars of all languages (Fitch, Hauser, & Chomsky, 2005; Hauser et al., 2002). Further, based on a detailed analysis of human language and animal communication systems, they proposed that recursion is the *only* property that is specific to human language. "The narrow language faculty includes recursion and this is the only uniquely human component of the faculty of language" (Hauser et al., 2002, p. 1569). Chomsky's team proposes that all other properties of language are either shared with non-language thought processes or with non-human communication systems. What are they talking about and why does it matter? *Recursion* is defined as "the ability to place one component inside another component of the same type." So, where language is concerned, recursion could happen if you could place one phrase inside another phrase of the same type or one sentence inside another sentence.[7]

English allows us to place one sentence inside another sentence. Here's a sentence:

Tom likes beans.

We can place that sentence inside another sentence:

Susan thinks (X)   (where X is a sentence)

The result would be:

Susan thinks Tom likes beans.

The degree to which this sort of recursion can go on is essentially infinite, and is limited only by the speaker's ability and willingness to continue:

John knows Dave believes Jenny hopes Carol recognizes Bob realizes … Susan thinks Tom likes beans.

Thus, recursion is one of the characteristics that gives language the property of *discrete infinity*, the ability to generate infinite messages (even infinitely long messages) from finite means.

Most of the languages that have been studied do have recursion, but there does appear to be at least one exception: Pirahã (Everett, 2005, 2008). In English, recursion is often used to create expressions that modify or change the meaning of one of the elements of the sentence. For example, to take the word *nails* and give it a more specific meaning, we could use an *object relative clause* such as *that Dan bought*, as in

Hand me the nails that Dan bought.

In this sentence, the relative clause *that Dan bought* (which could be glossed as "Dan bought the nails") is contained within a larger noun phrase: *the nails (that Dan bought (the nails))*. So the relative clause is nested within a larger phrase, kind of like a stack of bowls. Pirahã expresses the same meaning in a much different form, one that does not involve recursion. To express the meaning that goes with "Hand me the nails that Dan bought," a Pirahã speaker would say the equivalent of:

Give me the nails. Dan bought those very nails. They are the same. (Everett, 2008, p. 227).

In this case, none of the expressions are contained within other expressions of the same type. Pirahã even appears to lack a very simple form of recursion that happens when you use a *coordinate structure* to put two noun phrases together, as in <u>*Dan and Ted*</u> *went to Brazil* (E. Gibson, personal communication). In *Dan and Ted*, you have an overarching noun phrase (of the form *NP and NP*) that contains two separate noun phrases (*Dan, Ted*). To express a meaning like this, a Pirahã speaker would say the equivalent of, "Dan went to Brazil. Ted went to Brazil." Instead of having a stack of bowls, Pirahã has the linguistic equivalent of a string of pearls. All of the statements are connected to each other in an important way, but none of them is contained within any of the others. If recursion does not occur in Pirahã language, which is still definitely a language on a par with other languages in its ability to convey meaning, then recursion is not a necessary characteristic of human languages, despite the fact that most of them have it anyway.

Why does Pirahã lack recursion? Everett's (2008) answer is that Pirahã lacks recursion because recursion introduces statements into a language that do not make direct assertions about the world. When you say, *Give me the nails that Dan bought*, that statement *presupposes* that it is true that Dan bought the nails, but it does not say so outright. In Pirahã, each of the individual sentences is a direct statement or assertion about the world. "Give me the nails" is a command equivalent to "I want the nails" (an assertion about the speaker's mental state). "Dan bought the nails" is a direct assertion of fact, again expressing the speaker's mental state ("I know Dan bought those nails"). "They are the same" is a further statement of fact. Everett describes the Pirahã as being a very literal-minded people. They have no creation myths. They do not tell fictional stories. They do not believe assertions made by others about past events unless the speaker has direct knowledge of the events, or knows someone who does. As a result, they are very resistant to conversion to Christianity, or any other faith that requires belief in things unseen. Everett argues that these cultural principles determine the form of Pirahã grammar. Specifically, because the Pirahã place great store in first-hand knowledge, sentences in the language must be assertions. Nested statements, like relative clauses, require presuppositions (rather than assertions) and are therefore ruled out. If Everett is right about this, then Pirahã grammar is shaped by Pirahã culture. The form their language takes is shaped by their cultural values and the way they relate to one another socially. If this is so, then Everett's study of Pirahã grammar would overturn much of the received wisdom on where grammars come from and why they take the form they do. Which leads us to …

Grammar, Language Origins, and Non-Human Communication Systems

Many language scientists are concerned with the precise definition of language and with detailed descriptions of the grammars of different languages because having those two things nailed down can help us understand how humans think and how we compare with

other living creatures around us. One of the most basic questions that we might like to answer is: Why do humans have language? That question motivates research on the emergence of language abilities in the human evolutionary line as well as research on the language abilities of non-human animals. Figuring out how language abilities developed in the human evolutionary line requires us to analyze the language abilities of ancestors long dead. We need to understand how we are similar to and different from evolutionary ancestors in terms of both language and non-language characteristics. The major obstacle in this line of research is that we have no way of directly observing either the behavior of these ancestors or their nervous systems. As a result, researchers are forced to draw inferences from the fossil record and from artifacts found along with fossil remains. Understanding how we relate to other living animals is potentially easier, because we have living specimens to study. But there are complex issues here as well, some of which are addressed below.

There are two main, overarching ideas about how modern human language abilities emerged from evolutionary ancestors who lacked language, and the same ideas can be used to describe our relationship to living, non-human close relatives (e.g., chimpanzees, bonobos, gorillas, and other primates). These two ideas can be captured by the concepts *continuity* and *discontinuity* (Lenneberg, 1967; Lieberman, 2000; Penn, Holyoak, & Povinelli, 2008). According to the continuity hypothesis, modern human language is quantitatively different from precursor mental abilities, but it is not different in kind or in quality from more basic communication systems. According to the continuity hypothesis, human language abilities are closely related to pre-existing communicative abilities and represent a relatively modest upgrade from those abilities. One advantage of this approach is that we can apply general ideas about adaptation and natural selection to the development of human language, the same way we apply those ideas to other characteristics of humans. The discontinuity hypothesis, by contrast, proposes that some aspects of modern human language abilities represent a clean break from the past, that our language abilities are qualitatively different from more basic communication systems, either in our evolutionary ancestors or in living, non-human animal communication systems. That is, humans possess communication abilities that do not exist in other, more primitive systems. One of the challenges for discontinuity theorists is to identify language abilities that exist in humans but not other species (or in our ancestors), and to explain how the gap between human language abilities and more primitive communication systems was crossed. Let's consider the evolution of human language abilities within our direct ancestors after considering the language abilities of other living modern species (e.g., chimps, dolphins, monkeys).

Research on communication abilities in apes

There are no data which prove that other apes are unable to communicate linguistically.
E. SUE SAVAGE-RUMBAUGH

Studies of how animals communicate can help test ideas related to the continuity hypothesis. As Lenneberg (1967, p. 228) puts it, "[If human] forms of communication … descended from primitive animal forms of communication, [then] a study of the latter is likely to disclose that there is indeed a straight line of evolution of this feature." Apes and monkeys provide useful comparisons to humans because some apes, such as chimpanzees, are closely related biologically to humans. Monkeys and apes are also highly intelligent, which makes them good candidates to share some of the complex abilities that are involved in producing and understanding language. For example, understanding language requires a listener to recognize the meaning or semantic force of an utterance. It turns out that analogous abilities

are present in some species of monkeys. Diana monkeys make different calls for aerial predators and ground predators, as do other species of monkeys (Zuberbühler, 2003; Zuberbühler, Cheney, & Seyfarth, 1999). Are the calls just blindly, instinctively elicited by the sight of the predator, or do monkeys actually assign some meaning to the different calls? If you play to a Diana monkey a recording of another Diana monkey making an alarm call that goes with an aerial predator, like an eagle, and then follow that up with the sound that the eagle itself makes, the Diana monkeys are not surprised. At least, they don't act surprised when the "eagle" monkey call is followed by a stimulus associated with the actual eagle itself. By contrast, if you play the Diana monkey alarm call for "eagle" and then play the sound of a jaguar growling, the little guys go bonkers. It's as if they know "the eagle monkey sound *means* there's an eagle around." So, even though the sound of an eagle is much different than the sound of the "eagle" warning call, behaviorally, Diana monkeys treat the two as equivalent. This ability to treat an arbitrary sound as a "pointer" to an object in the environment is very similar to what people do when they associate an arbitrary collection of sounds (a word) with something else (a concept).

Apes may also make different vocalizations to point to different objects besides predators. For instance, one captive ape (called Kanzi) produces slightly different vocalizations in different contexts. When Kanzi's trainers commented on or asked about bananas, grapes, or juice, Kanzi would often include some kind of vocalization in his response. When those vocal responses were subjected to an acoustic analysis, the vocal response in each context was slightly different (Tagliatela, Savage-Rumbaugh, & Baker, 2003). While it would be premature to say that Kanzi has his own spoken "words" for grape, banana, and juice, the ability to produce different vocal responses in different circumstances is a necessary precursor to spoken language, because spoken language requires us to make different sounds when referring to different concepts.

A number of researchers have attempted to teach language to chimpanzees. The idea was to find out whether human language ability was determined by genetics, or whether it was the result of immersion in cultures where language use was constantly present. If chimps could learn to use language, then human language abilities could not be solely caused by human genetics. In the early days, a chimp named Vicki was trained to make vocal responses to receive rewards. Vicki was never very good at this, primarily because the chimp vocal apparatus is not well configured to make speech sounds, and because chimps do not have good voluntary control over vocalizations (Gardner & Gardner, 1969; Lieberman, 2000). After they figured this out, researchers began to train chimps to use gestures to communicate. Two of the most famous of these animals were Nim Chimpsky and Washoe. Because chimps have much greater voluntary control over gesturing than vocalization, chimps have been much more successful mastering aspects of gestural communication. Members of other species, such as the mountain gorilla Koko, have also learned to communicate in this way (Gardner & Gardner, 1975; Jensvold & Gardner, 2000).

It turns out that when apes learn to sign, they talk mostly about food (see Table 1.1). This makes sense, as chimps like Nim and Washoe were taught to sign using operant conditioning techniques, according to which the chimps were given treats when they produced target behaviors (Premack, 1990). However, according to their human companions, Washoe and Nim's signing went beyond the boundaries of their operant training and showed some of the characteristics of human languages. In one famous example, Washoe was said to make the signs "water" and "bird" to describe a duck that had landed on a pond in her enclosure (Fouts, 1975). This could reflect a generative use of previously learned symbols. That is, Washoe could have been combining signs that were learned to label other concepts in order to label a new concept, the same way you could combine old words to create a new meaning. Similar claims are made regarding bonobos (a species closely related to chimps). Kanzi was credited by his trainers with rudimentary grammatical knowledge on the basis of producing

Table 1.1 The ten most frequent three-sign sequences by Nim Chimpsky (adapted from Terrace et al., 1979, p. 894)

Signs	Frequency
play me Nim	81
eat me Nim	48
eat Nim eat	46
tickle me Nim	44
grape eat Nim	37
banana Nim eat	33
Nim me eat	22
banana eat Nim	26
eat me eat	22
me Nim eat	21

two-symbol strings such as "M&M grape" and "potato oil." However, it is also possible that apes just make signs in a sequence, and are not making the monkey equivalent of compound words. An ape could sign "water" and "bird" in succession, without necessarily intending to apply the combination to refer to the duck.

Chimps were also claimed to have mastered some aspects of grammar, including the ability to interpret *wh-questions* (e.g., *Who is that? What is that?*; Gardner & Gardner, 1975). They are also claimed to observe the basics of word order by producing signs that express specific kinds of meanings in specific places in multi-sign sequences (Terrace, Pettitto, Sanders, & Bever, 1979).[8] For example, when Nim the chimp produced the sign "more" in combination with any other sign, he was much more likely to place the sign "more" first (as in *more banana*) than to use the reverse order (*banana more*). Terrace and colleagues concluded that Nim must have had an internal rule that he applied when constructing sequences of signs containing "more" ("more X" roughly, where X could be replaced by any sign for an object). Similarly, *give X* appeared far more frequently than *X give*. This kind of grammatical knowledge is a vital component of some human languages, including English.

It is sometimes difficult to decide what caused an ape to produce a particular sequence of signs. If you wish to believe that apes have language abilities similar to human children (and for similar reasons), then you would view the following conversation as being highly coherent (from Jensvold & Gardner, 2000):

Washoe: fruit gimme
Trainer: Who funny?
Washoe: Roger
Trainer: Where cat?
Washoe: Roger gimme.
Trainer: Where cat?
Washoe: Gimme

Figure 1.1 Nim Chimpsky signs "me," "hug," and "cat" to his trainer as an increasingly worried tabby (*Felis catus*) looks on (from Terrace et al., 1979, p. 892)

The response to the trainer's first question (Who funny?) could express the idea that the chimp has heard and understood the question, thought about possible answers, and selected Roger as an appropriate funny person. But a skeptical observer might argue that the rest of the exchange casts doubt on this interpretation of Washoe's behavior. Washoe seems to be focused on fruit throughout (perhaps because Washoe has received fruit on previous occasions after producing similar behavior). So the sign "Roger" after "who funny" could just be a false start for the expression "Roger gimme," which could be paraphrased as "Roger give me fruit." In which case, the whole exchange appears to take the form of two ships passing in the night, rather than a coherent conversation.

Operant conditioning techniques succeeded in getting chimps to produce signs, and other training techniques have also been successful in both chimps and the closely related species, bonobo. Sue Savage-Rumbaugh and her colleagues pioneered observational learning techniques as an alternative to operant conditioning (Brakke & Savage-Rumbaugh, 1996a, b; Lyn, Greenfield, & Savage-Rumbaugh, 2006; Lyn & Savage-Rumbaugh, 2000; Savage-Rumbaugh & Fields, 2000; Sevcik & Savage-Rumbaugh, 1994; Shanker, Savage-Rumbaugh, & Taylor, 1999; Tagliatela et al., 2003; Williams & Savage-Rumbaugh, 1997). In observational learning contexts, an animal is exposed to humans who are modeling language behaviors, such as gesturing. The animal may choose to repeat or imitate some of these behaviors, but is not rewarded with food for doing so.

Savage-Rumbaugh adopted a framework that appeals to three main factors to explain why animals (including humans) have whatever communicative abilities they enjoy: biological characteristics of the species (*phylogeny*), maturational characteristics of

the individual (*ontogeny*), and culture or environment. Any, or more likely all, of these features can determine how much skill an individual will have producing and understanding language (e.g., humans are better than fish at using language, 12-year-old humans are better than 12-month-olds, and children in highly interactive households are likely to have greater language abilities than children from less interactive households). Savage-Rumbaugh proposed that some of the limitations in chimp language abilities observed in early studies of non-human communication could have resulted from the fact that training in language started relatively late in the life of individual chimps, the kind of language environment that the chimps were exposed to, and/or the chimp's genetic or biological characteristics.

In an attempt to gather further evidence regarding these possibilities, Savage-Rumbaugh raised a chimp named *Panpanzee* and a bonobo named *Panbanisha*, starting when they were infants, in a language-rich environment. Chimpanzees are the closest species to humans. The last common ancestor of humans and chimpanzees lived between about 5 million and 8 million years ago. Bonobos are physically similar to chimpanzees, although bonobos are a bit smaller on average. Bonobos as a group also have social characteristics that distinguish them from chimpanzees. They tend to show less intra-species aggression and are less dominated by male members of the species.[9] Despite the physical similarities, the two species are biologically distinct. By testing both a chimpanzee and a bonobo, Savage-Rumbaugh could hold environmental factors constant while observing change over time (ontogeny) and differences across the two species (phylogeny). If the two animals acquired the same degree of language skill, this would suggest that cultural or environmental factors have the greatest influence on their language development. Differences between them would most likely reflect phylogenetic biological differences between the two species. Differences in skill over time would most likely reflect ontogenetic or maturational factors.

Rather than reward the developing apes with food in return for signing, adult caregivers modeled language behaviors for the apes.

> *Caregivers communicated to [the apes] with spoken English and visuographic symbols called "lexigrams." The visual symbols were printed on a card, and the animals could point to various symbols that were associated with different concepts. During their training, the apes were exposed to spoken English (which they were not able to emulate), gestures (which they could copy), and lexigrams (which they were also able to use). Because the rearing paradigm stressed observational learning [learning by watching] and emphasized the young apes' language comprehension, [apes] were encouraged to attend to these communications but were not required to produce the symbols themselves in order to receive food or other reward. (Brakke & Savage-Rumbaugh, 1996b, p. 363)*

By comparing how much her chimp and her bonobo learned to how much operant-trained chimps learned in previous studies, Savage-Rumbaugh could estimate the effects of cultural/environmental factors on language learning in apes.

Over the course of the study, which lasted for just under four years, the apes developed communication skills using both gesture and the lexigrams (Brakke & Savage-Rumbaugh, 1996a, b). Communication via gesture developed before attempts to use lexigrams in both apes, and the chimpanzee continued to rely exclusively on gesture for a whole year after the bonobo had started to use lexigrams. Panpanzee the chimp did appear to imitate her trainers' use of the lexigrams, but she did not use them spontaneously. Panpanzee appeared more likely than Panbanisha the bonobo to combine using the lexigrams with gesturing throughout the study period, and the chimp was about 50% more likely to combine gesturing and pointing to lexigrams when she interacted with her trainers. Overall, the chimp produced fewer "words" during the study period. Because the chimp and the bonobo were both reared using the same methods, under essentially identical environmental

conditions, differences between the chimp and the bonobo are not likely to result from differences in the environment, but could be caused by biological/genetic differences between the species.[10]

Savage-Rumbaugh reports that, among the animals exposed to enriched language environments from infancy, four have acquired receptive vocabularies of 500 words or more, with productive vocabularies of 150 words or more. Further, she reports that bonobos raised in a language-enriched environment (including Kanzi and his half-sister Mulika) appear to use symbols more spontaneously than chimps raised under operant-learning conditions (who tend to sign mostly in human-initiated exchanges). If so, the immersion methods that Savage-Rumbaugh used to teach her animals may be responsible for the greater spontaneity of their signing behavior.

"Monkeys don't talk"

Chimpanzee signing should not be labeled linguistic.
ESTEBAN RIVAS

Keep your stinking paws off me, you damned, dirty ape.
CHARLTON HESTON, *PLANET OF THE APES*

Some researchers interpret ape signing behavior as being highly consistent with the linguistic behavior of young children, but in some ways, the behavior of signing apes differs greatly from the language-related behavior of young children. First and foremost, the acquisition of language-related (or language-like) behaviors in apes varies widely from one animal to the next. In contrast to children, who universally acquire a native language given normal brain function, a stable environment, and exposure to a model, some apes acquire the ability to interpret symbols and use them to communicate, and some do not, even when they are exposed to the same models (see, e.g., the difference between Kanzi and Matata; Sevcik & Savage-Rumbaugh, 1994). This fact, by itself, could indicate that apes are using different mechanisms than humans to acquire language skills. Furthermore, children do much more than copy the behaviors of their adult caregivers. Children actively experiment with the language (as when infants babble) and develop knowledge of the sound system of language before they begin to produce their first words (analogous to ape signs or lexigrams). Such prelinguistic babbling behavior has not been reported in apes, suggesting that the mechanisms of acquisition and development are different in humans than in apes.

The acquisition and use of grammar also appears to work differently in children than in apes. When children produce multi-word utterances, their longer utterances contain elements of their shorter utterances, but they also contain new elements. Repetition of elements within utterances is almost unheard of in child language, but it is common in the signing of apes. For example, some of Nim's sign sequences include expressions like "eat Nim eat Nim," and "banana me eat banana" (Terrace et al., 1979, p. 894). In fact, repetition appears to be a major mechanism contributing to Nim's (and other chimps') longer utterances (e.g., "give orange me give eat orange me eat orange give me eat orange give me you," Terrace et al., 1979, p. 895; Rivas, 2005). Repetition occurs in over 90% of some apes' (e.g., Koko's) signing behavior. These repetitious sequences resemble the "superstitious" behaviors produced by pigeons, dogs, and college undergraduates when there is a contingency between behavior and reward, but when the delivery of a reward is delayed (Bruner & Revusky, 1961). Imagine that Washoe gets rewarded for producing signs. Imagine that some time elapses between the time when Washoe makes a sign and the time when he gets his reward. During that interval, additional signs may be produced. Washoe may

come to "believe" that a sequence, rather than a particular sign, caused the reward to appear. Alternatively, intermittent reinforcement increases the rate of behaviors in animals generally, so if apes are reinforced on a variable schedule, they will tend to produce longer and longer sets of signs, not because they have acquired more complex grammatical rules, but simply because external rewards delivered on variable schedules draw more behaviors out of animals.[11]

Critical observers of ape language studies have also suggested that the way apes use signs is different from the way humans use words (Rivas, 2005; Seidenberg & Pettito, 1987; Tomasello, 2007). Humans use words to express *intentions* (ideas behind or motivations for speaking), while apes' use of symbols seems much shallower and less intentional. Humans also commonly use words to draw attention to objects or events in the environment, or to comment on those objects and events. Apes most commonly make signs in order to get something. To put it in more neutral terms, most ape signs are associated with objects (fruit, juice, M&Ms) and actions (tickling, chasing) that are rewarding to apes. As Tomasello (2007, p. 152) notes, "Most if not all ape gestures are imperative, intended to influence the behavior of others directly, whereas many human gestures are used for declarative or informative purposes." It is difficult, therefore, to determine the extent to which apes' signing behavior is maintained by extrinsic reward, and how much reflects an intention to communicate particular thoughts. This orientation toward reward appears to be present even in apes who were trained using observational methods. For example, although Kanzi the bonobo was trained using observational learning methods, rather than direct operant reward, his pattern of sign use closely matches that of apes that were trained using more traditional operant conditioning techniques. As many as 96% of Kanzi's signs can be interpreted as requests (Greenfield & Savage-Rumbaugh, 1990), consistent with the idea that much of his signing behavior is maintained by reward or the prospect of reward. Brakke and Savage-Rumbaugh (1996b, p. 365) report similar patterns of request-like behavior for the chimp *Panpanzee* and the bonobo *Panbanisha*, who were reared in a similar fashion to Kanzi.[12]

Apes also appear to apply such grammatical rules as they may have much less consistently than humans. For example, although Washoe signs "more X" far more often than he signs "X more," the difference is not as high as it should be if he were truly applying a grammatical rule. In languages like English, when a grammatical rule is in place, the related behavior is almost 100% consistent. An English-speaking child would *always* say, *I want more juice*, and would never say, *I want juice more*. There are limited exceptions to consistent application of grammatical rules, but these typically occur when the grammar offers two or more ways to express the same idea. For example, you can say, *Give Mary more juice* (which is called a *ditransitive* sentence) or you could say the equivalent *Give more juice to Mary* (which is called a *dative* sentence). An individual speaker may flip back and forth between these two options, but that is because that individual has two grammatical rules for how to form sentences involving the transfer of objects from one person to another (*X verb Y to Z, and X verb Z Y*, roughly). Different rules get triggered on different occasions, but once a rule gets triggered, it is followed to the letter (so you would never get a sentence like *Give to more juice Mary* or *John Mary give more juice*).[13] More recent research shows that, although non-human primates can learn some of the patterns characteristic of human language grammars, other patterns appear to be beyond their grasp (Hauser, Newport, & Aslin, 2001; Newport, Hauser, Spaepen, & Aslin, 2004; Ramus, Hauser, Miller, Morris, & Mehler, 2000).

Apes and people also differ greatly in the way they take turns during interactions. In dialogue, different cultures differ somewhat in the degree to which one speaker's utterances will overlap with another's, but interruptions are relatively infrequent (they certainly do happen, but the interruption is often marked as such by the interrupter). Apes interrupt people all of the time, usually to ask them for food (Terrace et al., 1979). This ubiquity of interruptions suggests that one of the prerequisites for full-blown language to emerge is the

ability of individuals to engage in impulse control. If individuals in a communicative exchange are not able to control the impulse to vocalize, that is to coordinate their verbal behavior with other members of the group, the result is a cacophony of overlapping sounds, making it difficult or impossible for any individual to be heard over the general din (think of a seal colony, or the internet).

To summarize, although apes display some behaviors that humans do, including producing signs to refer to specific objects and events, there are substantial differences between the ways humans and apes use language (although see Lieberman, 2000, for a vigorous defense of similarities between ape and human language abilities). Whether you view these differences as being fatal to the continuity hypothesis depends on how you view the relationship between ape language abilities and human language abilities. If you view the ape abilities as being more primitive versions of the human abilities, then the continuity hypothesis wins. If you view the differences between humans and apes as being so great that human behaviors must be generated by an entirely different set of mental processes, then the discontinuity hypothesis wins.

Language origins

Studying ape communication is a way to investigate why humans have the language abilities that they enjoy. Abilities that apes demonstrate are likely to have been present in the last common ancestor of apes and humans. It is possible, but less likely, that shared abilities of humans and apes developed independently after the two species split between 5 and 8 million years ago. Other means to study language origins involve comparing modern humans to our evolutionary ancestors. By assessing human ancestors' physical features and artifacts, theorists can develop ideas about how and when modern language abilities emerged. Much of this work is speculative, because the critical evidence needed to discriminate between different theories is unavailable. As with research that compares living humans and living apes, much of the theorizing in the evolutionary approach to language origins revolves around the concepts of continuity and discontinuity. Continuity theory views modern humans' language abilities as reflecting modifications of abilities that existed in our ancestors. Discontinuity theory views modern humans' language abilities as being distinct and separate from our ancestors' abilities.

One thing that all language scientists agree on is that the human capacity for speech is an *adaptation*, in the Darwinian sense. An adaptation is a characteristic of a species that has been selected for and maintained by environmental factors. That is, at some point in the past, there was variability within the species such that some individuals had more of the relevant characteristic than others did. Those individuals who had more of the characteristic were more likely to survive and reproduce (they were biologically more fit), and so individuals without the characteristic in question became scarcer and scarcer in the population until their numbers declined to zero. Speech is viewed as an adaptation in part because of its obvious advantages—it allows for the near instantaneous sharing of complex knowledge across individuals and the coordination of joint activities—but also in part because of its less obvious disadvantages (Aitchison, 2000; Darwin, 1859/1979). To produce a wide range of speech sounds, the larynx needs to be deeper in the throat in humans than it is in other species (including in our close relatives, chimpanzees and bonobos). As a result of its position deep in the throat, humans are not able to close off their airway when they eat. That makes it more likely that humans will choke to death accidentally. According to the Centers for Disease Control, about 150 children in the United States die each year from foreign objects lodging in their windpipes. Other animals do not choke as frequently as people do, because they can close off their airways when they eat. The fact that the descended

larynx persists is evidence that this characteristic is maintained because of its selective advantages, in spite of its potential drawbacks.

It is generally agreed that modern language abilities, most specifically the ability to produce spoken language, would not have been possible without two modern human characteristics: (a) A vocal apparatus that allows for the production of a variety of distinct speech sounds (phonemes); and (b) The ability to exert a high degree of very fine control over that vocal apparatus (Lieberman, 2000; MacLarnon & Hewitt, 1999).

When, exactly, did modern language emerge? We can only speculate. Because our ancestors lacked the kind of control necessary for speech (and if we discount the possibility that language was developed as a gestural system before it was developed as a vocal system; see Falk 2004, and replies therein; Hewes, 1973), it is likely that complex, modern language emerged some time during the last 200,000 to 70,000 years. While it is possible that the same processes caused language to appear at the same time as other features of *Homo sapiens*, it is also possible that the cultural and artistic revolution that occurred approximately 50,000 years ago coincided with the emergence of fully modern languages.

Philip Lieberman (2000) argues that human ancestors (e.g., *Homo erectus*) had the ability to speak, although their speech would not have been as refined as modern humans' speech. This conclusion is based on reasoning about why the human vocal tract has the shape it does. Lieberman notes that to produce vowel sounds such as /i/ (as in *meet*) and /u/ (as in *you*), the space above the larynx in the throat has to be about the same length as the horizontal space between the top of the throat and the mouth opening. For natural selection to produce and maintain this configuration, Lieberman argues, some rudimentary speech abilities must have been present beforehand. Natural selection could then have favored individuals who had physical characteristics that allowed them to produce a wider range of vowel sounds. Unless some rudimentary speech abilities were present prior to the advent of *Homo sapiens*, a lowered larynx, and the accompanying ability to produce more vowel sounds, would have to be the result of a massive and incredibly lucky mutation, rather than gradual evolution by natural selection.

Other researchers view speech as being absent until the advent of *Homo sapiens*. Speaking is an exercise in controlled exhalation. Rather than breathing out smoothly in one continuous motion, speaking requires us to rapidly change the flow of air out of the lungs in order to control fine aspects of speech, such as how much emphasis or stress (*accent*) we place on each word and syllable.[14] MacLarnon and Hewitt argue that speech could not have been present in our ancestors, because, like modern apes, our evolutionary ancestors lacked the ability to exert this fine degree of control, as evidenced by the relatively small diameter of the nerves that lead from the brain to the relevant upper-torso and throat muscles (MacLarnon & Hewitt, 1999). When early *Homo sapiens* fossils were compared to *Australopithecus afarensis*, *Homo erectus*, and *Neanderthal* specimens, only early *Homo sapiens* had the kinds of nerve tracts that are associated with modern humans. This enhancement of breathing control not only increases the range of speech sounds that people can produce, it also increases the absolute amount of time they can speak without stopping to catch their breath. Non-human vocalizations in primates are limited to about 5 seconds. Humans can go on for upwards of 10 seconds without stopping for breath.[15]

The fossil record shows that human ancestors before *Homo sapiens* emerged, between about 70,000 and 200,000 years ago, had some of the cultural and physical characteristics of modern humans, including making tools and cooking food. If we assume that modern language emerged sometime during the *Homo sapiens* era, then it would be nice to know why it emerged then, and not before. One possibility is that a general increase in brain size relative to body weight in *Homo sapiens* led to an increase in general intelligence, and this increase in general intelligence triggered a language revolution. On this account, big brain comes first and language emerges later. This hypothesis leaves a number of questions

unanswered, however, such as, what was that big brain doing before language emerged? If the answer is "not that much," then why was large brain size maintained in the species (especially when you consider that the brain demands a huge proportion of the body's resources)? And if language is an optional feature of big, *sapiens* brains, why is it a universal characteristic among all living humans? Also, why do some groups of humans who have smaller sized brains nonetheless have fully developed language abilities?

Another line of thought starts with the idea that word-like units had to be present before more complex sequences of words emerged (Aitchison, 2000). Before you begin to use words, you have to appreciate the idea that sounds can be associated with objects, the *naming insight*. But where did that naming insight develop? One possibility is that it was an extension of more primitive verbal systems. Non-human primates already have some aspects of semantics (meaning) in their call systems, using alarm calls to activate knowledge of particular kinds of animals, rather than treating them as mere noise. Candidates for the first word-like units in human languages include noises imitating predators or prey, grunting noises made in concert with physical exertion, or the equivalents of lip-smacks and hooting that apes make as greetings. The theory is that you would need some number of such *proto-words* before the language could develop a system of speech sounds, and you need a system of speech sounds before you could synthesize a larger set of words to express a wider range of concepts. Unfortunately, no existing data indicate which, if any, of these candidates gave rise to the first words. It is likely, however, that once the naming insight took hold, that the stock of words would have expanded rapidly.

The next step in language evolution could have been the development of something approximating a modern pidgin. Pidgins develop when adult speakers of different languages are placed in circumstances that require them to communicate (Bickerton, 1988). Pidgins are generally simpler than full-blown languages, with a restricted vocabulary and rudimentary grammar. For example, the following expressions are used in *Tok Pisin*, a pidgin used in Papua New Guinea (Aitchison, 2000, p. 124):

han bilong diwai	"branch of a tree"
han bilong pik	"front legs of a pig"
han bilong pisin	"bird's wing"

In a creole (a language that emerges from the combination of two or more pre-existing languages) or another kind of full-blown language, we would expect concepts such as branch, legs, and wings to have shorter names; and the language would not depend upon the extension of one term *hand* to cover a wide variety of objects (although there may be advantages to making the similarity between hands, branches, legs, and wings explicit). Pidgins tend to lack the grammatical features of true languages, including markers for past and present tense on verbs, number agreement between subjects and verbs, sets of distinct prepositions (*on, of, below*, etc.), and case marking (e.g., changes in the form of nouns depending on their position within a sentence or their semantic role). Thus, pidgins represent an intermediate form between having no grammar at all, and having the kind of complex grammar that is characteristic of all true languages.

Some theorists suggest that grammar is the only thing that distinguishes human language abilities from those of our ancestors and those of our living relatives (e.g., the apes). If so, where did grammar come from? One possibility is suggested by the language bioprogram hypothesis (Bickerton, 1988; Lenneberg, 1967; Pinker, 1994). According to the language bioprogram hypothesis, human genetics creates the mental equivalent of the heart, stomach, lungs, or other organ. Like the heart, which depends on other organs for support, the language organ relies on other mental abilities for support. But also like the heart, which specializes in taking deoxygenated blood from the body, passing it to the lungs, and

recirculating oxygenated blood back to the body, the language organ also specializes. In particular, it specializes in building the hierarchical, symbolic representations that underlie the ability to speak and the ability to understand speech. Is there any evidence that genetics contributes to human language abilities, and grammar in particular? (And remember, we're talking about *real*, descriptive grammar, not Mrs Heidemann's prescriptive grammar.)

Evidence for a genetic origin of grammar comes from two chief sources:[16] Studies of *creoles* (full-blown languages that develop out of pidgins) and studies of genetic anomalies that are associated with language disorders. If genetics contributes to our language abilities, then we would expect fully grammatical language to develop in children, whether they are exposed to a fully grammatical model language or not. Some studies have detected just such a pattern (Bickerton, 1988). In these studies, children who grow up listening to their parents speak pidgin appear to spontaneously add grammatical markers, such as case, tense, and agreement features, and wind up speaking a version of the language that is qualitatively different from the pidgin that their parents speak. Some of this research has been criticized because it relied on retrospective (historical) reports of elderly people talking about events from their childhood and because the individuals in question were being exposed to fully grammatical languages when their parents spoke their native language at home. However, more recent work on Nicaraguan Sign Language (LSN) is less subject to these criticisms (Emmorey, 2002; Senghas & Coppola, 2001).

In Nicaragua before the Sandinista revolution, children who were deaf tended to grow up isolated from one another and the vast majority were born to parents who could hear and speak, but could not sign. Thus, deaf children did not have a sign language model to follow. Most of them developed systems of "home sign," idiosyncratic systems of gesture that allowed them to communicate with their caregivers and families. Starting in 1977, a central school for the deaf was established, which brought together deaf children from all over the country for the first time in Nicaragua. Children were taught officially using oral methods (speaking and lip-reading training), but they communicated with each other in their spare time using gestures. Early on, the deaf children's system of gestures developed a standard, shared vocabulary, but the system lacked many of the grammatical features of full-blown sign languages. However, as younger children were added to the mix, they spontaneously added grammatical features found in other sign languages. This was accomplished even though the deaf children were not exposed to an adult language model. One way to explain this phenomenon is to propose that children's genetic heritage provides them with the mental tools they need to "invent" a grammatical system, as well as the drive to implement such a system if it is not already present in their environment.

Studies of individuals with *selective language impairment* (SLI) also suggest a genetic contribution to language abilities (Enard et al., 2002; Gopnik, 1990; 1994; Gopnik & Crago, 1991). One set of studies focused on a particular family living in England (the *KE family*). Half of the members of this family appear to be entirely normal in their general intelligence and their language abilities. The other half of the family also appear to have essentially normal intellectual abilities, but they have a number of problems producing and understanding language. For example, the affected members of the family have difficulty with past tense endings on verbs. They are likely to speak sentences such as "Yesterday he walks," or "After thinking about it for a while, she finally jump and fell." In writing sentences, the affected individuals do produce the correct tense endings some of the time, but generally this happens in response to explicit instruction on individual verbs. That is, they do not appear to apply the general rule "To make the past tense, add *-ed* to the end." Other verb-related markers are also not applied, or applied incorrectly (as in "Carol is cry in the church."). Affected individuals also have problems making plural nouns out of singular nouns. In the *wug test*, people are given nonsense words, like *wug*, and *zat*, and are asked to make them plural. "Here is one zat. Now there are two _____." When asked to fill in the

blank, one family member said, "zacko." Genetic testing revealed that the affected members of the family all have an unusual form of the *FOXP2* gene, while the unaffected members have the more common form (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001).

Although the deficit shown in the KE family has been attributed to grammar, specifically, some researchers favor an explanation under which the family suffers from a more general problem planning and executing sequences of behaviors (Vargha-Khadem et al., 1998; Watkins, Dronkers, & Vargha-Khadem, 2002). This latter interpretation helps explain why affected family members have difficulty moving facial muscles on command as well as problems repeating single words clearly. So, either the gene is directly responsible for instilling some aspects of grammar as an innate feature of human cognitive abilities (as per the genetic bioprogram hypothesis; Bickerton, 1988; Pinker, 1994) or the gene affects those parts of the brain which normally are involved in planning sequences of behavior, and our language production and comprehension processes normally tap into those resources as we speak and interpret language.

Language and Thought

The throat motor segment thus becomes the controlling segment of the body.
J. B. WATSON

You may have had the experience that when you are thinking about something or planning some kind of activity that you have a voice in your head, probably sounding much like your own voice, that is talking to you about whatever it is you are thinking about. This inner monologue is such a common experience when you are thinking that you might believe that "talking to yourself" and thinking are one and the same thing. You would be in good company if you thought this was how cognition worked, as the famed behaviorists J. B. Watson and B. F. Skinner both advocated versions of this idea (Skinner, 1957; Watson, 1924). Watson asked and answered the question in this way (1924, pp. 341, 347): "Do we think only in words, that is, in verbal motor contractions? My own answer has been: Yes … 'thinking' is largely sub-vocal talking." However, since Watson's time, plenty of evidence has accumulated to show that thinking and language are separate, though linked, abilities.

The beginning of the end for Watson's theory of language and thought came in 1947, when medical doctors strapped a healthy 34-year-old research volunteer to a gurney in Salt Lake City and injected him with curare (Smith, Brown, Toman, & Goodman, 1947; see Table 1.2). Curare paralyzes the muscles of the body completely, including the *pharyngeal* (throat) muscles that Watson believed were critical for thought processes. About four minutes after the curare injection was completed, the research volunteer lost the ability to speak because his throat muscles could no longer move. Despite the loss of speech, the volunteer could still perceive everything that was happening around him. After he recovered, the volunteer reported that he was "clear as a bell" during the entire time he was paralyzed. In addition, during the time that he could not speak, the volunteer answered yes-or-no questions by moving muscles that had not yet become completely paralyzed, including his eyebrow and eyelid muscles. His answers to these questions were "entirely correct." Based on the events during the experiment, and on the volunteer's subsequent description of his experience, Smith's research team concluded that de-activating the speech muscles had no effect on the volunteer's ability to perceive, think about, or remember, the events that occurred during total muscular paralysis.

The curare results are fatal for Watson's idea that thinking and moving the throat muscles, whether overtly or covertly, are the same thing, but he could still be right if "talking to